RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2023.023

ScienceAsia 49 (2023): **670–677**

A two-step smoothing Levenberg-Marquardt method for nonlinear complementarity problems

Linsen Song[a,*], **Yan Gao**[b]

[a] School of Mathematical Sciences, Henan Institute of Science and Technology, Xinxiang 453003 China
[b] School of Management, University of Shanghai for Science and Technology, Shanghai 200093 China

[*]Corresponding author, e-mail: slinsen@163.com

Received 15 Mar 2022, Accepted 27 Nov 2022
Available online 12 Aug 2023

ABSTRACT: By reformulating the nonlinear complementarity problem into a system of nonsmooth equations, we proposed a two-step smoothing Levenberg-Marquardt method with global convergence in this paper, where not only one step but also an approximate step is computed at each iteration. Under local error bound condition, the local convergence result of the present algorithm is also obtained. Finally, numerical experiments illustrate the effectiveness of the present method when compare it with one-step smoothing Levenberg-Marquardt methods.

KEYWORDS: nonsmooth analysis, generalized Jacobi, smoothing Levenberg-Marquardt method, two-step iterations

MSC2020: 90C33

INTRODUCTION

Consider the nonlinear complementarity problem (NCP) such that

$$x \geqslant 0, \quad F(x) \geqslant 0, \quad x^{\mathrm{T}}F(x) \geqslant 0,$$

where $x \in \mathbb{R}^n$, $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable. It has a large number of important applications and has attracted many interests in last decades [1–4]. For NCP, a popular method is to reformulate it into a system of nonsmooth equations such that

$$H(x) = (\varphi(x_1, F_1(x)), \ldots, \varphi(x_n, F_n(x)))^{\mathrm{T}} = 0, \quad (1)$$

where $\varphi : \mathbb{R}^2 \to \mathbb{R}$ is an NCP function [5, 6]. And a Levenberg-Marquardt (LM) algorithm always solves it by introducing a perturbation item λ_k, then generating the iterative direction d_k by solving the equation such that

$$(J_k^{\mathrm{T}}J_k + \lambda_k I)d_k = -J_k^{\mathrm{T}}H(x_k), \quad (2)$$

where $\lambda_k > 0$ is called LM parameter, J_k denotes the generalized Jacobi of the function H in nonsmooth LM algorithms [7–9], the Jacobi of the function H_ε, which is a smoothing approximating function of H in smoothing LM algorithms. Since it is always difficult or time consuming to obtain the generalized Jacobi of the function H for nonsmooth algorithms, the smoothing ones for NCP have been paid more attention for a long time [10–12]. By using smoothing technique to convert NCP to smoothing nonlinear systems, Yu and Pu present a smoothing Levenberg—Marquardt method. And Under the local error bound condition, which is much weaker than nonsingularity assumption or the strictly complementarity condition, they obtain the local superlinear convergence [11]. By investigating an element of related functions' B-differential, a smoothing Levenberg-Marquardt was proposed based on a Chen-Harker-Kanzow-Smale (CHKS) smoothing function, which satisfies a property called strongly Jacobi consistency [12]. These algorithms are effective in numerical experiments. However, they are both based on one-step LM iterative step. Considering that the two-step technique has been adopted for complementarity problems in recent years, such as Zheng and Liu proposed a two-step modulus-based matrix splitting iteration method for solving a class of nonlinear complementarity problems [13], Cao and Wang investigated the implicit complementarity problems by a two-step modulus-based matrix splitting iteration method [14], a two-step modulus-based matrix splitting iteration method for the horizontal nonlinear complementarity problem was proposed in [15] and so on, maybe it is also effective when applied into the LM algorithms. And maybe it is valuable to investigate the function F in a general form, since these methods with two-step technology are based on the property of matrix, and the correlation functions are expressed in the form of matrices.

Actually, for a smooth equation such that

$$G(x) = 0,$$

the LM algorithm always has quadratic convergence if the Jacobi matrix of the smooth function $G : \mathbb{R}^n \to \mathbb{R}^n$ is nonsingular and Lipschitzian continuous. As the nonsingularity is a strong condition, Yamashita and Fukushima investigated the convergence rate under the local error bound condition [16] when the LM parameter λ_k is chosen as $\|G(x_k)\|^2$. Fan and Pan obtained the convergence order when $\lambda_k = \|G(x_k)\|^\delta$ for $\delta \in (0, 2]$ [17]. In [18], Fan introduced a two-step Levenberg-Marquardt method (MLM) with cubic convergence when $\delta \in [1, 2]$, where each iteration

$d_k = d_{k_1} + d_{k_2}$ is obtained by solving two linear systems, the one denoted as d_{k_1} is the solution of the equation in the form of Eq. (2), where $H(x_k)$ is replaced by $G(x_k)$, the other one d_{k_2} is obtained by solving the linear system such that

$$(J_k^{\mathrm{T}} J_k + \lambda_k I)d = -J_k^{\mathrm{T}} G(y_k),$$

where $y_k = x_k + d_{k_1}$. Amini and Rostami proposed a three-step modified Levenberg-Marquardt method (MMLM) with a new line search, where the addition of the LM step and two approximate LM steps are adopted as the trial step at every iteration [19], the convergence order of the algorithm is biquadratic. It is noticeable that both the MLM and the MMLM algorithms have better convergence and they do not need compute more, but only the Jacobi of $G(x_k)$ once in the k-th iteration. Inspired by these, a two-step smoothing LM method and its local convergence under the error bound condition for the nonlinear complementarity problem are investigated in this paper.

PRELIMINARIES

In this section, we recall some definitions and propositions necessary in this paper.

