Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 18, No. 3, pp. 661–677, 2022

CONVERGENCE RATE OF THE ACCELERATED MODIFIED LEVENBERG-MARQUARDT METHOD UNDER HÖLDERIAN LOCAL ERROR BOUND

Haiyan Wang

Abstract: In this paper, we study the convergence properties of the accelerated modified Levenberg-Marquardt (LM) method presented in [Mathematics of Computation, **83** (2014), 1173-1187] under the conditions of 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.

Key words: *nonlinear equations, Levenberg-Marquardt method, Hölderian local error bound, convergence rate*

Mathematics Subject Classification: *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. It is natural to transform (1.1) to the nonlinear least squares problem

$$\min_{x \in R^n} \Phi(x) = \|F(x)\|^2. \tag{1.2}$$

Obviously, (1.1) has a solution if and only if the minimal value of (1.2) is zero. We assume that the solution set of (1.1), denoted by X^*, is nonempty. In this paper, $\|\cdot\|$ refers to the 2-norm both for vectors and matrices.

The Levenberg-Marquardt (LM) method is one of the most well-known methods for solving (1.2). At each iteration, it computes the step

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

where $F_k = F(x_k)$, $J_k = F'(x_k)$ is the Jacobian and the LM parameter $\lambda_k \geq 0$ is introduced to overcome the singularity or near singularity of $J_k^T J_k$. The LM method has quadratic convergence if the Jacobian $J(x)$ is Lipschitz continuous and nonsingular at the solution of (1.2) and if λ_k is chosen properly.

However, the condition of nonsingularity is too strong. Under the local error bound condition, which is weaker than the nonsingularity, Yamashita and Fukushima took $\lambda_k = \|F_k\|^2$ and proved that the LM method still converges quadratically [15]. More generally, Fan and Yuan took $\lambda_k = \|F_k\|^\alpha$ and proved by a different approach that the LM method has quadratic convergence for all $\alpha \in [1,2]$ (cf. [6]). In [3], Fan proposed the modified LM method, which computes not only a LM step but also an approximate LM step at each iteration. Furthermore, Fan presented the accelerated LM method, which performs the line search along the approximate LM step, and proved that the convergence order of the accelerated LM method with order $\min\{1+\alpha, 3\}$ under the local error bound condition and the Lipschitz continuity of the Jacobian [4]. Recently, Fan et al. proposed the adaptive LM method, which can decide automatically whether an iteration should evaluate the Jacobian matrix to compute an LM step, or use the latest evaluated Jacobian to compute an approximate LM step [5].

In real applications, some nonlinear equations may not satisfy the local error bound condition, but satisfy the more general Hölderian local error bound condition.

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

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

Obviously, when $\gamma = 1$, the Hölderian local error bound condition is reduced to the local error bound condition. Wang and Fan considered the convergence rate of the LM method under the Hölderian local error bound condition and the Hölderian continuity of the Jacobian [12]. The convergence properties of the inexact Levenberg-Marquardt method under the same conditions are also given in [13].

In this paper, we study the convergence properties of the accelerated Levenberg-Marquardt method proposed in [4] under the Hölderian local error bound condition and the Hölderian continuity of the Jacobian. At every iteration, it first solves the linear equations

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F_k \quad \text{with} \quad \lambda_k = \mu_k \|F_k\|^\delta, \quad \delta > 0 \tag{1.5}$$

to obtain the LM step d_k, where $\mu_k > 0$ is updated from iteration to iteration, then it solves

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F(y_k) \quad \text{with} \quad y_k = x_k + d_k \tag{1.6}$$

to get the approximate LM step $\hat{d}_k$. After that, it performs a line search along $\hat{d}_k$ and define the trial step

$$s_k = d_k + \alpha_k \hat{d}_k, \tag{1.7}$$

where α_k is the step size for $\hat{d}_k$.

This paper is organized as follows. In Section 2, we show that the accelerated LM method converges globally under the Hölderian continuity of the Jacobian. In Section 3, we study the convergence order of the algorithm under the Hölderian local error bound condition and the Hölderian continuity of the Jacobian. Finally, we conclude the paper in section 4.

2 Global Convergence of the Accelerated LM algorithm Under the Hölderian Continuity of the Jacobian

In this section, we first present the modified LM algorithm, then prove it converges globally under the Hölderian continuity of the Jacobian.

Note that the step d_k given by (1.5) is the minimizer of the problem:

$$\min_{d\in R^n} \|F_k + J_k d\|^2 + \lambda_k \|d\|^2 \triangleq \varphi_k(d). \tag{2.1}$$

Define

$$\Delta_k = \|d_k\| = \| -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k\|. \tag{2.2}$$

Then d_k is also a solution of the trust region problem:

$$\min_{d\in R^n} \|F_k + J_k d\|^2, \quad s.t.\ \|d\| \le \Delta_k. \tag{2.3}$$

Following Powell's result given in [11], we have

$$\|F_k\|^2 - \|F_k + J_k d_k\|^2 \ge \|J_k^T F_k\| \ \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}. \tag{2.4}$$

Similarly, $\hat{d}_k$ is not only the minimizer of the problem:

$$\min_{d\in R^n} \|F(y_k) + J_k d\|^2 + \lambda_k \|d\|^2 \triangleq \varphi_{k,1}(d), \tag{2.5}$$

but also a solution of the trust region problem:

$$\min_{d\in R^n} \|F(y_k) + J_k d\|^2, \quad s.t.\ \|d\| \le \Delta_{k,1} = \|\hat{d}_k\| = \| -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k)\|. \tag{2.6}$$

Hence, we have

$$\|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 \ge \|J_k^T F(y_k)\| \ \min\left\{\|\hat{d}_k\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|}\right\}. \tag{2.7}$$

When $\hat{d}_k$ is a decreasing direction of the merit function $\Phi(x)$ at y_k, we perform a line search along $\hat{d}_k$. That is, we solve the problem

$$\min_{\alpha>0} \left\| F\left(y_k + \alpha \hat{d}_k\right)\right\|^2. \tag{2.8}$$

We approximate it by

$$\min_{\alpha>0} \left\| F(y_k) + \alpha J_k \hat{d}_k \right\|^2, \tag{2.9}$$

which is equivalent to the problem

$$\max_{\alpha>0} \|F(y_k)\|^2 - \left\|F(y_k) + \alpha J_k \hat{d}_k\right\|^2 \triangleq \phi(\alpha), \tag{2.10}$$

where

$$\phi(\alpha) = -\hat{d}_k^T J_k^T J_k \hat{d}_k \alpha^2 + 2\hat{d}_k^T \left(J_k^T J_k + \lambda_k I\right) \hat{d}_k \alpha. \tag{2.11}$$

The maximizer of $\phi(\alpha)$ is

$$\tilde{\alpha}_k = \frac{\hat{d}_k^T \left(J_k^T J_k + \lambda_k I\right) \hat{d}_k}{\hat{d}_k^T J_k^T J_k \hat{d}_k} = 1 + \frac{\lambda_k \hat{d}_k^T \hat{d}_k}{\hat{d}_k^T J_k^T J_k \hat{d}_k} > 1 \tag{2.12}$$

if $J_k\hat{d}_k \neq 0$. Since the value of $\tilde{\alpha}_k$ could be very large when $J_k\hat{d}_k$ is close to 0, we give $\tilde{\alpha}_k$ an upper bound and compute the step size by solving

$$\max_{\alpha\in[1,\hat{\alpha}]} \|F(y_k)\|^2 - \left\|F(y_k) + \alpha J_k\hat{d}_k\right\|^2, \tag{2.13}$$

where $\hat{\alpha}$ is a positive constant.

