



# A modified Levenberg–Marquardt method for solving system of nonlinear equations

Liang Chen<sup>1,3</sup> · Yanfang Ma<sup>2</sup>

Received: 24 July 2022 / Revised: 16 November 2022 / Accepted: 20 November 2022 /

Published online: 30 November 2022

© The Author(s) under exclusive licence to Korean Society for Informatics and Computational Applied Mathematics 2022

## Abstract

A modified Levenberg–Marquardt methods for solving system of nonlinear equations is described and analysed in this paper. Specifically, we propose a convex combination of  $\|F_k\|^\delta$  and  $\|J_k^T F_k\|^\delta$  with  $\delta \in [1, 2]$  for the LM parameter and analyse the convergence of this modified Levenberg–Marquardt method under the  $\gamma$ -Hölderian local error bound of the underlying function and the  $v$ -Hölderian continuity of its Jacobian. The results show that, under some suitable relationships of exponents  $v$ ,  $\gamma$  and  $\delta$ , the modified Levenberg–Marquardt method converges to the solution set of the system of nonlinear equations at least superlinearly. Numerical experiments show the new algorithm is efficient.

**Keywords** System of nonlinear equations · Levenberg–Marquardt method ·  $\gamma$ -Hölderian local error bound ·  $v$ -Hölderian continuity

**Mathematics Subject Classification** 65K05 · 90C30

## 1 Introduction

In this paper, we consider the system of nonlinear equations

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

---

✉ Liang Chen  
clmyf2@163.com ; chenliang1977@gmail.com

<sup>1</sup> School of Sciences, Changzhou Institute of Technology, Changzhou 213032, Jiangsu, People's Republic of China

<sup>2</sup> School of Computer Science and Information Engineering, Changzhou Institute of Technology, Changzhou 213032, Jiangsu, People's Republic of China

<sup>3</sup> School of Mathematical Sciences, Huaibei Normal University, Huaibei 235000, Anhui, People's Republic of China

where  $F(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a continuously differentiable function. We assume that the system of nonlinear Eq. (1) has a nonempty solution set  $X_*$ , and refer  $\|\cdot\|$  to the Euclidean 2-norm.

As we all know, the system of nonlinear equations is one of the cornerstones in many fields of science, including engineering, finance, machine learning, etc. There have been many well-known iterative methods to find one or more solutions for the system of nonlinear Eq. (1), which may be summarized by two strategies: the line search and the trust region paradigms. The line search strategy computes a descent direction and a suitable stepsize to obtain a new estimation of the minimum point, while the trust region methods choose a step and a stepsize in a region around the current iterate within which we trust. Both these strategies contain the globally convergent line search and trust region methods, the Newton-based methods, the conjugate gradient methods, the quasi-Newton BFGS, DFP, and SR1 etc. They are operational in the system of nonlinear equations and unconstrained optimizations [1–3]. Moreover, the iterative technique is not only widely used for solving systems of nonlinear equations and unconstrained optimization, but also for solving matrix equations. Some gradient and least squares based iterative methods have been used to find the solutions to matrix equations that are often encountered in systems and control [4, 5].

Although there have been many efficient iterative methods to find a solution to the system of nonlinear equations, none of them is guaranteed to obtain the solutions for all systems of nonlinear equations. Many iterative methods are specialized at solving particular classes of systems of nonlinear equations. Therefore, the application of numerical methods is one of the important problems that many researchers are interested in. According to the characteristics of nonlinear problems, a suitable iterative method for solving a class of system of nonlinear equations that plays a key role in applied science, and there have been many works on these recently. For instance, in system identification, many identification methods and parameter estimation approaches have been developed for linear systems, bilinear systems [6] and nonlinear systems. Among them, there are many methods based on gradient and least squares that have been widely used for solving stochastic systems [7, 8], parameter identification [9], dual-rate systems [10, 11], and state space systems [12], as well as many identification methods for solving Hammerstein nonlinear systems [13].

In this paper, we focus on the numerical methods for solving the system of nonlinear equations. Due to the characteristics of the different nonlinear problems, there are many iterative algorithms have not yet been utilized efficiently. Therefore, many improved modifications of the classical iterative methods have been made [14–16]. Here, we improve the classical Levenberg–Marquardt method for solving the system of nonlinear equations and discuss the convergence under some conditions.

The rest of this paper is organized as follows. In the next section, we first propose the modified Levenberg–Marquardt algorithm for the system of nonlinear Eq. (1). Then, in Sect. 3, the convergence rate of the algorithm will be discussed with  $\gamma$ -Hölderian local error bound condition and the  $\nu$ -Hölderian continuity of the Jacobian. In Sect. 4, the global convergence result is given when the line search is used. Some numerical results are given in Sect. 5. Finally, we conclude this paper in Sect. 6.

## 2 The modified Levenberg–Marquardt method

One of the most well-known methods for solving system of nonlinear Eq. (1) is the Levenberg–Marquardt method (hereinafter referred to as the LM method) [17, 18], which computes the following trial step at every iteration

$$d_k = - \left( J_k^T J_k + \lambda_k I \right)^{-1} J_k^T F_k, \quad (2)$$

where  $F_k = F(x_k)$ ,  $J_k = J(x_k) = F'(x_k)$  is the Jacobian, and  $\lambda_k$  is the LM parameter updated by an appropriate rule. The LM method has quadratic convergence when the function  $F(x)$  is Lipschitz continuous and its Jacobian  $J(x)$  is nonsingular. To improve the computational efficiency and theoretical results of the LM method, there have been intense and interesting discussions on the choices of the LM parameter and convergence conditions in many works of literature recently [19].

Let  $N(x_*, b) = \{x \mid \|x - x_*\| \leq b\} \subset \mathbb{R}^n$ , with  $b < 1$  is a positive constant, be a neighbourhood of  $x_*$  and  $\forall x \in N(x_*, b)$ . By choosing the LM parameter as  $\lambda_k = \|F_k\|^2$ , Yamashita and Fukushima [20] show that the LM method has quadratic convergence under the following local error bound condition which is weaker than nonsingularity,

$$c \operatorname{dist}(x, X_*) \leq \|F(x)\|, \quad \forall x \in N(x_*, b), \quad (3)$$

where  $c$  is a positive constant, and  $\operatorname{dist}(x, X_*) = \inf_{\bar{x} \in X_*} \|x - \bar{x}\|$  means the distance from  $x$  to  $X_*$ . It is obvious that, if  $J(x_*)$  is nonsingular at a solution  $x_*$  of (1), then  $x_*$  is an isolated solution and  $\|F(x)\|$  provides a local error bound on some neighbourhood of  $x_*$ .

Although the local error bound condition (3) is weaker than nonsingularity of Jacobian  $J_k$ , it is not always applicable for some ill-conditioned nonlinear equations from application fields like biochemical systems. The following condition, called  $\gamma$ -Hölderian local error bound, is proposed,

$$c \operatorname{dist}(x, X_*) \leq \|F(x)\|^\gamma, \quad \forall x \in N(x_*, b), \quad (4)$$

where  $\gamma \in (0, 1]$ ,  $c > 0$  is a positive constant. For example, for the Powell singular function [21], the local error bound condition (3) is invalid, but the  $\gamma$ -Hölderian local error bound condition (4) is valid with  $\gamma = 1/2$  around the origin [22].