In nonsmooth analysis, there exists some generalized differentials with different forms. Let $G : \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitzian, Ω_G denotes a set where G is not differential, B-differential of G at $x \in \mathbb{R}^n$ is defined by

$$\partial_{\mathrm{B}} G(x) = \{\lim\, JG(x_i) : x_i \to x,\ x_i \notin \Omega_G\}.$$

Clarke generalized Jacobi of G at $x \in \mathbb{R}^n$ is defined as the convex hull of $\partial_{\mathrm{B}} G(x)$ such that

$$\partial_{\mathrm{Cl}} G(x) = \mathrm{conv}\ \partial_{\mathrm{B}} \mathrm{G(x)}.$$

Denote

$$\begin{aligned} H(x) &= H_{\min}(x) \\ &= (\varphi_{\min}(x_1, F_1(x)), \ldots, \varphi_{\min}(x_n, F_n(x)))^{\mathrm{T}}, \end{aligned} \quad (3)$$

in Eq. (1), we use the CHKS smoothing function

$$\varphi_\varepsilon(a,b) = \frac{a+b-\sqrt{\varepsilon^2+(a-b)^2}}{2},$$

to approximate to the NCP function "$\varphi_{\min}$" throughout this paper. Then, the smoothing approximating function $H_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ of the function H has some good properties, such as the strongly Jacobi consistency [12]

$$\lim_{\varepsilon \to 0} \mathrm{dist}(\nabla \mathrm{H}_\varepsilon(\mathrm{x}), \partial_{\mathrm{Cl}} \mathrm{H}_{\min}(\mathrm{x})) = 0.$$

We next list some others, which are necessary in the convergence analysis.

(i) For arbitrary $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, there is

$$\|H_{\varepsilon_1}(x) - H_{\varepsilon_2}(x)\| \leqslant \kappa |\sqrt{\varepsilon_1} - \sqrt{\varepsilon_2}|,$$

where $\kappa = \sqrt{2n}$.

(ii) For arbitrary $\varepsilon > 0$, we have

$$\|H(x) - H_\varepsilon(x)\| \leqslant \kappa\sqrt{\varepsilon}.$$

(iii) Denote $\alpha(x) = \{i : x_i = F_i(x)\}$. Suppose $x \in \mathbb{R}^n$ is not a solution of NCP. Define a function $\bar{\varepsilon} : \mathbb{R}^n \to \mathbb{R}$ such that

$$\bar{\varepsilon}(x,\delta) = \begin{cases} 1, & \text{if } \frac{n\tau^2(x)}{\delta^2} - \rho(x) \leqslant 0, \\ \frac{\rho(x)\delta}{\sqrt{n\tau^2(x) - \delta^2\rho(x)}}, & \text{otherwise} \end{cases} \quad (4)$$

where $\rho(x) = \min\{(x_i - F_i(x))^2 : i \notin \alpha(x)\}$, $\tau(x) = \frac{1}{2}\max\{\|(x_i - F_i(x))(e_i - \nabla F_i(x))\| : i \notin \alpha(x)\}$ and $\delta > 0$ is a given constant. Then, there is

$$\mathrm{dist}(\nabla \mathrm{H}_\varepsilon(\mathrm{x}), \partial_{\mathrm{Cl}} \mathrm{H}_{\min}(\mathrm{x})) \leqslant \delta,$$

for all ε such that $0 \leqslant \varepsilon \leqslant \bar{\varepsilon}(x,\delta)$.

Semismooth functions include many types of functions in nonsmooth analysis, such as smooth functions, convex functions, maximum functions and so on. The function $H_{\min}$ is semismooth, it has some properties as follows

$$H(y) - H(x) - V^{\mathrm{T}}(y-x) = o(\|y-x\|),\ \forall V \in \partial_{\mathrm{Cl}} H(x),$$

which means that

$$\|H(y) - H(x) - V^{\mathrm{T}}(y-x)\| \leqslant C\|y-x\|^{1+p}, \quad (5)$$

where $C > 0$ and $p \geqslant 0$ are constants. Suppose $V \in \partial_{\mathrm{Cl}} H(x)$ is Lipschitz continuous, there is

$$\|H(y) - H(x)\| \leqslant O(\|y-x\|^{(1+p)}). \quad (6)$$

Especially, we have $p \geqslant 1$ in Eq. (5) when H is strongly semismooth.

ALGORITHM AND ITS GLOBAL CONVERGENCE

In this section, a two-step smoothing Levenberg-Marquardt method for nonlinear complementarity problems is presented, and its global convergence is also shown.

Define $\Phi(x)$ as the merit function of (1) such that

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

and also its approximating merit function $\Phi_\varepsilon(x)$ such that

$$\Phi_\varepsilon(x) = \frac{1}{2}\|H_\varepsilon(x)\|^2.$$

In what follows, we present the two-step smoothing Levenberg-Marquardt method.

Algorithm 1

Step 0: Choose an initial point $x_0 \in \mathbb{R}^n$ and $\eta \in (0,1)$, $\alpha \in (0,1)$, $s \in (0,1)$, $m \in (0,1)$, $\sigma \in (0, 1-\alpha/4)$,

$\gamma > 0$, $\varepsilon > 0$. Set $\kappa = \sqrt{2n}$, $\beta_0 = \|H(x_0)\|$, $\varepsilon_0 = (\alpha\beta_0/2\kappa)^2$, $k = 1$.

Step 1: If $\|V(x_k)^{\mathrm{T}}H(x_k)\| \leqslant \varepsilon$, where $V(x_k) \in \partial_{\mathrm{Cl}}H(x_k)$, then terminates. Otherwise, go to Step 2.