Define the actual reduction of the merit function $\Phi(x)$ at the k-th iteration as

$$Ared_k = \|F_k\|^2 - \|F(x_k+s_k)\|^2 \tag{2.14}$$

and the predicted reduction as

$$Pred_k = \|F_k\|^2 - \|F_k + J_kd_k\|^2 + \|F(y_k)\|^2 - \|F(y_k) + \alpha_kJ_k\hat{d}_k\|^2. \tag{2.15}$$

By (2.4), (2.7) and (2.9)–(2.12),

$$\begin{aligned}
Pred_k &= \|F_k\|^2 - \|F_k + J_kd_k\|^2 + \|F(y_k)\|^2 - \left\|F(y_k) + \alpha_kJ_k\hat{d}_k\right\|^2 \\
&\geq \|F_k\|^2 - \|F_k + J_kd_k\|^2 + \|F(y_k)\|^2 - \left\|F(y_k) + J_k\hat{d}_k\right\|^2 \\
&\geq \|J_k^TF_k\| \ \min\left\{\|d_k\|, \frac{\|J_k^TF_k\|}{\|J_k^TJ_k\|}\right\} + \|J_k^TF(y_k)\| \ \min\left\{\|\hat{d}_k\|, \frac{\|J_k^TF(y_k)\|}{\|J_k^TJ_k\|}\right\} \\
&\geq \|J_k^TF_k\| \ \min\left\{\|d_k\|, \frac{\|J_k^TF_k\|}{\|J_k^TJ_k\|}\right\}.
\end{aligned} \tag{2.16}$$

The ratio of the actual reduction to the predicted reduction

$$r_k = \frac{Ared_k}{Pred_k} \tag{2.17}$$

plays a key role in deciding whether to accept the trial step and how to update the LM parameter λ_k.

The algorithm is presented as follows.

Algorithm 2.1. Given $x_1 \in R^n, \delta > 0, \mu_1 \geq \mu_0 > 0, 0 < p_0 < p_1 < p_2 < 1, a_1 > 1 > a_2 > 0$. Set $k := 1$.

Step 1. If $\|J_k^TF_k\| = 0$, then stop; compute

$$\lambda_k = \mu_k\|F_k\|^\delta.$$

Solve

$$(J_k^TJ_k + \lambda_kI)d = -J_k^TF_k \tag{2.18}$$

to obtain d_k; set

$$y_k = x_k + d_k;$$

solve

$$(J_k^TJ_k + \lambda_kI)d = -J_k^TF(y_k) \tag{2.19}$$

to obtain $\hat{d}_k$; set

$$s_k = d_k + \alpha_k\hat{d}_k. \tag{2.20}$$

Step 2. Compute $r_k = \frac{Ared_k}{Pred_k}$; set

$$x_{k+1} = \begin{cases} x_k + s_k, & if\ r_k \geq p_0, \\ x_k, & otherwise. \end{cases} \tag{2.21}$$

Step 3. Compute

$$\mu_{k+1} = \begin{cases} a_1\mu_k, & if\ r_k < p_1, \\ \mu_k, & if\ r_k \in [p_1, p_2], \\ \max\{a_2\mu_k, \mu_0\}, & otherwise; \end{cases} \tag{2.22}$$

set $k := k+1$; go to Step 1.

Assumption 2.1. (a) The Jacobian $J(x)$ is Hölderian continuous of order $v \in (0,1]$, i.e., there exists a positive constant κ_{hj} such that

$$\|J(x) - J(y)\| \leq \kappa_{hj}\|x - y\|^v, \quad \forall x, y \in R^n. \tag{2.23}$$

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

$$\|J(x)\| \leq \kappa_{bj}, \quad \forall x \in R^n. \tag{2.24}$$

By (2.23), we have

$$\begin{aligned} \|F(y) - F(x) - J(x)(y-x)\| &= \left\| \int_0^1 J(x + t(y-x))(y-x)dt - J(x)(y-x) \right\| \\ &\leq \|y - x\| \int_0^1 \|J(x + t(y-x)) - J(x)\| dt \\ &\leq \kappa_{hj}\|y - x\|^{1+v} \int_0^1 t^v dt \\ &= \frac{\kappa_{hj}}{1+v}\|y - x\|^{1+v}. \end{aligned} \tag{2.25}$$

Theorem 2.2. *Under Assumption 2.1, the sequence* $\{x_k\}$ *generated by Algorithm 2.1 satisfies*

$$\liminf_{k \to +\infty} \|J_k^T F_k\| = 0. \tag{2.26}$$

Proof. Suppose that (2.26) is not true. Then there exists a positive constant τ such that

$$\|J_k^T F_k\| \geq \tau, \quad \forall k. \tag{2.27}$$

Define the index set of successful iterations as

$$S = \{k \mid r_k \geq p_0\}. \tag{2.28}$$

We derive the contradictions in two cases.

Case 1: S is infinite. By (2.7), (2.24) and (2.27),

$$\begin{aligned}\|F_1\|^2 &\geq \sum_{k=0}^{\infty}(\|F_k\|^2 - \|F_{k+1}\|^2) \geq \sum_{k\in S}(\|F_k\|^2 - \|F_{k+1}\|^2) \\ &\geq \sum_{k\in S} p_0 Pred_k \\ &\geq \sum_{k\in S} p_0\|J_k^T F_k\| \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\} \\ &\geq \sum_{k\in S} p_0 \tau \min\left\{\|d_k\|, \frac{\tau}{\kappa_{bj}^2}\right\}. \end{aligned} \tag{2.29}$$

Hence,

$$\lim_{k\in S, k\to+\infty} d_k = 0. \tag{2.30}$$

We can get from (1.5) and (2.22) that

$$\mu_k \to +\infty, \quad \lambda_k \to +\infty, \quad k \notin S,$$

and it follows from $\|F_k\|$ is non-increasing, (2.24), (2.27) and the definition of d_k that

$$\lim_{k\notin S, k\to+\infty} d_k = 0. \tag{2.31}$$

Hence we have

$$\lim_{k\to\infty} d_k = 0. \tag{2.32}$$

By (2.24) and (2.25), we have

$$\begin{aligned}\|\hat{d}_k\| &= \| -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k)\| \\ &\leq \| \ -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k\| + \| -(J_k^T J_k + \lambda_k I)^{-1} J_k^T J_k d_k\| \\ &+ \frac{\kappa_{hj}}{1+v}\| -(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \|d_k\|^{1+v} \\ &\leq 2\|d_k\| + \frac{\kappa_{bj}\kappa_{hj}}{(1+v)\lambda_k}\|d_k\|^{1+v}. \end{aligned} \tag{2.33}$$