The  $\gamma$ -Hölderian local error bound condition (4) is obviously a generalization of the local error bound condition (3), which extends the exponent  $\gamma$  from 1 to an interval  $(0, 1]$ . Recently, under the  $\gamma$ -Hölderian local error bound condition, the convergence of some modified LM methods has been discussed with several choices of LM parameter [23–26].

From the definition of the LM trial step (2), it can be seen that the choice of the LM parameter  $\lambda_k$  has a great influence on the computational efficiency of the LM method. Therefore, in order to improve the computational efficiency, many choices of the LM parameter, besides  $\lambda_k = \|F_k\|^2$  proposed above, have been given recently. For

example, Fan and Yuan [27] chose  $\lambda_k = \|F_k\|$ , Dennis and Schnable [28] employed  $\lambda_k = O(\|J_k^T F_k\|)$  while Fischer [29] took  $\lambda_k = \|J_k^T F_k\|$ , Ma and Jiang [30] proposed the convex combination of these two choices, i.e.  $\lambda_k = \theta \|F_k\| + (1 - \theta) \|J_k^T F_k\|$  with  $\theta \in [0, 1]$ . And some other choices are given [31–33]. To make the LM parameter belongs to an interval which is proportional to  $\|F_k\|$ , some algebraic rules for the LM parameter also have been given [34, 35].

The choice of the LM parameter may cause some unsatisfactory properties when the sequence  $\{x_k\}$  generated by the LM method is close to or far away from the solution set, Fan et al. [36] and Chen [37] extended the choice  $\lambda_k = \|F_k\|^2$  to  $\lambda = \|F_k\|^\delta$  with  $\delta \in [1, 2]$ . Inspired by those observations, we consider our choice as

$$\lambda_k = \theta \|F_k\|^\delta + (1 - \theta) \|J_k^T F_k\|^\delta, \tag{5}$$

where  $\delta \in [1, 2]$  and  $\theta \in [0, 1]$ . It’s clear that, if we take  $\delta = 1$ , the choice (5) will be that which took by Ma and Jiang in [30], and if  $\delta = \theta = 1$ , the choice (5) will reduce to that proposed by Fan and Yuan in [27], etc. It also can be considered as the convex combination of  $\|F_k\|^\delta$  and  $\|J_k^T F_k\|^\delta$ . Therefore, choice (5) can be considered as the extension of many other choices.

Now, we consider the following modified LM algorithm with the LM parameter as (5) for the system of nonlinear Eq. (1).

---

**Algorithm 1** The modified LM algorithm

---

- Require:**  $x_0 \in X; \delta \in [1, 2]; \theta \in [0, 1]$ .  
 1:  $k \leftarrow 0; \lambda_0 \leftarrow \theta \|F_0\|^\delta + (1 - \theta) \|J_0^T F_0\|^\delta$ ;  
 2: **while**  $\|J_k^T F_k\| > 0$  **do**  
 3:   Solve the equation (2) to obtain  $d_k$ ;  
 4:    $x_{k+1} \leftarrow x_k + d_k$ ;  
 5:   Update  $\lambda_k$  by computing (5);  
 6:    $k \leftarrow k + 1$ ;  
 7: **end while**
- 

To generalize the theoretical application of the modified LM algorithm 1, we analyse the convergence of our new LM method under the  $\gamma$ -Hölderian local error bound (4) instead of the commonly used local error bound condition (3).

### 3 Convergence rate

The following assumptions are needed to establish convergence of Algorithm 1 throughout the paper.

**Assumption 1** The following conditions hold for some  $x_* \in X_*$ .

- (a)  $F(x)$  admits a  $\gamma$ -Hölderian local error bound property in the neighbourhood of  $x_* \in X_*$ , where  $\gamma \in (0, 1]$ , i.e., there exists positive constants  $c > 0$  that makes

$$c \operatorname{dist}(x, X_*) \leq \|F(x)\|^\gamma, \quad \forall x \in N(x_*, b). \tag{6}$$

(b)  $J(x)$  is  $v$ -Hölderian continuous, where  $v \in (0, 1]$ , i.e., there exists a constant  $\kappa_1 > 0$  such that

$$\|J(x) - J(y)\| \leq \kappa_1 \|x - y\|^v, \quad \forall x, y \in N(x_*, b). \tag{7}$$

**Remark 1** If  $\gamma = 1$ , then the  $\gamma$ -Hölderian local error bound condition (6) will be the local error bound condition (3). If  $v = 1$ , then the  $v$ -Hölderian continuity of the Jacobian (7) will be Lipschitz continuity of the Jacobian.

It follows from Assumption 1 and the mean value theorem that

$$\begin{aligned} \|F(y) - F(x) - J(x)(y - x)\| &= \left\| \int_0^1 J(x + t(y - x))(y - x) dt - J(x)(y - x) \right\| \\ &\leq \|y - x\| \int_0^1 \|J(x + t(y - x)) - J(x)\| dt \\ &\leq \frac{\kappa_1}{1 + v} \|y - x\|^{1+v} \quad \text{for all } x, y \in N(x_*, b). \end{aligned} \tag{8}$$

It can also be seen that there is a constant  $\kappa_2 > 0$  such that

$$\begin{aligned} \|J(x)\| &\leq \kappa_2, \quad \forall x \in N(x_*, b). \\ \|F(y) - F(x)\| &\leq \kappa_2 \|y - x\|, \quad \forall x, y \in N(x_*, b). \end{aligned} \tag{9}$$

Suppose that Algorithm 1 starts with an initial  $x_0$  that is sufficiently close to  $x_*$  such that the sequence  $\{x_k\}_{k=0}^\infty$  stays in the feasible set  $X_*$  for  $F(x)$ . It is worth emphasizing that  $x_*$  may not be an isolated zero of  $F(x)$ , and in the end  $x_k$  may converge to some point in  $X_*$  if it converges. Denote by  $\tilde{x} \in X_*$  which satisfies

$$\|\tilde{x} - x\| = \text{dist}(x, X_*). \tag{10}$$

Now, we will study the local convergence of our modified LM method, i.e., Algorithm 1.

Denote by

$$d(x) = - \left( J(x)^T J(x) + \lambda(x)I \right)^{-1} J(x)^T F(x) \tag{11}$$

and

$$\lambda(x) = \theta \|F(x)\|^\delta + (1 - \theta) \|J(x)^T F(x)\|^\delta. \tag{12}$$

**Lemma 2** *Let Assumption 1 hold. If  $v \geq 2(1/\gamma - 1)$ , then there exist some positive constants  $c_1, c_2$  such that*

$$c_1 \text{dist}(x, X_*)^{\delta/\gamma} \leq \lambda(x) \leq c_2 \text{dist}(x, X_*)^\delta, \quad \forall x \in N(x^*, b). \tag{13}$$

**Proof** First we will show the second inequality in (13). From (9), we have

$$\begin{aligned} \lambda(x) &= \theta \|F(x)\|^\delta + (1 - \theta) \|J(x)^T F(x)\|^\delta \leq (\theta + (1 - \theta) \|J(x)\|^\delta) \|F(x)\|^\delta \\ &= (\theta + (1 - \theta) \|J(x)\|^\delta) \|F(\bar{x}) - F(x)\|^\delta \leq (\theta + (1 - \theta) \|J(x)\|^\delta) \kappa_2^\delta \|\bar{x} - x\|^\delta \\ &\leq \kappa_2^\delta (\theta + (1 - \theta) \kappa_2^\delta) \|\bar{x} - x\|^\delta. \end{aligned}$$

This proves the second inequality in (13) with  $c_2 = \kappa_2^\delta (\theta + (1 - \theta) \kappa_2^\delta)$ .