Step 2: Set $\lambda_k = \|H(x_k)\|^{\delta_k}$, where

$$\delta_k = \begin{cases} \frac{1}{\|H(x_k)\|}, & \text{if } \Phi(x_k) \geqslant 1, \\ 1 + \frac{1}{k}, & \text{otherwise.} \end{cases}$$

Obviously, $\delta_k \in (0, 2]$.

Step 3: Solve the following linear system

$$(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)d_k = -J(x_k)^{\mathrm{T}}H_{\varepsilon_k}(x_k), \qquad (7)$$

to obtain its solution d_{k_1}, where $J(x_k)$ is the Jacobi of the function $H_{\varepsilon_k}(x_k)$. Set $y_k = x_k + d_{k_1}$.

Step 4: Solve the linear system

$$(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)d_k = -J(x_k)^{\mathrm{T}}H_{\varepsilon_k}(y_k), \qquad (8)$$

to obtain d_{k_2}. Set $d_k = d_{k_1} + d_{k_2}$.

Step 5: Set $\sigma_k = \min\{\sigma, \lambda_k/4\}$. Select the smallest m_k from $\mathbb{N}$ which satisfies the following inequality

$$\Phi_{\varepsilon_k}(x_k + s^{m_k}d_k) - \Phi_{\varepsilon_k}(x_k) \leqslant -\sigma_k s^{m_k}\|d_k\|^2. \qquad (9)$$

Set $t_k = s^{m_k}$, $x_{k+1} = x_k + t_k d_k$.

Step 6: If

$$\|H(x_{k+1})\| \leqslant \max\{\eta\beta_k, \frac{\|H(x_{k+1}) - H_{\varepsilon_k}(x_{k+1})\|}{\alpha}\},$$

set $\beta_{k+1} = \|H(x_{k+1})\|$ and choose ε_{k+1} such that

$$0 < \varepsilon_{k+1} \leqslant \min\{(\frac{\alpha}{2\kappa}\beta_{k+1})^2, m\varepsilon_k, \bar{\varepsilon}(x_{k+1}, \gamma\beta_{k+1})\},$$

where $\bar{\varepsilon}(.,.)$ is defined by (4). Otherwise, set $\beta_{k+1} = \beta_k$, $\varepsilon_{k+1} = m\varepsilon_k$.

Step 7: Set $k := k + 1$, return to Step 1.

Remark 1 Algorithm 1 is well defined since the line search condition such that

$$\Phi_{\varepsilon_k}(x_k + t_k d_k) - \Phi_{\varepsilon_k}(x_k) \leqslant -\sigma_k t_k \|d_k\|^2,$$

where $\sigma_k = \min\{\sigma, \lambda_k/4\} > 0$ and $t_k > 0$ small enough.

Remark 2 It is important to force the smoothing parameter ε_k to go to zero in Algorithm 1. Since $\varepsilon_{k+1} = m\varepsilon_k$, the sequence ε_k generated is not increasing.

We next investigate the global convergence of Algorithm 1.

Assumption 1 The solution set X of NCP is not empty.

Assumption 2 There exists a neighborhood $\Omega \subset X$ such that both $H_\varepsilon(x)$ and its Jacobi $J(x)$ are Lipschitz continuous on it, i.e., there exists constant $L_1 > 0$, $L_2 > 0$ such that $\|J(y) - J(x)\| \leqslant L_1\|y - x\|$, for all $x, y \in \Omega$, and

$$\|H_\varepsilon(y) - H_\varepsilon(x)\| \leqslant L_2\|y - x\|, \ \forall x, y \in \Omega.$$

Then, by the Lipschitzness of the Jacobi, there is

$$\|H_\varepsilon(y) - H_\varepsilon(x) - J(x)(y - x)\| \leqslant L_1\|y - x\|^2, \quad (10)$$

and

$$\|J(x)\| \leqslant L_2. \qquad (11)$$

In Algorithm 1, if $\|V(x_k)^{\mathrm{T}}H(x_k)\| \leqslant \varepsilon$ occurs at some iteration k, then Algorithm 1 terminates finitely a solution of the problem (1). Otherwise, Algorithm 1 generates a sequence $\{(x_k, \varepsilon_k)\}$, where $\{\varepsilon_k\}$ is not increasing, based on which, we give the following assumption.

Assumption 3 The sequence generated by Algorithm 1 has at least one accumulation point, which is a solution of NCP.

By the virtue of [11], we have Lemma 1 to Lemma 3 as follows.

Lemma 1 *For all $k \geqslant 0$, there is*

$$\|H(x_k) - H_{\varepsilon_k}(x_k)\| \leqslant \alpha\|H(x_k)\|^2.$$

Lemma 2 *For all $x \in \mathbb{R}^n$, $k \in \mathbb{N}$, the following statement holds*

$$\|H_{\varepsilon_{k+1}}(x_{k+1})\| + \kappa\sqrt{\varepsilon_{k+1}} \leqslant \|H_{\varepsilon_k}(x_k)\| + \kappa\sqrt{\varepsilon_k},$$

where $\kappa = \sqrt{2n}$.