It follows that

$$\begin{aligned} &|\|F(x_k + d_k)\|^2 - \|F_k + J_k d_k\|^2| \\ &\leq \frac{2\kappa_{hj}}{1+v}\|F_k + J_k d_k\| \|d_k\|^{1+v} + \frac{\kappa_{hj}^2}{(1+v)^2}\|d_k\|^{2+2v}, \end{aligned} \tag{2.34}$$

and

$$\begin{aligned} &|\|F(x_k + d_k + \alpha_k \hat{d}_k)\|^2 - \|F(x_k + d_k) + J_k \hat{d}_k\|^2| \\ &\leq \frac{2\kappa_{hj}}{1+v}\|F_k + J_k(d_k + \alpha_k \hat{d}_k)\| \|d_k + \alpha_k \hat{d}_k\|^{1+v} + \frac{\kappa_{hj}^2}{(1+v)^2}\|d_k + \alpha_k \hat{d}_k\|^{2+2v} \\ &+ \frac{2\kappa_{hj}}{1+v}\|F_k + J_k(d_k + \alpha_k \hat{d}_k)\| \|d_k\|^{1+v} + \frac{\kappa_{hj}^2}{(1+v)^2}\|d_k\|^{2+2v}. \end{aligned} \tag{2.35}$$

It then follows from (2.16), (2.24), (2.25), (2.27), (2.32), (2.32) and (2.35) that

$$\begin{aligned}
|r_k - 1| &= \left|\frac{Ared_k - Pred_k}{Pred_k}\right| \\
&\le \left|\frac{\|F(y_k)\|^2 - \|F_k + J_k d_k\|^2 + \|F(x_k+s_k)\|^2 - \|F(y_k) + J_k\hat{d}_k\|^2}{\|J_k^T F_k\| \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}}\right| \\
&\le \Big|\Big(\frac{2\kappa_{hj}}{1+v}\|F_k + J_k(d_k+\alpha_k\hat{d}_k)\|\|d_k+\alpha_k\hat{d}_k\|^{1+v} + \frac{\kappa_{hj}^2}{(1+v)^2}\|d_k+\alpha_k\hat{d}_k\|^{2+2v} \\
&+ \frac{2\kappa_{hj}}{1+v}\|F_k + J_k(d_k+\alpha_k\hat{d}_k)\|\|d_k\|^{1+v} + \frac{2\kappa_{hj}^2}{(1+v)^2}\|d_k\|^{2+2v} \\
&+ \frac{2\kappa_{hj}}{1+v}\|F_k + J_k d_k\|\|d_k\|^{1+v}\Big)\Big/\Big(\|J_k^T F_k\| \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}\Big)\Big| \\
&\le \Big|\Big(\frac{2\kappa_{hj}}{1+v}\|F_k + J_k(d_k+\alpha_k\hat{d}_k)\|\|d_k+\alpha_k\hat{d}_k\|^{1+v} + \frac{\kappa_{hj}^2}{(1+v)^2}\|d_k+\alpha_k\hat{d}_k\|^{2+2v} \\
&+ \frac{2\kappa_{hj}}{1+v}\|F_k + J_k(d_k+\alpha_k\hat{d}_k)\|\|d_k\|^{1+v} + \frac{2\kappa_{hj}^2}{(1+v)^2}\|d_k\|^{2+2v} \\
&+ \frac{2\kappa_{hj}}{1+v}\|F_k + J_k d_k\|\|d_k\|^{1+v}\Big)\Big/\Big(\tau \min\left\{\|d_k\|, \frac{\tau}{\kappa_{bj}^2}\right\}\Big)\Big| \\
&\to 0. \qquad (2.36)
\end{aligned}$$

Hence,

$$\lim_{k\to+\infty} r_k = 1. \tag{2.37}$$

In view of (2.22), we know that there exists a positive constant $\tilde{\mu}$ such that

$$\mu_k < \tilde{\mu} \tag{2.38}$$

holds for all large k. This contradicts to (2.31). So (2.27) cannot be true when S is infinite.

Case 2: S is finite. Then there exists a $\tilde{k}$ such that $r_k < p_0$ for all $k > \tilde{k}$. Hence,

$$\mu_k \to +\infty. \tag{2.39}$$

By the definition of d_k, we have

$$d_k \to 0. \tag{2.40}$$

By the same analysis as (2.36), we have

$$r_k \to 1. \tag{2.41}$$

Thus, there exists a positive constant $\hat{\mu}$ such that $\mu_k < \hat{\mu}$ holds for all large k, which gives a contradiction to (2.39). Therefore, (2.27) cannot be true when S is finite.

Summarizing the above, we know that (2.26) holds true. The proof is completed. □

3 Convergence Rate of Algorithm 2.1 Under the Hölderian Local Error Bound Condition and the Hölderian Continuity of the Jacobian.

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

Assumption 3.1. (a) $F(x)$ provides a Hölderian local error bound of order $\gamma \in (0,1]$ in some neighbourhood of $x^* \in X^*$, i.e., there exist constants $c > 0$ and $0 < b < 1$ such that

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

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

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

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

Similarly to (2.25), we have

$$\|F(y) - F(x) - J(x)(y - x)\| \leq \frac{\kappa_{hj}}{1+v}\|y - x\|^{1+v}, \quad \forall x, y \in N(x^*, \frac{b}{2}). \tag{3.3}$$

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

$$\|F(y) - F(x)\| \leq \kappa_{bf}\|y - x\|, \quad \forall x, y \in N(x^*, \frac{b}{2}). \tag{3.4}$$

Denote by $\bar{x}_k$ the point in X^* that is closest to x_k, i.e.,

$$\|\bar{x}_k - x_k\| = dist(x_k, X^*). \tag{3.5}$$

In the following, we use the singular value decomposition technique to study the relationship between the norm of trial step s_k and $dist(x_k, X^*)$.

Due to the results given by Behling and Iusem in [1], we assume that $\text{rank}(J(\bar{x})) = r$ for all $\bar{x} \in N(x^*, b) \cap X^*$. Suppose that the singular value decomposition (SVD) of $J(\bar{x}_k)$ is

$$J(\bar{x}_k) = \bar{U}_k\bar{\Sigma}_k\bar{V}_k^T = (\bar{U}_{k,1}, \bar{U}_{k,2})\begin{pmatrix} \bar{\Sigma}_{k,1} & \\ & 0 \end{pmatrix}\begin{pmatrix} \bar{V}_{k,1}^T \\ \bar{V}_{k,2}^T \end{pmatrix} = \bar{U}_{k,1}\bar{\Sigma}_{k,1}\bar{V}_{k,1}^T, \tag{3.6}$$

where $\bar{\Sigma}_{k,1} = diag(\bar{\sigma}_{k,1}, \ldots, \bar{\sigma}_{k,r}) > 0$. Suppose the SVD of J_k is

$$\begin{aligned} J_k =& U_k\Sigma_kV_k^T = (U_{k,1}, U_{k,2})\begin{pmatrix} \Sigma_{k,1} & \\ & \Sigma_{k,2} \end{pmatrix}\begin{pmatrix} V_{k,1}^T \\ V_{k,2}^T \end{pmatrix} \\ =& U_{k,1}\Sigma_{k,1}V_{k,1}^T + U_{k,2}\Sigma_{k,2}V_{k,2}^T, \end{aligned} \tag{3.7}$$

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

$$J_k = U_1\Sigma_1V_1^T + U_2\Sigma_2V_2^T \tag{3.8}$$

for convenience.