Now, we show the first inequality in (13). From

$$\begin{aligned} \|F(x)\|^2 &= F(x)^T F(x) \\ &= F(x)^T (F(\bar{x}) + J(x)(x - \bar{x})) + F(x)^T (F(x) - F(\bar{x}) - J(x)(x - \bar{x})), \end{aligned}$$

we have

$$F(x)^T J(x)(x - \bar{x}) = \|F(x)\|^2 - F(x)^T (F(x) - F(\bar{x}) - J(x)(x - \bar{x})).$$

From the properties of the norm, we get

$$\|F(x)^T J(x)(x - \bar{x})\| \geq \|F(x)\|^2 - \|F(x)^T\| \|F(x) - F(\bar{x}) - J(x)(x - \bar{x})\|. \tag{14}$$

Since  $v \geq 2(1/\gamma - 1)$ , it follows from (6), (8), (9) and (14) that

$$\|J(x)^T F(x)\| \|\bar{x} - x\| \geq c^{2/\gamma} \|\bar{x} - x\|^{2/\gamma} - \frac{\kappa_1 \kappa_2}{1 + v} \|\bar{x} - x\|^{2+v}.$$

Thus, we have

$$\|J(x)^T F(x)\| \geq c^{2/\gamma} \|\bar{x} - x\|^{2/\gamma-1} - \frac{\kappa_1 \kappa_2}{1 + v} \|\bar{x} - x\|^{1+v} \geq c_3 \|\bar{x} - x\|^{2/\gamma-1}$$

with  $c_3 > 0$  is some positive constant. Hence

$$\begin{aligned} \lambda(x) &= \theta \|F(x)\|^\delta + (1 - \theta) \|J(x)^T F(x)\|^\delta \\ &\geq \theta c^{\delta/\gamma} \|\bar{x} - x\|^{\delta/\gamma} + c_3 (1 - \theta) \|\bar{x} - x\|^{\delta(2/\gamma-1)} \\ &\geq c_1 \|\bar{x} - x\|^{\delta/\gamma}. \end{aligned}$$

This proves the first inequality in (13) with  $c_1$  being a positive constant. □

**Lemma 3** *Let Assumption 1 hold, then the following inequality holds with a positive constant  $c_4 > 0$ .*

$$\|d(x)\| \leq c_4 \text{dist}(x, X^*)^{\min\{1, 1+v-\delta/2\gamma\}}, \quad \forall x \in N(x^*, b/2). \tag{15}$$

**Proof** For all  $x \in N(x^*, b/2)$ , we get

$$\|\bar{x} - x^*\| \leq \|\bar{x} - x\| + \|x - x^*\| \leq b. \tag{16}$$

Thus,  $\bar{x} \in N(x^*, b)$ .

Define the quadratical convex function  $\varphi(d)$  as

$$\varphi(d) = \|F(x) + J(x)d\|^2 + \lambda(x)\|d\|^2. \tag{17}$$

Hence,  $d(x)$  is the minimizer of  $\varphi(d)$ . By Assumption 1, Lemma 2, (16) and (17), we get

$$\begin{aligned} \|d(x)\|^2 &\leq \frac{\varphi(d(x))}{\lambda(x)} \leq \frac{\varphi(\bar{x} - x)}{\lambda(x)} = \frac{\|F(x) + J(x)(\bar{x} - x)\|^2 + \lambda(x)\|\bar{x} - x\|^2}{\lambda(x)} \\ &= \frac{\|F(\bar{x}) - F(x) - J(x)(\bar{x} - x)\|^2}{\lambda(x)} + \|\bar{x} - x\|^2 \\ &\leq \frac{\kappa_1^2}{c_1(1+v)^2} \|\bar{x} - x\|^{2(1+v)-\delta/\gamma} + \|\bar{x} - x\|^2 \\ &\leq \left( \frac{\kappa_1^2}{c_1(1+v)^2} + 1 \right) \|\bar{x} - x\|^{2\min\{1, 1+v-\delta/2\gamma\}}. \end{aligned}$$

Let  $c_4 = \sqrt{\kappa_1^2/c_1(1+v)^2 + 1}$ , we obtain (15). □

**Lemma 4** Assume  $x, x + d(x) \in N(x_*, b/2)$ . Then the following inequality holds with a positive constant  $c_5$ .

$$\text{dist}(x + d(x), X_*) \leq c_5 \text{dist}(x, X_*)^{\gamma \min\{1+v, 1+\delta/2, 2(1+v-\delta/2\gamma)\}}. \tag{18}$$

**Proof** From the definition of  $\varphi(d)$  (17) and Lemma 2, we can show

$$\begin{aligned} \|F(x) + J(x)d(x)\|^2 &\leq \varphi(d(x)) \leq \varphi(\bar{x} - x) \leq \|F(x) + J(x)(\bar{x} - x)\|^2 + \lambda_k \|\bar{x} - x\|^2 \\ &\leq \frac{\kappa_1^2}{(1+v)^2} \|\bar{x} - x\|^{2(1+v)} + c_2 \|\bar{x} - x\|^{2+\delta} \\ &\leq \left( \frac{\kappa_1^2}{(1+v)^2} + c_2 \right) \|\bar{x} - x\|^{2\min\{1+v, 1+\delta/2\}}. \end{aligned}$$

Then we have

$$\|F(x) + J(x)d(x)\| \leq c_6 \|\bar{x} - x\|^{\min\{1+v, 1+\delta/2\}}, \tag{19}$$

where positive constant  $c_6 = \sqrt{\kappa_1^2/(1+v)^2 + c_2}$ .