Denote an index set K such that

$$K = \{0\} \cup \Big\{k \in \mathbb{N} : \|H(x_k)\| \leqslant \max\Big\{\eta\beta_{k-1}, \frac{\|H(x_k) - H_{\varepsilon_{k-1}}(x_k)\|}{\alpha}\Big\}\Big\}.$$

Lemma 3 *If the index set K is infinite, each accumulation point of $\{x_k\}$ is a solution of NCP.*

Lemma 4 ([7]) *Let $\{x_k\} \subseteq \mathbb{R}^n$ and $\{\varepsilon_k\}$ be two sequences with $x_k \to x^*$ for some $x^* \in \mathbb{R}^n$ and $\varepsilon_k \downarrow 0$. Then, there exists a subsequence $\{\nabla\Phi_{\varepsilon_k}(x_k)\}_L$ such that $\lim_{k\to\infty,\, k\in L} \nabla\Phi_{\varepsilon_k}(x_k) \in \partial_{\mathrm{Cl}}\Phi(x^*)$.*

Theorem 1 *If $\{x_k\}$ is a sequence generated by Algorithm 1 and x^* is an accumulation point of $\{x_k\}$, then $0 \in \partial_{\mathrm{Cl}}\Phi(x^*)$.*

Proof: If K is infinite, the conclusion follows by Lemma 3 immediately. Hence, we just need considering the case when K contains finitely many indices. Suppose K is finite, $\tilde{k}$ is the largest element in K, and x^* is an accumulation point of the sequence such that $x_k \to x^*$, $k \in L$. Without loss of generality, we suppose $K \cap L = \varnothing$. Then, from Step 6 of Algorithm 1, there is

$$\beta_k = \beta_{\tilde{k}} = \|H(x_{\tilde{k}})\|,$$

and

$$\|H(x_k)\| > \eta\beta_k > 0,$$

which implies that

$$\|\Phi(x_k)\| > \eta^2\|\Phi(x_{\tilde{k}})\| > 0, \quad k > \tilde{k}.$$

Next, we prove the statement by contradiction. Assume that x^* is not a stationary point of $\Phi(x)$, which implies that $\Phi(x^*) \neq 0$.

(i): We first show that there exists an index set L such that $\{d_k\}_{k\in L} \subset \{d_k\}$, and it is bounded. From the LM Eq. (7), we have

$$\|d_{k_2}\| = \|-(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1}J(x_k)^{\mathrm{T}}H_{\varepsilon_k}(y_k)\|,$$

together with (10), there is

$$\begin{aligned}\|d_{k_2}\| \leqslant{}& \|(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1}J(x_k)^{\mathrm{T}}H_{\varepsilon_k}(x_k)\| \\ &+ \|(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1}J(x_k)^{\mathrm{T}}d_{k_1}\| \\ &+ L_1\|d_{k_1}\|^2\|(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1}J(x_k)^{\mathrm{T}})\|. \quad (12)\end{aligned}$$

Since d_{k_1} is the solution of the LM Eq. (7), there is

$$d_{k_1} = \|(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1}J(x_k)^{\mathrm{T}}H_{\varepsilon_k}(x_k)\|. \quad (13)$$

Suppose the singular value decomposition (SVD) of $J(x_k)$ is as follows

$$J(x_k) = U_k\Sigma_k S_k^{\mathrm{T}} = U_k \begin{pmatrix} \sigma_{k,1} & & & \\ & \sigma_{k,2} & & \\ & & \ddots & \\ & & & \sigma_{k,n} \end{pmatrix} S_k^{\mathrm{T}}, \quad (14)$$

where $U_k, S_k \in \mathbb{R}^{n\times n}$ are orthogonal matrices and $\sigma_{k,1} \geqslant \cdots \geqslant \sigma_{k,n} \geqslant 0$, there is

$$\begin{aligned}\|(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1}\| &= \|S_k(\Sigma_k^2 + \lambda_k I)^{-1}S_k^{\mathrm{T}}\| \\ &= \|(\Sigma_k^2 + \lambda_k I)^{-1}\| \\ &= \max_{i=1,2,\ldots,n}(\sigma_{k,n}^2 + \lambda_k I)^{-1} \leqslant \lambda_k^{-1}. \quad (15)\end{aligned}$$

Hence, from (12), (13) and (15), we have

$$\begin{aligned}\|d_{k_2}\| &\leqslant \|d_{k_1}\| + \|(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1}\|\|J(x_k)^{\mathrm{T}}\|\|d_{k_1}\| \\ &\quad + L_1\|d_{k_1}\|^2\|(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1}\|\|J(x_k)^{\mathrm{T}}\| \\ &\leqslant \|d_{k_1}\| + \frac{L_2}{\lambda_k}\|d_{k_1}\| + \frac{L_1L_2}{\lambda_k}\|d_{k_1}\|^2. \quad (16)\end{aligned}$$

Furthermore, since

$$\|d_{k_1}\| \leqslant \frac{\|J(x_k)H_{\varepsilon_k}(x_k)\|}{\|\lambda_k\|}, \quad (17)$$

from the LM Eq. (7) and $\{\lambda_k\}$ is bounded for any $k \in L$, there exists an index set denoted as L that $\{d_{k_1}\}_{k\in L}$ is bounded. Then, combining (16) with (17), we have $d_k = d_{k_1} + d_{k_2}$ are bounded. Conveniently, suppose that there exists a convergent subsequence $\{d_k\}_{k\in L}$ such that $d_k \to d^*$.