Without loss of generality, we assume that both x_k and $x_k + d_k$ lie in $N(x^*, \frac{b}{4})$.

Lemma 3.1. *Under Assumption 3.1, there exists a constant $\tilde{c}_1 > 0$ such that*

$$\|s_k\| \leq \tilde{c}_1 dist(x_k, X^*)^{\min\left\{1, 1+v-\frac{\delta}{2\gamma}, 1+v+\max\{v-\frac{\delta}{\gamma}, -\frac{\delta}{2\gamma}\}, (1+v)(1+v-\frac{\delta}{2\gamma})+\max\{v-\frac{\delta}{\gamma}, -\frac{\delta}{2\gamma}\}\right\}}. \tag{3.9}$$

Proof. Since $x_k \in N(x^*, \frac{b}{4})$, we have

$$\|\bar{x}_k - x^*\| \le \|\bar{x}_k - x_k\| + \|x_k - x^*\| \le \frac{b}{2}. \tag{3.10}$$

So, $\bar{x}_k \in N(x^*, \frac{b}{2})$. Then it follows from (3.1) that

$$\lambda_k = \mu_k \|F_k\|^\delta \ge \mu_0 c^{\frac{\delta}{\gamma}} \|\bar{x}_k - x_k\|^{\frac{\delta}{\gamma}}. \tag{3.11}$$

Since d_k is the minimizer of $\varphi_k(d)$, we get from (3.3) and (3.11) that

$$\begin{aligned}
\|d_k\|^2 &\le \frac{\varphi_k(d_k)}{\lambda_k} \\
&\le \frac{\varphi_k(\bar{x}_k - x_k)}{\lambda_k} \\
&= \frac{\|F_k + J_k(\bar{x}_k - x_k)\|^2 + \lambda_k \|\bar{x}_k - x_k\|^2}{\lambda_k} \\
&\le \frac{\kappa_{hj}^2 c^{-\frac{\delta}{\gamma}}}{\mu_0(1+v)^2} \|\bar{x}_k - x_k\|^{2+2v-\frac{\delta}{\gamma}} + \|\bar{x}_k - x_k\|^2 \\
&\le \Big(\frac{\kappa_{hj}^2 c^{-\frac{\delta}{\gamma}}}{\mu_0(1+v)^2} + 1\Big) \|\bar{x}_k - x_k\|^{2\min\{1, 1+v-\frac{\delta}{2\gamma}\}},
\end{aligned} \tag{3.12}$$

which gives

$$\|d_k\| \le \tilde{c} \|\bar{x}_k - x_k\|^{\min\{1, 1+v-\frac{\delta}{2\gamma}\}}, \tag{3.13}$$

where $\tilde{c} > 0$. As analyzed in (2.33), we know

$$\|\hat{d}_k\| \le 2\|d_k\| + \frac{\kappa_{hj}}{1+v} \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \|d_k\|^{1+v}. \tag{3.14}$$

Using the SVD of J_k, we have

$$\begin{aligned}
&\|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \\
&= \|(V_1, V_2) \begin{pmatrix} (\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 & \\ & (\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 \end{pmatrix} \begin{pmatrix} U_1^T \\ U_2^T \end{pmatrix} \| \\
&\le \| \begin{pmatrix} \Sigma_1^{-1} & \\ & (\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 \end{pmatrix} \| \\
&\le \|\Sigma_1^{-1}\| + \|(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2\|.
\end{aligned} \tag{3.15}$$

By the theory of matrix perturbation [14],

$$\|diag(\Sigma_1 - \bar{\Sigma}_{k,1}, \Sigma_2)\| \le \|J_k - J(\bar{x}_k)\| \le \kappa_{hj} \|\bar{x}_k - x_k\|^v. \tag{3.16}$$

The above inequalities imply

$$\|\Sigma_1 - \bar{\Sigma}_{k,1}\| \le \kappa_{hj} \|\bar{x}_k - x_k\|^v, \quad \|\Sigma_2\| \le \kappa_{hj} \|\bar{x}_k - x_k\|^v. \tag{3.17}$$

Without loss of generality, we assume that $\kappa_{hj}\|\bar{x}_k - x_k\|^v \le \frac{\bar{\sigma}}{2}$ holds for all large k. By (3.16),

$$\|\Sigma_1^{-1}\| \le \frac{1}{\bar{\sigma} - \kappa_{hj}\|\bar{x}_k - x_k\|^v} \le \frac{2}{\bar{\sigma}}, \tag{3.18}$$

moreover, we have from (3.11) that

$$(\Sigma_2^2+\lambda_k I)^{-1}\Sigma_2 \le \frac{\Sigma_2}{\lambda_k} \le \frac{1}{\mu_0 c^{\frac{\delta}{\gamma}}}\|\bar{x}_k - x_k\|^{v-\frac{\delta}{\gamma}}, \tag{3.19}$$

and

$$(\Sigma_2^2+\lambda_k I)^{-1}\Sigma_2 \le \frac{1}{2\sqrt{\lambda_k}} \le \frac{1}{\sqrt{\mu_0} c^{\frac{\delta}{2\gamma}}}\|\bar{x}_k - x_k\|^{-\frac{\delta}{2\gamma}}. \tag{3.20}$$

Then the above inequalities together with (3.13), (3.14) and (3.15) show that there exists a posotive $\tilde{c}_1$ such that

$$\|\hat{d}_k\| \le 2\|d_k\| + \frac{\kappa_{hj}}{(1+v)}\|d_k\|^{1+v}\big(\frac{2}{\bar{\sigma}} + \max\{\frac{1}{\mu_0 c^{\frac{\delta}{\gamma}}}, \frac{1}{2\sqrt{\mu_0} c^{\frac{\delta}{2\gamma}}}\}\|\bar{x}_k - x_k\|^{\max\{v-\frac{\delta}{\gamma}, -\frac{\delta}{2\gamma}\}}\big). \tag{3.21}$$