It follows from Assumption 1, Lemma 3 and (19) that

$$\begin{aligned} c \operatorname{dist}(x + d(x), X_*) &\leq \|F(x + d(x))\|^\gamma \leq \left( \|F(x) + J(x)d(x)\| + \kappa_1 \|d(x)\|^2 \right)^\gamma \\ &\leq \left( c_6 \|\tilde{x} - x\|^{\min\{1+v, 1+\delta/2\}} + \kappa_1 c_4^2 \|\tilde{x} - x\|^{2 \min\{1, 1+v-\delta/2\}} \right)^\gamma \\ &\leq \left( c_6 + \kappa_1 c_4^2 \right)^\gamma \|\tilde{x} - x\|^{\gamma \min\{1+v, 1+\delta/2, 2(1+v-\delta/2)\}}. \end{aligned}$$

Hence the inequality (18) holds with  $c_5 = c^{-1} (c_6 + \kappa_1 c_4^2)^\gamma$ . □

It is clear that, from (11) and (12), we have  $d_k = d(x_k)$ ,  $\lambda_k = \lambda(x_k)$  and  $x_{k+1} = x_k + d_k$ . Then it follows from Lemmas 2, 3 and 4 that, at every iteration,

$$\begin{aligned} c_1 \operatorname{dist}(x_k, X^*)^{\delta/\gamma} &\leq \lambda_k \leq c_2 \operatorname{dist}(x_k, X^*)^\delta, \\ \|d_k\| &\leq c_4 \operatorname{dist}(x_k, X^*)^{\min\{1, 1+v-\delta/2\}}, \\ \operatorname{dist}(x_{k+1}, X^*) &\leq c_5 \operatorname{dist}(x_k, X^*)^{\gamma \min\{1+v, 1+\delta/2, 2(1+v-\delta/2)\}} \end{aligned}$$

hold with  $x_k, x_{k+1} \in N(x_*, b/2)$ .

**Remark 2** Let

$$\begin{aligned} \alpha &= \min\{\gamma(1 + v), \gamma(1 + \delta/2), 2\gamma(1 + v) - \delta\}, \\ \beta &= \min\{1, 1 + v - \delta/2\gamma\}. \end{aligned}$$

To analyse the convergence rate of the sequence  $\{x_k\}$  generated by Algorithm 1, the value of above  $\alpha$  and  $\beta$  in Lemmas 3 and 4 should be discussed. Obviously, a larger value of  $\alpha$  will lead to derive a better convergence rate. Also, to deduce a convergence result, one needs to have  $\alpha > 1$ . This holds if and only if

$$\gamma > \frac{1}{1 + v} \quad \text{and} \quad 2 \left( \frac{1}{\gamma} - 1 \right) < \delta < 2\gamma(1 + v) - 1, \tag{20}$$

which imposes an additional requirement on the values of  $\gamma, v$  and  $\delta$ .

Since  $\alpha > 1$ , we have that  $\beta\alpha^k > k$  for sufficiently large  $k$ . As  $\sum_{l=0}^\infty \left(\frac{1}{2}\right)^l = 2$ , we deduce that

$$c_7 = \sum_{l=0}^\infty \left(\frac{1}{2}\right)^{\beta\alpha^l} < \infty.$$

Let

$$r = \min \left\{ \frac{1}{2c_5^{1/(\alpha-1)}}, \frac{b}{2 \left( 1 + 2c_4c_5^{(1-\beta)/(\alpha-1)}c_7 \right)} \right\}, \tag{21}$$



where  $c_4, c_5, c_7, \alpha, \beta$  are defined above.

Now, we show that if the initial point  $x_0$  in Algorithm 1 is close enough to  $X_*$ , the assumption of Lemma 4 will be satisfied.

**Lemma 5** *Assume (20) holds. If the initial point  $x_0 \in N(x_*, r)$ , where  $r$  is given by (21), then we have  $x_k \in N(x_*, b/2)$  for every  $k$ .*

**Proof** The proof is conducted by induction of  $k$ . We start with  $k = 0$ . By the assumption, we have  $x_0 \in N(x_*, r)$ . Since  $r \leq b/2$ , then  $x_l \in N(x_*, b/2)$  for all  $l = 0, 1, \dots, k$ . In order to show that  $x_{k+1} \in N(x_*, b/2)$ , we first from Lemma 3 note that

$$\begin{aligned} \|x_{k+1} - x_*\| &= \|x_k + d_k - x_*\| \leq \|x_k - x_*\| + \|d_k\| \leq \|x_{k-1} - x_*\| + \|d_{k-1}\| + \|d_k\| \\ &\leq \dots \leq \|x_0 - x_*\| + \sum_{l=0}^k \|d_l\| \leq r + c_4 \sum_{l=0}^k \text{dist}(x_l, X_*)^\beta. \end{aligned} \tag{22}$$

By Lemma 4, since  $r \leq 1/2c_5^{1/(\alpha-1)}$  defined in (21), we get

$$\begin{aligned} \text{dist}(x_l, X_*) &\leq c_5 \text{dist}(x_{l-1}, X_*)^\alpha \leq c_5 c_5^\alpha \text{dist}(x_{l-2}, X_*)^{\alpha^2} \\ &\leq \dots \\ &\leq c_5 c_5^\alpha \dots c_5^{\alpha^{l-1}} \text{dist}(x_0, X_*)^{\alpha^l} = c_5^{1+\alpha+\dots+\alpha^{l-1}} \text{dist}(x_0, X_*)^{\alpha^l} \\ &\leq c_5^{\frac{\alpha^l-1}{\alpha-1}} \|x_0 - x_*\|^{\alpha^l} \leq c_5^{\frac{\alpha^l-1}{\alpha-1}} r^{\alpha^l} \leq r \left( c_5^{\frac{1}{\alpha-1}} r \right)^{\alpha^l-1} \\ &\leq r \left( \frac{1}{2} \right)^{\alpha^l-1}, \end{aligned} \tag{23}$$

where  $l = 0, 1, \dots, k$ .

By (21), (22) and (23), we have

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq r + c_4 \sum_{l=0}^k \text{dist}(x_l, X_*)^\beta \leq r + c_4 \sum_{l=0}^k \left( r \left( \frac{1}{2} \right)^{\alpha^l-1} \right)^\beta \\ &= r + c_4 (2r)^\beta \sum_{l=0}^k \left( \frac{1}{2} \right)^{\beta\alpha^l} \leq r \left( 1 + 2^\beta c_4 r^{\beta-1} \sum_{l=0}^\infty \left( \frac{1}{2} \right)^{\beta\alpha^l} \right) \\ &\leq r \left( 1 + 2c_4 c_5^{(1-\beta)/(\alpha-1)} c_7 \right) \leq \frac{b}{2}. \end{aligned}$$

The above inequality indicates that  $x_{k+1} \in N(x_*, b/2)$ . □

From Lemmas 4 and 5 that we obtain the following main theorem.

**Theorem 6** *Under Assumption 1, if (20) holds, then the sequence  $\{\text{dist}(x_k, X_*)\}$  super-linearly converges to zero with order  $\min\{\gamma(1 + v), \gamma(1 + \delta/2), 2\gamma(1 + v) - \delta\}$ .*

Moreover, the sequence  $\{x_k\}$  generated by Algorithm 1 with initial point  $x_0 \in N(x_*, r)$  converges to a solution  $\tilde{x} \in X_* \cap N(x_*, b/2)$  of nonlinear Eq. (1).

**Proof** The first part of this theorem, that is the convergence rate of sequence  $\{\text{dist}(x_k, X_*)\}$ , is clearly derived by Lemmas 4 and 5. So, only the proof of second part should be given.