(ii): We next show that $\liminf_{k\in L} t_k = 0$. If $\liminf_{k\in L} t_k = t^* > 0$, the line search rules in Algorithm 1 shows that

$$\Phi_{\varepsilon_k}(x_{k+1}) - \Phi_{\varepsilon_k}(x_k) \leqslant -t_k\sigma_k\|d_k\|^2,$$

for all $k \in L$. Since

$$d_k = d_{k_1} + d_{k_2},$$

It follows from LM Eqs. (7) and (8) that

$$\begin{aligned}d_{k_1} &= -[J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I]^{-1}J(x_k)H_{\varepsilon_k}(x_k), \\ d_{k_2} &= -[J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I]^{-1}J(x_k)H_{\varepsilon_k}(y_k),\end{aligned}$$

then

$$d_k = -[J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I]^{-1}J(x_k)[H_{\varepsilon_k}(x_k) + H_{\varepsilon_k}(y_k)]. \quad (18)$$

According to (14) such that

$$J(x_k) = U_k\Sigma_k S_k^{\mathrm{T}} = U_k \begin{pmatrix} \sigma_{k,1} & & & \\ & \sigma_{k,2} & & \\ & & \ddots & \\ & & & \sigma_{k,n} \end{pmatrix} S_k^{\mathrm{T}},$$

where $U_k, S_k \in \mathbb{R}^{n\times n}$ are orthogonal matrices and $\sigma_{k,1} \geqslant \cdots \geqslant \sigma_{k,n} \geqslant 0$, we obtain

$$(J(x_k)^{\mathrm{T}}J(x_k) + \lambda_k I)^{-1} = S_k(\Sigma_k^2 + \lambda_k I)^{-1}S_k^{\mathrm{T}},$$

then

$$d_k = -S_k \wedge_k S_k^{\mathrm{T}}J(x_k)[H_{\varepsilon_k}(x_k) + H_{\varepsilon_k}(y_k)]. \quad (19)$$

where

$$\wedge_k = \begin{pmatrix} \frac{1}{\sigma_{k,1}^2 + \lambda_k} & & & \\ & \frac{1}{\sigma_{k,2}^2 + \lambda_k} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_{k,n}^2 + \lambda_k} \end{pmatrix}.$$

Denote $C(x_k) = J(x_k)[H_{\varepsilon_k}(x_k) + H_{\varepsilon_k}(y_k)]$, then

$$\|d_k\|^2 = C(x_k)^{\mathrm{T}}S_k \wedge_k^2 S_k^{\mathrm{T}}C(x_k), \quad (20)$$

Since $x_k \to x^*(k \in L)$, $\varepsilon_k \to 0$ and $\sigma_k = \min\{\sigma, \lambda_k/4\}$, there exists some $m_1 > 0$ such that

$\sigma_{k,i}^2 \leqslant m_1$ for $k \in L, i = 1,\ldots,n$, which implies that some convergent subsequences such that

$$\sigma_{k,i}^2 \to \sigma_i^2,$$

exists. Considering that

$$\lambda_k = \|H(x_k)\|^{\delta_k},$$

where $\delta \in (0,2]$ holds for any $k \in L$, there exists a convergence subsequence of $\{\lambda_k\}_{k\in L}$. Denote it as $\{\lambda_k\}_{k\in L}$ itself and $\lambda_k \to \lambda^*$, $k \in L \to \infty$. Moreover, suppose $\lambda_k \leqslant m_2$ for convenience, there is

$$\|d_k\|^2 \geqslant M\|C(x_k)\|^2, \tag{21}$$

where $M = 1/(m_1+m_2)$. So, we obtain from the line search rule in Step 5 that

$$\begin{aligned}\Phi_{\varepsilon_k}(x_{k+1}) - \Phi_{\varepsilon_k}(x_k) &\leqslant -\sigma_k t_k \|d_k\|^2 \\ &\leqslant -\sigma_k t_k M \|C(x_k)\|^2,\end{aligned}$$

for any $k \in L$ large enough. However, since K is finite, there is

$$\{\sigma_k\} \to \sigma^* = \min\{\sigma, \frac{1}{4}\|H(x^*)\|^{\delta_k}\} > 0.$$

Letting $c = \sigma^* t^* M\|C(x^*)\|^2 > 0$ and combining that $\{\varepsilon_k\} \to 0$, we have

$$\Phi(x_{k+1}) - \Phi(x_k) \leqslant -\frac{c}{2},$$

then, similar to the proof of Theorem 3.10 in [9], deducing from Lemmas 1 and 2, we obtain

$$\Phi(x_{l_{j+1}}) - \Phi(x_{l_j}) \leqslant -\frac{c}{4},$$

for all l_j sufficiently large, which contradicts the nonnegativity of $\Phi(x)$. Hence, we obtain $\liminf_{k\in L} t_k = 0$.

(iii): Assume $\liminf_{k\in L} t_k = 0$, the reminder proof is similar to Theorem 3.4 in [12], since

$$\frac{\Phi_{\varepsilon_k}(x_k + s^{m_k-1}d_k) - \Phi_{\varepsilon_k}(x_k)}{s^{m_k-1}} > -\sigma_k\|d_k\|^2.$$

for all $K \in L$ large enough, we have $\nabla\Phi(x^*)^{\mathrm{T}}d^* \geqslant -\sigma^*\|d^*\|^2$, namely,

$$V^{*\mathrm{T}}d^* \geqslant -\sigma^*\|d^*\|^2,$$

where d^* is the limit of $\{d_k\}_L$ and σ^* is the limit of $\{\sigma_k\}_L$, $V^* \in \partial_{\mathrm{Cl}}\Phi(x^*)$. Suppose $\{S_k\}_L \to S^*$, $\wedge^* = \mathrm{diag}(\frac{1}{\sigma_1^2+\lambda^*}, \cdots, \frac{1}{\sigma_n^2+\lambda^*})$, together with (19), one has

$$(V^*)^{\mathrm{T}}S^*(\sigma^*(\wedge^*)^2 - \wedge^*)(S^*)^{\mathrm{T}}V^* \geqslant 0,$$

which implies $V^* = 0$ and $\Phi(x^*) = 0$. This contradicts $\Phi(x^*) \neq 0$. The proof is completed. □

Remark 3 The Armijo line search plays an important rule in the proof of Theorem 1. Although the two-step and three-step LM methods have been proved with global convergence [18, 19], they are all applied for solving smoothing systems of equations, the global convergence is valuable when the two-step technology is used for solving a nonsmooth system.