So we obtain that

$$\begin{aligned}
\|s_k\| &= \|d_k + \alpha_k \hat{d}_k\| \\
&\le \|d_k\| + \hat{\alpha}\|\hat{d}_k\| \\
&\le \|d_k\| + 2\hat{\alpha}\|d_k\| + \frac{2\hat{\alpha}\kappa_{hj}}{(1+v)\bar{\sigma}}\|d_k\|^{1+v} \\
&+ \frac{\kappa_{hj}}{1+v}\max\{\frac{1}{\mu_0 c^{\frac{\delta}{\gamma}}}, \frac{1}{2\sqrt{\mu_0} c^{\frac{\delta}{2\gamma}}}\}\|d_k\|^{1+v}\|\bar{x}_k - x_k\|^{\max\{v-\frac{\delta}{\gamma}, -\frac{\delta}{2\gamma}\}} \\
&\le (1+2\hat{\alpha})\tilde{c}\|\bar{x}_k - x_k\|^{\min\{1, 1+v-\frac{\delta}{2\gamma}\}} + \frac{2\hat{\alpha}\kappa_{hj}}{(1+v)\bar{\sigma}}\tilde{c}^{1+v}\|\bar{x}_k - x_k\|^{\min\{1+v, (1+v)(1+v-\frac{\delta}{2\gamma})\}} \\
&+ \frac{\kappa_{hj}}{1+v}\max\{\frac{1}{\mu_0 c^{\frac{\delta}{\gamma}}}, \frac{1}{2\sqrt{\mu_0} c^{\frac{\delta}{2\gamma}}}\}\tilde{c}^{1+v}\|\bar{x}_k - x_k\|^{\min\{1+v, (1+v)(1+v-\frac{\delta}{2\gamma})\} + \max\{v-\frac{\delta}{\gamma}, -\frac{\delta}{2\gamma}\}} \\
&\le \tilde{c}_1\|\bar{x}_k - x_k\|^{\min\left\{1, 1+v-\frac{\delta}{2\gamma}, 1+v+\max\{v-\frac{\delta}{\gamma}, -\frac{\delta}{2\gamma}\}, (1+v)(1+v-\frac{\delta}{2\gamma}) + \max\{v-\frac{\delta}{\gamma}, -\frac{\delta}{2\gamma}\}\right\}}
\end{aligned} \tag{3.22}$$

for $\tilde{c}_1 > 0$. The proof is completed. □

In view of Algorithm 2.1, we show that μ_k is bounded above.

Lemma 3.2. *Under Assumption 3.1, if* $v > \max\left\{\frac{1}{\gamma} - 1, \frac{1}{\gamma(1+v)-\frac{\delta}{2}} - 1, \frac{1-\gamma}{\gamma(1+v)-\frac{\delta}{2}}\right\}$, *then there exists a constant* $\bar{\mu} > 0$ *such that*

$$\mu_k \le \bar{\mu} \tag{3.23}$$

holds for all large k.

Proof. We discuss in two cases.

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

$$\begin{aligned}
\|F_k\| - \|F_k + J_k d_k\| &\ge \|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\| \\
&\ge c^{\frac{1}{\gamma}}\|\bar{x}_k - x_k\|^{\frac{1}{\gamma}} - \frac{\kappa_{hj}}{1+v}\|\bar{x}_k - x_k\|^{1+v} \\
&\ge c_1\|\bar{x}_k - x_k\|^{\frac{1}{\gamma}} \\
&\ge c_2\|d_k\|^{\max\left\{\frac{1}{\gamma}, \frac{1}{\gamma(1+v)-\frac{\delta}{2}}\right\}}
\end{aligned} \tag{3.24}$$

holds for some $c_1, c_2 > 0$.

Case 2: $\|\bar{x}_k - x_k\| > \|d_k\|$. By (3.24),

$$\begin{aligned}\|F_k\| - \|F_k + J_k d_k\| &\geq \|F_k\| - \left\| F_k + \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} J_k(x_k - \bar{x}_k) \right\| \\ &\geq \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}(\|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\|) \\ &\geq c_1 \|d_k\| \|\bar{x}_k - x_k\|^{\frac{1}{\gamma}-1} \\ &\geq c_3 \|d_k\|^{\max\left\{\frac{1}{\gamma}, \frac{1-\gamma}{\gamma(1+v)-\frac{\delta}{2}}+1\right\}} \end{aligned} \tag{3.25}$$

holds for some $c_3 > 0$.

Thus, by (3.24) and (3.25),

$$\begin{aligned} Pred_k &= \|F_k\|^2 - \|F_k + J_k d_k\|^2 + \|F(y_k)\|^2 - \left\| F(y_k) + \alpha_k J_k \hat{d}_k \right\|^2 \\ &\geq (\|F_k\| + \|F_k + J_k d_k\|)(\|F_k\| - \|F_k + J_k d_k\|) \\ &\geq \|F_k\|(\|F_k\| - \|F_k + J_k d_k\|) \\ &\geq c_4 \|F_k\| \|d_k\|^{\max\left\{\frac{1}{\gamma}, \frac{1}{\gamma(1+v)-\frac{\delta}{2}}, \frac{1-\gamma}{\gamma(1+v)-\frac{\delta}{2}}+1\right\}} \end{aligned} \tag{3.26}$$

holds for some $c_4 > 0$.

Since $\|F_k + J_k d_k\| \leq \|F_k\|$, $v > \max\left\{\frac{1}{\gamma} - 1, \frac{1}{\gamma(1+v)-\frac{\delta}{2}} - 1, \frac{1-\gamma}{\gamma(1+v)-\frac{\delta}{2}}\right\}$, combining with the analysis of (2.34), (2.35) and (2.36) , we have

$$\begin{aligned} |r_k - 1| &= \left| \frac{Ared_k - Pred_k}{Pred_k} \right| \\ &\leq \left| \frac{\|F(y_k)\|^2 - \|F_k + J_k d_k\|^2 + \|F(x_k + s_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2}{c_4 \|F_k\| \|d_k\|^{\max\left\{\frac{1}{\gamma}, \frac{1}{\gamma(1+v)-\frac{\delta}{2}}, \frac{1-\gamma}{\gamma(1+v)-\frac{\delta}{2}}+1\right\}}} \right| \\ &\leq \Big| \Big(\frac{2\kappa_{hj}}{1+v} \|F_k + J_k(d_k + \alpha_k \hat{d}_k)\| \|d_k + \alpha_k \hat{d}_k\|^{1+v} + \frac{\kappa_{hj}^2}{(1+v)^2} \|d_k + \alpha_k \hat{d}_k\|^{2+2v} \\ &\quad + \frac{2\kappa_{hj}}{1+v} \|F_k + J_k(d_k + \alpha_k \hat{d}_k)\| \|d_k\|^{1+v} + \frac{2\kappa_{hj}^2}{(1+v)^2} \|d_k\|^{2+2v} \\ &\quad + \frac{2\kappa_{hj}}{1+v} \|F_k + J_k d_k\| \|d_k\|^{1+v} \Big) \Big/ \Big(c_4 \|F_k\| \|d_k\|^{\max\left\{\frac{1}{\gamma}, \frac{1}{\gamma(1+v)-\frac{\delta}{2}}, \frac{1-\gamma}{\gamma(1+v)-\frac{\delta}{2}}+1\right\}} \Big) \Big| \\ &\to 0. \end{aligned} \tag{3.27}$$

Therefore,

$$\lim_{k \to +\infty} r_k = 1. \tag{3.28}$$

By the updating rule (2.22), we know (3.23) holds. The proof is completed. □

Next we investigate the convergence rate of Algorithm 2.1. The estimations of $\|U_1 U_1^T F_k\|$ and $\|U_2 U_2^T F_k\|$ are given as follows.