From the hypothesis, for every  $k$ , we get  $x_k \in N(x_*, b/2)$ . Hence, we just have to show that  $\{x_k\}$  converges. From Lemma 3 and (23), we have

$$\|d_k\| \leq c_4 \text{dist}(x_k, X_*)^\beta \leq c_4 r^\beta \left(\frac{1}{2}\right)^{\beta(\alpha^l - 1)}. \tag{24}$$

Denoting by

$$s_k = \sum_{l=0}^k \|d_l\| \leq c_4 (2r)^\beta \sum_{l=0}^k \left(\frac{1}{2}\right)^{\beta\alpha^l} \leq c_4 c_5^{\frac{\beta}{1-\alpha}} \sum_{l=0}^k \left(\frac{1}{2}\right)^{\beta\alpha^l} \leq c_4 c_5^{\frac{\beta}{1-\alpha}} c_7 < \infty,$$

we obtain that  $\{s_k\}$  is a Cauchy sequence. Thus, we have

$$\begin{aligned} \|x_{k+p} - x_k\| &\leq \|d_{k+p-1}\| + \|x_{k+p-1} - x_k\| \\ &\leq \dots \leq \sum_{l=k}^{k+p-1} \|d_l\| \\ &= |s_{k+p-1} - s_{k-1}|, \end{aligned} \tag{25}$$

for any  $k, p \in \mathbb{N}$ . Which indicates that the sequence  $\{x_k\}$  is also a Cauchy sequence. Then,  $\{x_k\}$  converges to some  $\tilde{x}$ . Since  $x_k \in N(x_*, b/2)$  for all  $k$  and  $\{\text{dist}(x_k, X_*)\}$  convergence to zero, we have  $\tilde{x} \in X_* \cap N(x_*, b/2)$ .  $\square$

**Theorem 7** Under the assumptions of Theorem 6, if  $v \geq \delta/2\gamma$ , then sequence  $\{x_k\}$ , generated by Algorithm 1 with  $x_0 \in N(x_*, r)$ , converges to the solution  $\tilde{x}$  of the nonlinear Eq. (1) with order  $\gamma(1 + \delta/2)$ .

**Proof** It follows from  $v \geq \delta/2\gamma$  that

$$\begin{aligned} &\gamma(1 + v) - \gamma \left(1 + \frac{\delta}{2}\right) \\ &= \gamma \left(v - \frac{\delta}{2}\right) \geq \gamma \left(\frac{\delta}{2\gamma} - \frac{\delta}{2}\right) \\ &= \frac{\gamma\delta}{2} \left(\frac{1}{\gamma} - 1\right) \geq 0 \end{aligned}$$

and

$$\begin{aligned}
 & (2\gamma(1 + v) - \delta) - \gamma \left( 1 + \frac{\delta}{2} \right) \\
 &= \gamma + 2\gamma v - \delta - \frac{\gamma\delta}{2} \geq \gamma - \frac{\gamma\delta}{2} \\
 &= \frac{\gamma}{2}(2 - \delta) \geq 0.
 \end{aligned}$$

So, from the the assumptions of Theorem 6, we have

$$\alpha = \min\{\gamma(1 + v), \gamma(1 + \delta/2), 2\gamma(1 + v) - \delta\} = \gamma(1 + \delta/2) > 1. \tag{26}$$

Moreover, it follows from (26) that  $\text{dist}(x_{k+1}, X_*) \leq \frac{1}{2}\text{dist}(x_k, X_*)$  holds for all sufficiently large  $k$ . From  $v \geq \delta/2\gamma$ , we have  $\beta = 1$  in Lemma 3. By letting  $p \rightarrow \infty$  in (25), we deduce

$$\begin{aligned}
 \|\tilde{x} - x_k\| &\leq \sum_{l=k}^{\infty} \|d_l\| \leq c_4 \sum_{l=k}^{\infty} \text{dist}(x_l, X_*) \leq c_4 \sum_{l=k}^{\infty} \left(\frac{1}{2}\right)^{l-k} \text{dist}(x_k, X_*) \\
 &\leq 2c_4 \text{dist}(x_k, X_*) \leq 2c_4 c_5 \text{dist}(x_{k-1}, X_*)^\alpha \leq 2c_4 c_5 \|\tilde{x} - x_{k-1}\|^\alpha, \tag{27}
 \end{aligned}$$

which implies that  $\{x_k\}$  converges with rate  $\gamma(1 + \delta/2)$ . □

**Remark 3** (1) As it was shown in Theorem 7, if  $\gamma = v = 1$ , we have  $\alpha = 1 + \delta/2$ . From (27), we obtain

$$\|\tilde{x} - x_k\| \leq 2c_4 c_5 \|\tilde{x} - x_{k-1}\|^{1+\delta/2}$$

which shows that the sequence  $\{x_k\}$  at least converges superlinearly and converges quadratically while  $\delta = 2$ . This result is the same as the result in [20] with  $\delta = 2$  and  $\theta = 1$ .

(2) Under Assumption 1, let  $\gamma = v = 1$ . Algorithm 1 will reduce to the algorithm proposed by Fan and Yuan in [27] when  $\theta = 1, \delta = 1$ , the algorithm proposed by Ma and Jiang in [30, Algorithm 2.1] when  $\delta = 1$ , and the algorithm proposed by Fan and Yuan in [36] when  $\theta = 1, \delta \in [1, 2]$ . The results of Theorem 7 is also the same as theirs.

**Remark 4** To obtain fast convergence order, singular value decomposition (SVD) technology was employed by some papers [27, 30, 36], but we omit it here.

### 4 Global convergence

In this section, we discuss the global convergence properties of the modified LM method. To guarantee the global convergence, the Armijo line search technique is employed in the modified LM method.

Take

$$\Phi(x) = \frac{1}{2} \|F(x)\|^2$$

as the merit function for nonlinear equations (1).

We consider the following algorithm.

---

**Algorithm 2** The modified LM algorithm with line search

---

**Require:**  $x_0 \in X; \delta \in [1, 2]; \theta \in [0, 1]; \tilde{\alpha}, \tilde{\beta}, \tilde{\gamma} \in (0, 1)$ .  
 1:  $k \leftarrow 0; \lambda_0 \leftarrow \theta \|F_0\|^\delta + (1 - \theta) \|J_0^T F_0\|^\delta$ ;  
 2: **while**  $\|J_k^T F_k\| > 0$  **do**  
 3:     Solve the equation (2) to obtain  $d_k$ ;  
 4:     **if**  $\|F(x_k + d_k)\| \leq \tilde{\gamma} \|F_k\|$  **then**  
 5:          $x_{k+1} \leftarrow x_k + d_k$ ;  
 6:     **else**  
 7:          $m \leftarrow 0$ ;  
 8:         **while**  $\Phi(x_k + \tilde{\beta}^m d_k) - \Phi(x_k) > \tilde{\alpha} \tilde{\beta}^m \nabla \Phi(x_k)^T d_k$  **do**  
 9:              $m \leftarrow m + 1$ ;  
 10:         **end while**  
 11:          $x_{k+1} \leftarrow x_k + \tilde{\beta}^m d_k$ ;  
 12:     **end if**  
 13:     Update  $\lambda_k$  by computing (5);  
 14:      $k \leftarrow k + 1$ ;  
 15: **end while**