LOCAL CONVERGENCE UNDER LOCAL ERROR BOUND

In this section, we investigate the local convergence of the algorithm under the local error bound.

Give the error bound assumption firstly.

Assumption 4 H provides a local error bound on some neighborhood of $\bar{x} \in X$, i.e., there exists constants $c_1 > 0$ such that

$$c_1 \mathrm{dist}(x, X) \leqslant \|H(x)\|, \quad \forall x \in N(\bar{x}, r), \tag{22}$$

where $\mathrm{dist}(x,X) = \inf_{y\in X}\|y - \bar{x}\|$ and $N(\bar{x}, r)$ is a neighborhood of $\bar{x}$.

We recall the SVD of $J(x_k)$, which is the Jacobi of the function $H_\varepsilon(x_k)$ and is necessary for the discussion of the local convergence. More specially, denote

$$U_k = (U_{k,1}, U_{k,2}, U_{k,3}),$$
$$\Sigma_k = \begin{pmatrix}\Sigma_{k,1} & & \\ & \Sigma_{k,2} & \\ & & 0\end{pmatrix}, \quad S_k = \begin{pmatrix}S_{k,1}^{\mathrm{T}} \\ S_{k,2}^{\mathrm{T}} \\ S_{k,3}^{\mathrm{T}}\end{pmatrix},$$

where $\Sigma_{k,1} = \mathrm{diag}(\sigma_{k,1},\ldots,\sigma_{k,r})$ with $\sigma_{k,1} \geqslant \sigma_{k,2} \geqslant \cdots \geqslant \sigma_{k,r} > 0$, and $\Sigma_{k,2} = \mathrm{diag}(\sigma_{k,r+1},\ldots,\sigma_{k,r+q})$ with $\sigma_{k,r+1} \geqslant \sigma_{k,r+2} \geqslant \cdots \geqslant \sigma_{k,r+q} > 0$, when we neglect the subscription k in $\Sigma_{k,i}$ for convenience, $J(x_k)$ denoted as J_k is such that

$$J_k = U_1\Sigma_1 S_1^{\mathrm{T}} + U_2\Sigma_2 S_2^{\mathrm{T}}.$$

Furhermore, denote $H_{\varepsilon_k} = H_{\varepsilon_k}(x_k)$. Then, from (7) and (8) in Algorithm 1, we conclude that

$$\begin{aligned}d_{k_1} &= -S_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^{\mathrm{T}} H_{\varepsilon_k} \\ &\quad - S_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^{\mathrm{T}} H_{\varepsilon_k}, \\ d_{k_2} &= -S_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^{\mathrm{T}} H_{\varepsilon_k}(y_k) \\ &\quad - S_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^{\mathrm{T}} H_{\varepsilon_k}(y_k),\end{aligned}$$

and

$$\begin{aligned}H_{\varepsilon_k} + J_k d_{k_1} &= H_{\varepsilon_k} - U_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^{\mathrm{T}} H_{\varepsilon_k} \\ &\quad - U_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^{\mathrm{T}} H_{\varepsilon_k} \\ &= \lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^{\mathrm{T}} H_{\varepsilon_k} \\ &\quad + \lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^{\mathrm{T}} H_{\varepsilon_k} + U_3 U_3^{\mathrm{T}} H_{\varepsilon_k},\end{aligned}$$

$$H_{\varepsilon_k}(y_k)+J_k d_{k_2} = \lambda_k U_1(\Sigma_1^2+\lambda_k I)^{-1}\Sigma_1 U_1 H_{\varepsilon_k}(y_k) + \lambda_k U_2(\Sigma_2^2+\lambda_k I)^{-1}\Sigma_2 U_2 H_{\varepsilon_k}(y_k)+U_3U_3^{\mathrm{T}}H_{\varepsilon_k}(y_k).$$

Then, similar to the proof of Lemma 3.4 in [15], we conclude Lemma 5 and Lemma 6 immediately.

Lemma 5 *Suppose Assumptions 1, 2, and 4 hold. If $x_k \in N(\bar{x}, r/2)$, then there is*

(a) $\|U_1U_1^{\mathrm{T}}H_{\varepsilon_k}(y_k)\| \leqslant c_2\|\overline{x}_k - x_k\|^2$,

(b) $\|U_2U_2^{\mathrm{T}}H_{\varepsilon_k}(y_k)\| \leqslant c_3\|\overline{x}_k - x_k\|^3$,

(c) $\|U_3U_3^{\mathrm{T}}H_{\varepsilon_k}(y_k)\| \leqslant c_4\|\overline{x}_k - x_k\|^3$,

where c_2, c_3, c_4 are positive constants.