Lemma 3.3. *Under Assumption 3.1, we have*

1. $\|U_1U_1^TF_k\| \le \kappa_{bf}\|\bar{x}_k - x_k\|$;
2. $\|U_2U_2^TF_k\| \le 2\kappa_{hj}\|\bar{x}_k - x_k\|^{1+v}$.

Proof. Result (1) follows immediately from (3.4).

Let $\tilde{J}_k = U_1\Sigma_1V_1^T$ and $\tilde{s}_k = -\tilde{J}_k^+F_k$, where $\tilde{J}_k^+$ is the pesudo-inverse of $\tilde{J}_k$. Then $\tilde{s}_k$ is the least squares solution of $\min_{s\in R^n}\|F_k + \tilde{J}_ks\|$. By (3.3) and (3.17),

$$\begin{aligned}\|U_2U_2^TF_k\| &= \|F_k + \tilde{J}_k\tilde{s}_k\| \\ &\le \|F_k + \tilde{J}_k(\bar{x}_k - x_k)\| \\ &\le \|F_k + J_k(\bar{x}_k - x_k)\| + \|(\tilde{J}_k - J_k)(\bar{x}_k - x_k)\| \\ &\le \frac{\kappa_{hj}}{1+v}\|\bar{x}_k - x_k\|^{1+v} + \|U_2\Sigma_2V_2^T(\bar{x}_k - x_k)\| \\ &\le \frac{\kappa_{hj}}{1+v}\|\bar{x}_k - x_k\|^{1+v} + \kappa_{hj}\|\bar{x}_k - x_k\|^v\|\bar{x}_k - x_k\| \\ &\le 2\kappa_{hj}\|\bar{x}_k - x_k\|^{1+v}. \end{aligned} \tag{3.29}$$

The proof is completed. □

By the SVD of J_k, we get

$$d_k = -V_1(\Sigma_1^2 + \lambda_kI)^{-1}\Sigma_1U_1^TF_k - V_2(\Sigma_2^2 + \lambda_kI)^{-1}\Sigma_2U_2^TF_k, \tag{3.30}$$

$$\hat{d}_k = -V_1(\Sigma_1^2 + \lambda_kI)^{-1}\Sigma_1U_1^TF(y_k) - V_2(\Sigma_2^2 + \lambda_kI)^{-1}\Sigma_2U_2^TF(y_k), \tag{3.31}$$

$$\begin{aligned}F_k + J_kd_k &= F_k - U_1\Sigma_1(\Sigma_1^2 + \lambda_kI)^{-1}\Sigma_1U_1^TF_k - U_2\Sigma_2(\Sigma_2^2 + \lambda_kI)^{-1}\Sigma_2U_2^TF_k \\ &= \lambda_kU_1(\Sigma_1^2 + \lambda_kI)^{-1}U_1^TF_k + \lambda_kU_2(\Sigma_2^2 + \lambda_kI)^{-1}U_2^TF_k, \end{aligned} \tag{3.32}$$

and

$$\begin{aligned}F(y_k) + J_k\hat{d}_k &= F(y_k) - U_1\Sigma_1(\Sigma_1^2 + \lambda_kI)^{-1}\Sigma_1U_1^TF(y_k) - U_2\Sigma_2(\Sigma_2^2 + \lambda_kI)^{-1}\Sigma_2U_2^TF(y_k) \\ &= \lambda_kU_1(\Sigma_1^2 + \lambda_kI)^{-1}U_1^TF(y_k) + \lambda_kU_2(\Sigma_2^2 + \lambda_kI)^{-1}U_2^TF(y_k). \end{aligned} \tag{3.33}$$

Lemma 3.4. *Under the assumptions of Lemma 3.2, we have*

1. $\|U_1U_1^TF(y_k)\| \le c_6\|\bar{x}_k - x_k\|^{\min\{1+\delta, 1+v, (1+v)(1+v-\frac{\delta}{2\gamma})\}}$;
2. $\|U_2U_2^TF(y_k)\| \le c_7\|\bar{x}_k - x_k\|^{\beta_0}$,

where
$$\begin{aligned}\beta_0 = \min\Big\{&\gamma(1+\delta)(1+v), \gamma(1+v)^2, \gamma(1+v)^2(1+v-\frac{\delta}{2\gamma}), v+\gamma(1+\delta), \\ &v+\gamma(1+v), v+\gamma(1+v)(1+v-\frac{\delta}{2\gamma}), v(1+v-\frac{\delta}{2\gamma})+\gamma(1+\delta), \\ &v(1+v-\frac{\delta}{2\gamma})+\gamma(1+v), v(1+v-\frac{\delta}{2\gamma})+\gamma(1+v)(1+v-\frac{\delta}{2\gamma})\Big\}.\end{aligned}$$

Proof. By Lemmas 3.2 and (3.4),

$$\lambda_k = \mu_k \|F_k\|^\delta \le \bar{\mu}\kappa_{bf}^\delta \|\bar{x}_k - x_k\|^\delta. \tag{3.34}$$

It follows from (3.18), (3.32), (3.34) and Lemma 3.3 that

$$\begin{aligned} \|F_k + J_k d_k\| &\le \frac{4\bar{\mu}\kappa_{bf}^{1+\delta}}{\bar{\sigma}^2}\|\bar{x}_k - x_k\|^{1+\delta} + 2\kappa_{hj}\|\bar{x}_k - x_k\|^{1+v} \\ &\le c_5\|\bar{x}_k - x_k\|^{\min\{1+\delta,1+v\}}, \end{aligned} \tag{3.35}$$

where $c_5 = \frac{4\bar{\mu}\kappa_{bf}^{1+\delta}}{\bar{\sigma}^2} + 2\kappa_{hj}$.

The above inequality, together with (3.3) and (3.13) implies that

$$\begin{aligned} \|F(y_k)\| &= \|F(x_k + d_k)\| \\ &\le \|F_k + J_k d_k\| + \kappa_{hj}\|d_k\|^{1+v} \\ &\le c_5\|\bar{x}_k - x_k\|^{\min\{1+\delta,1+v\}} + \kappa_{hj}\|d_k\|^{1+v} \\ &\le c_6\|\bar{x}_k - x_k\|^{\min\{1+\delta,1+v,(1+v)(1+v-\frac{\delta}{2\gamma})\}}, \end{aligned} \tag{3.36}$$

where $c_6 = c_5 + \kappa_{hj} > 0$. So we have

$$\|U_1U_1^T F(y_k)\| \le \|F(y_k)\| \le c_6\|\bar{x}_k - x_k\|^{\min\{1+\delta,1+v,(1+v)(1+v-\frac{\delta}{2\gamma})\}}. \tag{3.37}$$

Thus the Hölderian local error bound condition yields

$$\begin{aligned} \|\bar{y}_k - y_k\| &\le \frac{1}{c}\|F(y_k)\|^\gamma \\ &\le \frac{c_6}{c}\|\bar{x}_k - x_k\|^{\min\{\gamma(1+\delta),\gamma(1+v),\gamma(1+v)(1+v-\frac{\delta}{2\gamma})\}}. \end{aligned} \tag{3.38}$$