---

**Theorem 8** *Suppose Assumption 1 holds. Then the sequence  $\{x_k\}$  be generated by Algorithm 2 converges to a stationary point of  $\Phi(x)$ . Moreover, if the stationary point  $x_*$  is a solution of nonlinear Eq. (1), then  $\{x_k\}$  converges to the solution with convergence rate  $\min\{\gamma(1 + v), \gamma(1 + \delta/2), 2\gamma(1 + v) - \delta\}$ .*

**Proof** The Armijo line search technique is employed in Algorithm 2. That is, the next iteration will be computed might be

$$x_{k+1} = x_k + \tilde{\beta}^m d_k,$$

where  $\tilde{\beta}^m$  is a step size and  $m \geq 0$  is the smallest nonnegative integer that satisfies the following inequality.

$$\Phi(x_k + \tilde{\beta}^m d_k) - \Phi(x_k) \leq \tilde{\alpha} \tilde{\beta}^m \nabla \Phi(x_k)^T d_k.$$

Combine with [38, Eq. (2.10)], we get

$$\|F(x_{k+1})\|^2 \leq \|F_k\|^2 - \tilde{\alpha} \hat{\alpha} \frac{(F_k^T J_k d_k)^2}{\|d_k\|^2}, \tag{28}$$

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

Therefore, from the condition  $\|F(x_k + d_k)\| \leq \tilde{\gamma}\|F_k\|$  and (28), we obtain that  $\{\|F(x_k)\|\}$  is monotonically decreasing. Hence, the sequence  $\{x_k\}$  converges to a stationary point of  $\Phi(x)$ .

Next we show the remainder part. We deduce from the first part that  $\|F(x_k + d_k)\| \leq \tilde{\gamma}\|F_k\|$  holds for sufficient large  $k$ , that is,  $\tilde{\beta}^m = 1$  holds for all sufficient large  $k$ . Since  $\{x_k\}$  converges to a stationary point  $x_*$  which is a solution of nonlinear Eq. (1), there exists a large  $K$  such that

$$\|F(x_K)\|^{\alpha\gamma-1} \leq \frac{c^\alpha \tilde{\gamma}}{\kappa_2 c_5}, \tag{29}$$

and

$$\|x_K - x_*\| \leq r$$

where  $c, \kappa_2$  and  $c_5$  are defined in Sect. 3,  $r$  is defined by (21).

Let sequence  $\{y_k\}$  be generated by Algorithm 1 with unit step size and  $y_0 = x_K$ . Then, by Theorem 6, the sequence  $\text{dist}(y_l, X_*)$  converges to zero with convergence rate  $\min\{\gamma(1+v), \gamma(1+\delta/2), 2\gamma(1+v)-\delta\}$ . So, we just have to prove that  $x_{K+l} = y_l$  holds for each  $l$ , i.e.  $\{y_l\}$  satisfies

$$\|F(y_{l+1})\| \leq \tilde{\gamma}\|F(y_l)\|.$$

Let  $\bar{y}_{l+1} \in X_*$  such that  $\|y_{l+1} - \bar{y}_{l+1}\| = \text{dist}(y_{l+1}, X_*)$ . It follows from Assumption 1 (a), Lemma 4, (9) and (29) that

$$\begin{aligned} \|F(y_{l+1})\| &= \|F(y_{l+1}) - F(\bar{y}_{l+1})\| \\ &\leq \kappa_2 \text{dist}(y_{l+1}, X_*) \leq \kappa_2 c_5 \text{dist}(y_l, X_*)^\alpha \\ &\leq \frac{\kappa_2 c_5 \|F(y_l)\|^{\alpha\gamma-1}}{c^\alpha} \|F(y_l)\| \\ &\leq \tilde{\gamma}\|F(y_l)\| \end{aligned} \tag{30}$$

holds for  $\tilde{\gamma} < 1$  and each  $l$ . (30) implies that  $\tilde{\beta}^m = 1$  holds for all sufficient large  $k$  in Algorithm 2. Thus, we have  $\{x_k\}$  converges to the solution with convergence rate  $\min\{\gamma(1+v), \gamma(1+\delta/2), 2\gamma(1+v)-\delta\}$ . □

### 5 Numerical example

Some numerical experiments are carried out to verify Algorithm 1 proposed in Sect. 3 in this section. We first test the following Powell singular function [21] proposed in Sect. 1.

$$F(x) = \left(x_1 + 10x_2, \sqrt{5}(x_3 - x_4), (x_2 - 2x_3)^2, \sqrt{10}(x_1 - x_4)^2\right)^T, \tag{31}$$

where  $x \in \mathbb{R}^4$ . As proposed in Sect. 1, the local error bound (3) does not satisfy (31) but the  $\gamma$ -Hölderian local error bound (4) does.

The other test problems are systems of nonlinear equations created by the modifying singular problems proposed in [21, 39],

$$\hat{F}(x) = F(x) - J(x_*)A \left( A^T A \right)^{-1} A^T (x - x_*), \tag{32}$$

where  $x_*$  is its root, function  $F(x)$  is the standard nonsingular function, and matrix  $A \in \mathbb{R}^{n \times k}$  has full column rank with  $1 \leq k \leq n$ . It is easy to check that  $\hat{F}(x_*) = 0$  and  $\hat{J}(x_*) = J(x_*) \left( I - A \left( A^T A \right)^{-1} A^T \right)$  has rank  $n - k$ .

We chose the rank of  $\hat{J}(x_*)$  to be  $n - 1$  by using

$$A \in \mathbb{R}^{n \times 1}, \quad A^T = (1, 1, \dots, 1) \tag{33}$$

and  $n - 2$  by using

$$A \in \mathbb{R}^{n \times 2}, \quad A^T = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & -1 & 1 & -1 & \dots & \pm 1 \end{pmatrix} \tag{34}$$

respectively.

We test several choices of the LM parameter in Algorithm 1. Based on the range of  $\theta$  and  $\delta$  defined in (5), we use  $\theta = 0, 0.5$  and  $1$  and  $\delta = 1, 1.5$  and  $2$  for the LM method with unit step size in Sect. 3. The algorithm is terminated when the norm of  $\|J_k^T F_k\| \leq 10^{-6}$ , or when the number of the iterations exceeds  $100(n + 1)$ . However, for Powell singular function, we take the number  $10^6$  instead of  $100(n + 1)$ . The results for the Powell singular function are tabulated in Table 1, while the results for systems of nonlinear equations with rank  $n - 1$  and  $n - 2$  case are tabulated in Tables 2 and 3, respectively. We test three starting points  $x_0, 10x_0$  and  $100x_0$  where  $x_0$  is suggested by Moré, Garbow and Hillstrome in [21]. The notation “–” is used if the method fails to find the solution within the maximum iterations.

The results in Table 1 show that Algorithm 1 is efficient for Powell singular function even it does not satisfy the local error bound (3). But it is not perfect because of the big number of the iteration when initial point is not close enough to  $x_*$ . From the results in the all tables, we can see that Algorithm 1 is always outstanding when  $\delta = 1$  no matter what the value of  $\theta$  is. In fact, when  $\{x_k\}$  is close to the solution set  $X_*$ , the LM parameter  $\lambda_k$  with  $\delta = 2$  may be very small, even less than the machine accuracy, and it will lose its role. Conversely, when  $\{x_k\}$  is far away from the solution set  $X_*$ , the LM parameter  $\lambda_k$  with  $\delta = 2$  may be very large, and the trial step  $d_k$  will be small, then it prevents  $\{x_k\}$  from converging rapidly. Also Algorithm 1 with  $\lambda_k = \|F_k\|^\delta$  is always outperforms or at least performs as well as Algorithm 1 with  $\lambda_k = \|J_k^T F_k\|^\delta$ .