Lemma 6 *Suppose Assumptions 1 and 4 hold. If $\mathrm{dist}(x_k, X) = \|x_k - \overline{x}_k\|$, $\overline{x}_k \in X$ and $x_k, y_k \in N(\overline{x}, r/2)$, there are positive constants c_5, c_6, c_7 such that*

$$\|d_{k_1}\| \leqslant c_5\|\overline{x}_k - x_k\|, \quad \|d_{k_2}\| \leqslant c_6\|\overline{x}_k - x_k\|,$$

and

$$\|H_{\varepsilon_k}(y_k)+J_k d_{k_2}\| \leqslant c_7\|\overline{x}_k - x_k\|^{2+\delta_k}.$$

Proof: Consider d_{k_1} is the minimizer of an optimization problem such that

$$\min_{d\in\mathbb{R}^n} q(d) = \|H_{\varepsilon_k} + J_k d\|^2 + \lambda_k\|d\|^2,$$

where J_k is the Jacobi of H_{ε_k}, it is obvious that

$$\begin{aligned}\|d_{k_1}\|^2 &\leqslant \frac{1}{\lambda_k}q(d_k)\\ &\leqslant \frac{1}{\lambda_k}q(\overline{x}_k - x_k)\\ &= \frac{\|H_{\varepsilon_k} + J_k(\overline{x}_k - x_k)\|^2}{\lambda_k} + \|\overline{x}_k - x_k\|^2,\\ &\leqslant \frac{L_1^2\|\overline{x}_k - x_k\|^4}{\lambda_k} + \|\overline{x}_k - x_k\|^2.\end{aligned}$$

Since

$$\lambda_k = \|H(x_k)\|^{\delta_k} \geqslant c_1^{\delta_k}\|\overline{x}_k - x_k\|^{\delta_k}, \tag{23}$$

we have

$$\|d_{k_1}\|^2 \leqslant L_1^2 c_1^{-\delta_k}\|\overline{x}_k - x_k\|^{4-\delta_k} + \|\overline{x}_k - x_k\|^2,$$

which gives

$$\|d_{k_1}\| \leqslant c_5\|\overline{x}_k - x_k\|,$$

where $c_5 > 0$. Then, combine (16) such that

$$\|d_{k_2}\| \leqslant \|d_{k_1}\| + \frac{L_2}{\lambda_k}\|d_{k_1}\| + \frac{L_1L_2}{\lambda_k}\|d_{k_1}\|^2,$$

with (23), we have $\|d_{k_2}\| \leqslant c_6\|\overline{x}_k - x_k\|$, where $c_6 > 0$.

Next, we show that $\|H_{\varepsilon_k}(y_k)+J_k d_{k_2}\| \leqslant c_7\|\overline{x}_k - x_k\|^2$. In fact, similar the proof of the inequality (3.36) in [19], we have

$$\begin{aligned}&\|H_{\varepsilon_k}(y_k)+J_k d_{k_2}\|\\ &= \lambda_k U_1(\Sigma_1^2+\lambda_k I)^{-1}\Sigma_1 U_1^{\mathrm{T}}H_{\varepsilon_k}(y_k)\\ &\quad + \lambda_k U_2(\Sigma_2^2+\lambda_k I)^{-1}\Sigma_2 U_2^{\mathrm{T}}H_{\varepsilon_k} + U_3U_3^{\mathrm{T}}H_{\varepsilon_k}(y_k)\\ &\leqslant \lambda_k\|\Sigma_1^{-1}\|\|U_1U_1^{\mathrm{T}}H_{\varepsilon_k}(y_k)\| + \lambda_k\|\Sigma_2^{-1}\|\|U_2U_2^{\mathrm{T}}H_{\varepsilon_k}(y_k)\|\\ &\quad + \|U_3U_3^{\mathrm{T}}H_{\varepsilon_k}(y_k)\|\\ &\leqslant O(\|\overline{x}_k - x_k\|)^{2+\delta_k} + O(\|\overline{x}_k - x_k\|)^{3+\delta_k} + O(\|\overline{x}_k - x_k\|)^3\\ &\leqslant O(\|\overline{x}_k - x_k\|)^{2+\delta_k}.\end{aligned} \tag{24}$$

Since $\delta_k \in (0, 2]$, one has

$$\|H_{\varepsilon_k}(y_k)+J_k d_{k_2}\| \leqslant c_7\|\overline{x}_k - x_k\|^{2+\delta_k},$$

where $c_7 > 0$. □

Theorem 2 *Suppose $\{x_k\}$ is generated by Algorithm 1 and Assumptions 1 and 4 hold, then $\{x_k\}$ converges to some solutions quadratically.*

Proof: By the virtue of Assumptions 1 and 4 and Lemma 1, there is

$$\begin{aligned}c_1\|\overline{x}_{k+1} - x_{k+1}\| &\leqslant \|H(x_{k+1})\|\\ &\leqslant \|H_{\varepsilon_k}(x_{k+1})\| + \alpha\|H(x_k)\|^2\\ &\leqslant \|H_{\varepsilon_k}(y_k + d_{k_2})\| + \alpha\|H(x_k)\|^2\\ &\leqslant \|H_{\varepsilon_k}(y_k) + J(y_k)d_{k_2}\| + L_1\|d_{k_2}\|^2 + \alpha\|H(x_k)\|^2\\ &\leqslant \|H_{\varepsilon_k}(y_k) + J_k d_{k_2}\| + \|J(y_k) - J_k\|\|d_{k_2}\| + L_1\|d_{k_2}\|^2\\ &\quad + \alpha\|H(x_k)\|^2,\end{aligned}$$

then, together with (24) and (5) such that

$$\|H(x_k)\|^2 \leqslant O(\|\overline{x}_k - x_k\|^{2(1+p)}),$$

where $p \geqslant 0$, and the Lipschitz continuous of $J(x)$ that

$$\|J(y_k) - J_k\| \leqslant L_1\|\overline{x}_k - x_k\|,$$

we have

$$\begin{aligned}c_1\|\overline{x}_{k+1} - x_{k+1}\| &\leqslant O(\|\overline{x}_k - x_k\|^{2+\delta_k})\\ &\quad + O(\|\overline{x}_k - x_k\|^2) + O(\|\overline{x}_k - x_k\|^{2(1+p)}),\end{aligned}$$

which implies that $\|\overline{x}_{k+1} - x_{k+1}\| \leqslant O(\|\overline{x}_k - x_k\|^2)$. □

Remark 4 In [11], Yu and Pu proposed one-step smoothing LM method for NCP, where the local convergence order is superlinear under the local error bound when H is strong semismooth. However, we obtained that the local convergence of two-step LM method is quadratic when H is semismooth.