Let $\tilde{J}_k = U_1\Sigma_1V_1^T$ and $\tilde{p}_k = -\tilde{J}_k^+ F_{k,1}$. Since $\tilde{p}_k$ is the least squares solution of $\min\limits_{p\in R^n} \|F_k + \tilde{J}_k p\|$, we deduce from (3.2), (3.3), (3.13), (3.17) and (3.38) that

$$\begin{aligned} &\|U_2U_2^T F(y_k)\| \\ &= \|F(y_k) + \tilde{J}_k\tilde{p}_k\| \\ &\le \|F(y_k) + \tilde{J}_k(\bar{y}_k - y_k)\| \\ &\le \|F(y_k) + J(y_k)(\bar{y}_k - y_k)\| + \|(\tilde{J}_k - J(y_k))(\bar{y}_k - y_k)\| \\ &\le \frac{\kappa_{hj}}{1+v}\|\bar{y}_k - y_k\|^{1+v} + \|(J_k - J(y_k) - U_2\Sigma_2V_2^T)(\bar{y}_k - y_k)\| \\ &\le \frac{\kappa_{hj}}{1+v}\|\bar{y}_k - y_k\|^{1+v} + \|(J_k - J(y_k))(\bar{y}_k - y_k)\| + \|U_2\Sigma_2V_2^T(\bar{y}_k - y_k)\| \\ &\le \frac{\kappa_{hj}}{1+v}\|\bar{y}_k - y_k\|^{1+v} + \kappa_{hj}\|d_k\|^v\|\bar{y}_k - y_k\| + \kappa_{hj}\|\bar{x}_k - x_k\|^v\|\bar{y}_k - y_k\| \\ &\le \frac{\kappa_{hj}c_6^{1+v}}{(1+v)c^{1+v}}\|\bar{x}_k - x_k\|^{(1+v)\min\{\gamma(1+\delta),\gamma(1+v),\gamma(1+v)(1+v-\frac{\delta}{2\gamma})\}} \\ &+ \frac{\tilde{c}\kappa_{hj}c_6}{c}\|\bar{x}_k - x_k\|^{\min\{v,v(1+v-\frac{\delta}{2\gamma})\}+\min\{\gamma(1+\delta),\gamma(1+v),\gamma(1+v)(1+v-\frac{\delta}{2\gamma})\}} \\ &+ \frac{c_6\kappa_{hj}}{c}\|\bar{x}_k - x_k\|^{v+\min\{\gamma(1+\delta),\gamma(1+v),\gamma(1+v)(1+v-\frac{\delta}{2\gamma})\}} \\ &\le c_7\|\bar{x}_k - x_k\|^{\beta_0}, \end{aligned} \tag{3.39}$$

where $c_7 = \frac{\kappa_{hj} c_6^{1+v}}{(1+v)c^{1+v}} + \frac{\tilde{c}\kappa_{hj} c_6}{c} + \frac{c_6 \kappa_{hj}}{c} > 0.$

□

Theorem 3.5. *Under conditions of Lemma 3.2, the sequence $\{x_k\}$ generated by Algorithm 2.1 converges to the solution set of (1.1) with order $\gamma\beta_1$, where*

$$\beta_1 = \min\Big\{1+2\delta, 1+\delta+v, \delta+(1+v)(1+v-\frac{\delta}{2\gamma}), \beta_0, 1+2v, v+(1+v)(1+v-\frac{\delta}{2\gamma}), 2v-\frac{\delta}{\gamma}+\beta_0, 1+\delta+v(1+v-\frac{\delta}{2\gamma}), 1+v+v(1+v-\frac{\delta}{2\gamma}), (1+2v)(1+v-\frac{\delta}{2\gamma}), v(1+v-\frac{\delta}{2\gamma})+v-\frac{\delta}{\gamma}+\beta_0, (1+v)(v-\frac{\delta}{\gamma}+\beta_0)\Big\}.$$

Also, if $v > \frac{1}{\gamma} - 1$ and $v < \delta < 2\gamma v$, the sequence x_k generated by 2.1 converges to some solution of (1.1) superlinearly with order $\min\{\gamma(1+\delta+v), \gamma v + \gamma^2(1+v)\}$.

Proof. It follows from (3.34), (3.17),(3.18), (3.19), (3.20), (3.31) and Lemma 3.4 that

$$\begin{aligned}\|\hat{d}_k\| &\le \|\Sigma_1^{-1}\|\|U_1^T F(y_k)\| + \|\lambda_k^{-1}\Sigma_2\|\|U_2^T F(y_k)\| \\ &\le \frac{2c_6}{\bar{\sigma}}\|\bar{x}_k - x_k\|^{\min\{1+\delta,1+v,(1+v)(1+v-\frac{\delta}{2\gamma})\}} + \frac{\kappa_{hj}c_7}{\mu_0 c^{\frac{\delta}{\gamma}}}\|\bar{x}_k - x_k\|^{v-\frac{\delta}{\gamma}+\beta_0} \\ &\le c_8\|\bar{x}_k - x_k\|^{\min\{1+\delta,1+v,(1+v)(1+v-\frac{\delta}{2\gamma}),v-\frac{\delta}{\gamma}+\beta_0\}}, \end{aligned} \tag{3.40}$$

where $c_8 = \frac{2c_6}{\bar{\sigma}} + \frac{\kappa_{hj}c_7}{\mu_0 c^{\frac{\delta}{\gamma}}} > 0$. Then we can similarly get from (3.33) that

$$\begin{aligned}&\|F(y_k) + \alpha_k J_k \hat{d}_k\| \\ &\le \|F(y_k) + J_k\hat{d}_k\| \\ &= \|\lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1}U_1^T F(y_k)\| + \|\lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1}U_2^T F(y_k)\| \\ &\le \|\lambda_k \Sigma_1^{-2}\|\|U_1^T F(y_k)\| + \|U_2^T F(y_k)\| \\ &\le \frac{4\bar{\mu}\kappa_{bf}^{\delta}c_6}{\bar{\sigma}^2}\|\bar{x}_k - x_k\|^{\delta+\min\{1+\delta,1+v,(1+v)(1+v-\frac{\delta}{2\gamma})\}} + c_7\|\bar{x}_k - x_k\|^{\beta_0} \\ &\le c_9\|\bar{x}_k - x_k\|^{\min\{1+2\delta,1+\delta+v,\delta+(1+v)(1+v-\frac{\delta}{2\gamma}),\beta_0\}}, \end{aligned} \tag{3.41}$$

where $c_9 = \frac{4\bar{\mu}\kappa_{bf}^{\delta}c_6}{\bar{\sigma}^2} + c_7 > 0$. Thus, we have from (3.1)-(3.3), (3.41) and (3.42) that