**Table 1** Results on Powell singular function

Function	$n$	$x_0$	Iters											
			$\theta = 0$			$\theta = 0.5$			$\theta = 1$			$\theta = 2$		
			$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$
Powell singular	4	1	15	27	122	15	21	70	14	13	15	14	13	15
		10	64	3004	80.7724	50	1554	404107	34	100	485	34	100	485
		100	510	876.402	–	356	439.534	–	199	2660	48.909	199	2660	48.909

**Table 2** Results on the first singular test set with  $\text{rank}(F'(x_*)) = n - 1$

Function	$n$	$x_0$	Iters											
			$\theta = 0$				$\theta = 0.5$				$\theta = 1$			
			$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	
Rosenbrock	2	1	31	36	151	29	31	97	42	26	23	64	124	
		10	64	-	-	62	-	-	62	64	124	293	-	
		100	293	-	-	266	-	-	233	-	-	61	76	
Powell badly scaled	2	1	46	-	-	41	-	-	61	35	76	49	29	
		10	43	-	-	54	-	-	49	34	29	24	27	
		100	26	272	-	27	151	-	24	20	27	30	-	
Wood	4	1	79	-	-	56	-	-	30	39	-	61	392	
		10	-	-	-	319	-	-	61	392	-	17	14	
		100	22	75	-	23	46	-	17	14	15	21	25	
Helical valley	3	1	37	329	-	34	180	-	21	25	67	31	213	
		10	138	-	-	89	-	-	31	213	-	8	9	
		100	9	11	17	9	10	13	8	8	9	22	22	
Brown almost-linear	10	1	35	-	-	29	-	-	22	22	34	44	44	
		10	200	-	-	123	-	-	44	44	-	2	2	
		100	3	4	2	3	2	2	3	2	2	310	61	
Discrete Boundary value	10	1	21	14	14	162	38	18	99	147	514	12	5	
		10	62	405	-	80	276	-	12	7	5	86	42	
		100	7	6	5	8	6	5	5	5	5	48	42	
Discrete integral equation	30	1	24	18	17	53	33	30	27	99	1391	27	27	
		10	97	-	-	62	-	-	27	99	1391	27	27	
		100	-	-	-	-	-	-	-	-	-	-	-	



Table 2 continued

Function	$n$	$x_0$	Iters											
			$\theta = 0$			$\theta = 0.5$			$\theta = 1$			$\theta = 2$		
			$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$
Variably dimensioned	10	1	16	-	-	14	604	-	13	13	14	13	13	14
		10	35	-	-	26	-	-	15	26	111	15	26	111
		100	249	-	-	144	-	-	35	431	-	35	431	-
Broyden tridiagonal	30	1	12	11	13	13	11	12	13	10	10	13	10	10
		10	66	-	-	45	1958	-	22	57	667	22	57	667
		100	564	-	-	297	-	-	29	469	-	29	469	-
Broyden banded	30	1	15	25	118	14	19	66	13	12	13	13	12	13
		10	71	-	-	46	-	-	19	26	520	19	26	520
		100	571	-	-	299	-	-	25	59	-	25	59	-

**Table 3** Results on the second singular test set with  $\text{rank}(F'(x_*)) = n - 2$

Function	$n$	$x_0$	Iters											
			$\theta = 0$				$\theta = 0.5$				$\theta = 1$			
			$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	
Rosenbrock	2	1	12	53	-	11	32	-	10	10	11	10	10	
		10	27	-	-	20	-	-	13	15	57	15	57	
		100	139	-	-	78	-	-	16	36	-	36	-	
Powell badly scaled	2	1	24	14	14	241	79	19	-	186	34	-	34	
		10	2	2	2	2	2	2	2	93	13	2	13	
		100	2	3	57	2	3	32	2	2	2	2	2	
Wood	4	1	20	313	-	17	165	-	13	14	22	14	22	
		10	62	-	-	40	-	-	16	25	-	25	-	
		100	448	-	-	235	-	-	20	102	-	102	-	
Helical valley	3	1	26	67	-	27	48	-	28	21	20	28	20	
		10	40	336	-	41	189	-	34	33	73	34	73	
		100	166	-	-	131	-	-	82	307	-	307	-	
Brown almost-linear	10	1	9	11	17	9	10	13	8	8	9	8	9	
		10	35	-	-	29	-	-	22	22	34	22	34	
		100	200	-	-	123	-	-	42	44	-	44	-	
Discrete Boundary value	10	1	3	4	2	3	2	2	3	2	2	3	2	
		10	21	14	14	162	38	18	310	60	22	310	22	
		100	63	405	-	87	278	-	112	150	516	112	516	

Table 3 continued

Function	$n$	$x_0$	Iters											
			$\theta = 0$			$\theta = 0.5$			$\theta = 1$			$\theta = 2$		
			$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$
Discrete integral equation	30	1	7	6	5	8	6	5	12	7	5	12	7	5
		10	24	18	17	53	33	30	86	48	42	86	48	42
		100	112	-	-	122	-	-	136	161	1401	136	161	1401
Variably dimensioned	10	1	16	-	-	14	604	-	13	13	14	13	13	14
		10	35	-	-	26	-	-	15	25	108	15	25	108
		100	246	-	-	143	-	-	35	421	-	35	421	-
Broyden tridiagonal	30	1	12	11	13	13	11	12	13	10	10	13	10	10
		10	66	-	-	45	1958	-	22	57	667	22	57	667
		100	564	-	-	297	-	-	29	469	-	29	469	-
Broyden banded	30	1	15	25	118	14	19	66	13	12	13	13	12	13
		10	71	-	-	46	-	-	19	26	520	19	26	520
		100	571	-	-	299	-	-	25	59	-	25	59	-

## 6 Conclusion

We have presented a modified Levenberg–Marquardt method for solving system of nonlinear equations with adaptive choice of LM parameter which is a convex combination of  $\|F_k\|^\delta$  and  $\|J_k^T F_k\|^\delta$ . Under the  $\gamma$ -Hölderian local error bound condition of the underlying function and the  $\nu$ -Hölderian continuity of its Jacobian, its convergence including local and global has been analysed. These convergence properties hold in many applied problems, as they are satisfied by any real differentiable function. In our numerical experiments, we clearly obtained a superior performance of our choice of LM parameter.

**Acknowledgements** This research was supported by NFSC Grant 11901061, the University Natural Science Research Project of Anhui Province Grant KJ2020ZD008 and the Natural Science Foundation of Anhui province Grant 2108085MF204. Part of this work was completed during authors' visit to University of Texas at Arlington. They would like to thank Prof. Ren-Cang Li for his hospitality during the visit.

## Declarations

**Conflict of interest** The authors declare that they have no known conflict of interest.

## References

1. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, Heidelberg (1999)
2. Sun, W., Yuan, Y.: Optimization Theory and Methods: Nonlinear Programming. Springer, Heidelberg (2006)
3. Andrei, N.: Modern Numerical Nonlinear Optimization. Springer, Cham (2022)
4. Xie, L., Ding, J., Ding, F.: Gradient based iterative solutions for general linear matrix equations. Comput. Math. Appl. **58**(7), 1441–1448 (2009)
5. Xie, L., Liu, Y., Yang, H.: Gradient based and least squares based iterative algorithms for matrix equations  $AXB + CX^T D = F$ . Appl. Math. Comput. **217**(5), 2191–2199 (2010)
6. Li, M., Liu, X.: Iterative identification methods for a class of bilinear systems by using the particle filtering technique. Int. J. Adapt. Control Signal Process. **35**(10), 2056–2074 (2021)
7. Ding, F., Chen, T.: Parameter estimation of dual-rate stochastic systems by using an output error method. IEEE Trans. Automat. Control **50**(9), 1436–1441 (2005)
8. Li, M., Liu, X.: Maximum likelihood hierarchical least squares-based iterative identification for dual-rate stochastic systems. Int. J. Adapt. Control Signal Process. **35**(2), 240–261 (2020)
9. Ding, F., Ling, X., Meng, D., Jin, X.-B., Alsaedi, A., Hayat, T.: Gradient estimation algorithms for the parameter identification of bilinear systems using the auxiliary model. J. Comput. Appl. Math. **369**, 112575 (2020)
10. Ding, F., Chen, T.: Combined parameter and output estimation of dual-rate systems using an auxiliary model. Automatica **40**(10), 1739–1748 (2004)
11. Liu, Y., Ding, F., Shi, Y.: An efficient hierarchical identification method for general dual-rate sampled-data systems. Automatica **50**(3), 962–970 (2014)
12. Ding, F., Liu, X., Ma, X.: Kalman state filtering based least squares iterative parameter estimation for observer canonical state space systems using decomposition. J. Comput. Appl. Math. **301**, 135–143 (2016)
13. Ding, F., Liu, X.P., Liu, G.: Identification methods for Hammerstein nonlinear systems. Digital Signal Process. **21**(2), 215–238 (2011)
14. Deepho, J., Abubakar, A.B., Malik, M., Argyros, I.K.: Solving unconstrained optimization problems via hybrid cd-dy conjugate gradient methods with applications. J. Comput. Appl. Math. **405**, 113823 (2022)

15. Chen, L., Ma, Y.: Shamanskii-like Levenberg–Marquardt method with a new line search for systems of nonlinear equations. *J. Syst. Sci. Complexity* **33**(5), 1694–1707 (2020)
16. Chen, L.: A modified Levenberg–Marquardt method with line search for nonlinear equations. *Comput. Optim. Appl.* **65**(3), 753–779 (2016)
17. Levenberg, K.: A method for the solution of certain nonlinear problems in least squares. *Q. Appl. Math.* **2**, 164–168 (1944)
18. Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **11**, 431–441 (1963)
19. Yuan, Y.: Recent advances in trust region algorithms. *Math. Program.* **151**(1), 249–281 (2015)
20. Yamashita, N., Fukushima, M.: On the rate of convergence of the Levenberg–Marquardt method. In: Alefeld, G., Chen, X. (eds.) *Topics in Numerical Analysis: With Special Emphasis on Nonlinear Problems*, pp. 239–249. Springer Vienna, Vienna (2001)
21. Moré, J.J., Garbow, B.S., Hillstrome, K.H.: Testing unconstrained optimization software. *ACM Trans. Math. Softw.* **7**(1), 17–41 (1981)
22. Ahookhosh, M., Aragón Artacho, F.J., Fleming, Ronan M. T., Vuong, P.T.: Local convergence of the Levenberg–Marquardt method under Hölder metric subregularity. *Adv. Comput. Math.* **45**(5–6), 2771–2806 (2019)
23. Wang, H., Fan, J.: Convergence rate of the Levenberg–Marquardt method under Hölderian local error bound. *Optim. Methods Softw.* **35**(4), 767–786 (2020)
24. Guo, L., Lin, G.-H., Jane, J.Y.: Solving mathematical programs with equilibrium constraints. *J. Optim. Theory Appl.* **166**(1), 234–256 (2015)
25. Zheng, L., Chen, L., Ma, Y.: A variant of the Levenberg–Marquardt method with adaptive parameters for systems of nonlinear equations. *AIMS Math.* **7**(1), 1241–1256 (2022)
26. Zheng, L., Chen, L., Tang, Y.: Convergence rate of the modified Levenberg–Marquardt method under Hölderian local error bound. *Open Math.* **20**(1), 998–1012 (2022)
27. Fan, J., Yuan, Y.: On the convergence of a new Levenberg–Marquardt method. In *Technical Report, AMSS, Chinese Academy of Sciences* (2001)
28. Dennis, J.E., Jr., Schnable, R.B.: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics, Philadelphia (1983)
29. Fischer, A.: Local behavior of an iterative framework for generalized equations with nonisolated solutions. *Math. Program.* **94**, 91–124 (2002)
30. Ma, C., Jiang, L.: Some research on Levenberg–Marquardt method for the nonlinear equations. *Appl. Math. Comput.* **184**(2), 1032–1040 (2007)
31. Fan, J., Pan, J.: A note on the Levenberg–Marquardt parameter. *Appl. Math. Comput.* **207**(2), 351–359 (2009)
32. Huang, B., Ma, C.: A shamanskii-like self-adaptive Levenberg–Marquardt method for nonlinear equations. *Comput. Math. Appl.* **77**(2), 357–373 (2019)
33. Amini, K., Rostami, F., Caristi, G.: An efficient Levenberg–Marquardt method with a new LM parameter for systems of nonlinear equations. *Optimization* **67**(5), 637–650 (2018)
34. Karas, E.W., Santos, S.A., Svaiter, B.F.: Algebraic rules for computing the regularization parameter of the Levenberg–Marquardt method. *Comput. Optim. Appl.* **65**(3), 1–29 (2016)
35. Musa, Y.B., Waziri, M.Y., Halilu, A.S.: On computing the regularization parameter for the Levenberg–Marquardt method via the spectral radius approach to solving systems of nonlinear equations. *J. Numer. Math. Stochast.* **9**(1), 80–94 (2017)
36. Fan, J., Yuan, Y.: On the quadratic convergence of the Levenberg–Marquardt method without nonsingularity assumption. *Computing* **74**(1), 23–39 (2005)
37. Chen, L., Ma, Y., Su, C.: An efficient m-step Levenberg–Marquardt method for nonlinear equations. *ChinaXiv*, p. 15, (2016)
38. Yuan, Y.-X.: Problems on convergence of unconstrained optimization algorithms. In: *Numerical Linear Algebra and Optimization*, pp. 95–107. Science Press, Beijing, New York (1999)
39. Schnabel, R.B., Frank, P.D.: Tensor methods for nonlinear equations. *SIAM J. Numer. Anal.* **21**(5), 815–843 (1984)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.