$$\begin{aligned}&\|F(x_{k+1})\| \\ &= \|F(y_k + \alpha_k\hat{d}_k)\| \\ &\le \|F(y_k) + \alpha_k J(y_k)\hat{d}_k\| + \kappa_{hj}\alpha_k^{1+v}\|\hat{d}_k\|^{1+v} \\ &\le \|F(y_k) + \alpha_k J_k\hat{d}_k\| + \|(J(y_k) - J_k)\hat{d}_k\| + \kappa_{hj}{\alpha_k}^{1+v}\|\hat{d}_k\|^{1+v} \\ &\le \|F(y_k) + \alpha_k J_k\hat{d}_k\| + \kappa_{hj}\hat{\alpha}\|d_k\|^v\|\hat{d}_k\| + \kappa_{hj}\hat{\alpha}^{1+v}\|\hat{d}_k\|^{1+v} \\ &\le c_9\|\bar{x}_k - x_k\|^{\min\{1+2\delta,1+\delta+v,\delta+(1+v)(1+v-\frac{\delta}{2\gamma}),\beta_0\}} \\ &+ \kappa_{hj}\hat{\alpha}\tilde{c}^v c_8\|\bar{x}_k - x_k\|^{\min\{v,v(1+v-\frac{\delta}{2\gamma})\}+\min\{1+\delta,1+v,(1+v)(1+v-\frac{\delta}{2\gamma}),v-\frac{\delta}{\gamma}+\beta_0\}} \\ &+ \kappa_{hj}\hat{\alpha}c_8^{1+v}\|\bar{x}_k - x_k\|^{(1+v)\min\{1+\delta,1+v,(1+v)(1+v-\frac{\delta}{2\gamma}),v-\frac{\delta}{\gamma}+\beta_0\}} \\ &\le c_{10}\|\bar{x}_k - x_k\|^{\beta_1}, \end{aligned} \tag{3.42}$$

where $c_{10} = c_9 + \kappa_{hj}\hat{\alpha}\tilde{c}^v c_8 + \kappa_{hj}\hat{\alpha}c_8^{1+v} > 0$ and

$$\begin{aligned}\beta_1 = \min\Big\{&1+2\delta, 1+\delta+v, \delta+(1+v)(1+v-\frac{\delta}{2\gamma}), \beta_0, 1+2v, v+(1+v)(1+v-\frac{\delta}{2\gamma}),\\ &2v-\frac{\delta}{\gamma}+\beta_0, 1+\delta+v(1+v-\frac{\delta}{2\gamma}), 1+v+v(1+v-\frac{\delta}{2\gamma}), (1+2v)(1+v-\frac{\delta}{2\gamma}),\\ &v(1+v-\frac{\delta}{2\gamma})+v-\frac{\delta}{\gamma}+\beta_0, (1+v)(v-\frac{\delta}{\gamma}+\beta_0)\Big\}. \end{aligned} \tag{3.43}$$

Then,

$$\begin{aligned}\|\bar{x}_{k+1} - x_{k+1}\| &\le \frac{1}{c}\|F(x_{k+1})\|^{\gamma}\\ &\le \frac{c_{10}}{c}\|\bar{x}_k - x_k\|^{\gamma\beta_1}.\end{aligned} \tag{3.44}$$

This means that $\{x_k\}$ converges to the solution set X^* of (1.1) with order $\gamma\beta_1$.

Also, when $v > \frac{1}{\gamma} - 1$ and $v < \delta < 2\gamma v$, we have

$$\beta_0 = v + \gamma(1+v) \quad \textit{and} \quad \beta_1 = \min\{1+\delta+v, v+\gamma(1+v)\}. \tag{3.45}$$

Then the sequence converges with the order

$$\min\{\gamma(1+\delta+v), \gamma v + \gamma^2(1+v)\}, \tag{3.46}$$

and

$$\gamma(1+\delta+v) > \gamma(1+v) > 1,\ \gamma v + \gamma^2(1+v) > \gamma v + \gamma = \gamma(1+v) > 1. \tag{3.47}$$

We can get from the above two equations that the sequence x_k generated by 2.1 converges to some solution of (1.1) superlinearly with order $\min\{\gamma(1+\delta+v), \gamma v + \gamma^2(1+v)\}$. The proof is completed. □

Remark 3.6. Under Assumption 3.1, if $\gamma = v = 1$ and $\delta \in (0, 2]$, then the Hölderian local error bound condition degenerates into the local error bound condition and the sequence $\{x_k\}$ generated by Algorithm 2.1 with $\lambda_k = \mu_k\|F_k\|^{\delta}$ satisfies

$$\|\bar{x}_{k+1} - x_{k+1}\| \le O\|\bar{x}_k - x_k\|^3, \tag{3.48}$$

which is the same as the result given in [4].

4 Conclusions

In this paper, we consider an accelerated modified Levenberg-Marquardt (LM) algorithm for the system of nonlinear equations. We study the convergence properties of this algorithm 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. We add line search under these more general assumptions, which means that the Jacobian evaluated only after every 2 computations of the step. We show that the LM algorithm converges globally and the convergence order of the algorithm is $\gamma\beta_1$ defined in Theorem 3.5. Also, if $v > \frac{1}{\gamma} - 1$ and $v < \delta < 2\gamma v$, the sequence x_k generated by Algorithm 2.1 converges to some solution of (1.1) superlinearly with order $\min\{\gamma(1+\delta+v), \gamma v + \gamma^2(1+v)\}$.

References

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

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

[3] J.Y. Fan, The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence, *Math. Comp.* 81 (2012) 447–466.

[4] J.Y. Fan, Accelerating the modified levenberg-Marquardt method for nonlinear equations *Math. Comp.* 83 (2014) 1173–1187.

[5] J.Y. Fan, J.C. Huang and J.Y. Pan, An adaptive multi-step Levenberg-Marquardt method, *J. Sci. Comput.* 78 (2019) 531–548.

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

[7] C.T. Kelley, *Solving Nonlinear Equations with Newton's method*, Fundamentals of Algorithms, SIAM, Philadelphia, 2003.

[8] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, *Quart. Appl. Math.* 2 (1944) 164–166.

[9] D.W. Marquardt, An algorithm for least-squares estimation of nonlinear inequalities, *SIAM J. Appl. Math.* 11 (1963) 431–441.

[10] J.J. Moré, The Levenberg-Marquardt algorithm: implementation and theory, in: *Lecture Notes in Mathematics 630: Numerical Analysis*, G.A. Watson (ed.), Springer-Verlag, Berlin, 1978, pp. 105–116.

[11] M.J.D. Powell, Convergence properties of a class of minimization algorithms, *Nonlinear programming*, 2 (1975) 1–27.

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

[13] 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) 2265–2275.

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

[15] N. Yamashita and M. Fukushima, On the rate of convergence of the Levenberg-Marquardt method, *Computing*, (Supplement 15) (2001) 237–249.

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

[17] 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) 791–804.

Manuscript received 26 April 2021
revised 30 September 2021
accepted for publication 6 November 2021

Haiyan Wang
School of Mathematical Sciences
Shanghai Jiao Tong University
Shanghai 200240, China
E-mail address: haiyanjady@sjtu.edu.cn