NUMERICAL RESULT

In this section, we give some examples to illustrate the effectiveness of the present algorithm and compare it with the smoothing LM methods with one-step iterative direction proposed in [11] (denoted by SLM) and [12] (denoted by MSLM) at different initial points. Here, we code the algorithm in Matlab R2010a, Windows 10(64), 4G Memory, 2.4 GHZ. In what follows, we set $\eta = 0.8$, $\alpha = 0.7$, $\sigma = 0.015$, $s = 0.5$, $\gamma = 10$, $m = 0.75$. The stop criterion is $\|V(x_k)^{\mathrm{T}}H(x_k)\| \leqslant$ 1.0e-6.

Example 1 Let $F(x) = (f_1(x), f_2(x), f_3(x))^{\mathrm{T}}$, where

$$\begin{aligned} f_1(x) &= x_1 - 2, \\ f_2(x) &= x_2 - x_3 + x_2^3 + 3, \\ f_3(x) &= x_2 + x_3 + 2x_3^3 - 3. \end{aligned}$$

The problem has the solution $(2, 0, 1)^{\mathrm{T}}$.

Example 2 (Kojima-Shindo Problem) Let $F(x) = (f_1(x), f_2(x), f_3(x), f_4(x))^{\mathrm{T}}$, where

$$\begin{aligned} f_1(x) &= 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6, \\ f_2(x) &= 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2, \\ f_3(x) &= 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9, \\ f_4(x) &= x_1^2 + 2x_2^2 + 2x_3 + 3x_4 - 3. \end{aligned}$$

There exist solutions $(\sqrt{6}/2, 0, 0, 1/2)^{\mathrm{T}}$ and $(1, 0, 3, 0)^{\mathrm{T}}$ in this problem.

Example 3 Denote $G(x) = (g_1(x), g_2(x), \ldots, g_n(x))^{\mathrm{T}}$, where

$$g_i(x) = -(n+1) + x_i + \sum_{j=1}^{n} x_j, \quad i = 1, \ldots, n-1,$$

$$g_n(x) = -1 + \prod_{j=1}^{n} x_j.$$

Let $F(x) = (f_1(x), f_2(x), \ldots, f_n(x))^{\mathrm{T}}$, where

$$f_i(x) = \begin{cases} g_i(x) - g_i(x^*) + 1, & \text{if } i \text{ is odd}, \\ g_i(x) - g_i(x^*), & \text{otherwise}, \end{cases}$$

$x^* = (0, 1, 0, 1, \ldots)^{\mathrm{T}} \in \mathbb{R}^n$. Obviously, x^* is a unregenerate solution of NCP.

Numerical results are listed in Table 1–Table 3 respectively, where "–" in the tables means that the number of iterations is more than 50. Meanwhile, we take Kojima-Shindo problem as an example, and describe the trend of $\|H(x_k)\|$ with the iterative steps. One random comparison results of MSLM and the present algorithm are shown in Fig. 1, when the initial points are chosen from 10rand(1, 4), 100rand(1, 4), and 1000rand(1, 4).

As seen from Table 1–Table 3, the number of iterations and CPU time of the present algorithm are always less, when compared with SLM and MSLM. However, it can be seen that there are two cases, such as the initial points 10rand(1, 3) in Table 1 and $(100, 100, 100, 100)^{\mathrm{T}}$ in Table 2, where the CPU time of the present algorithm is a little more than MSLM, either the iterative steps is less or the calculation accuracy is improved.

CONCLUSION

In this paper, we proposed a two-step smoothing Levenberge-Marquardt method. Compared with the one-step smoothing LM methods proposed in [11] and [12], the present algorithm is effective, and always need less iterations and CPU time. However, a line search rule is adopted in this paper and there also exists cases that the CPU time of the present algorithm is a little more. Maybe a two-step LM algorithm with trust region technology will be better, which is our future topic to research.

Table 1 Numerical results for Example 1.

Initial point	SLM			MSLM			Algorithm 1		
	Iter	$\|H(x)\|$	CPU (s)	Iter	$\|H(x)\|$	CPU (s)	Iter	$\|H(x)\|$	Cpu (s)
rand(1, 3)	6	2.2712e-09	3.86	4	7.6921e-07	3.88	4	2.4252e-15	3.02
5rand(1, 3)	8	4.8887e-09	5.13	7	1.1556e-09	3.61	5	1.9281e-12	3.53
10rand(1, 3)	14	1.7476e-10	8.53	8	1.0413e-09	4.50	7	1.3958e-07	4.72
100rand(1, 3)	39	4.8428e-07	22.19	14	8.0567e-08	8.26	8	1.2880e-07	5.50

Table 2 Numerical results for Example 2.

Initial point	SLM			MSLM			Algorithm 1		
	Iter	$\|H(x)\|$	CPU (s)	Iter	$\|H(x)\|$	CPU (s)	Iter	$\|H(x)\|$	CPU (s)
$(1, 2, 1, 2)^{\mathrm{T}}$	10	4.4497e-09	8.78	10	1.2116e-09	8.36	6	5.5643e-13	6.17
$(2, 1, 1, 2)^{\mathrm{T}}$	7	1.1143e-09	5.48	6	1.5078e-06	5.19	7	1.6882e-12	4.84
$(10, 10, 10, 10)^{\mathrm{T}}$	14	6.0516e-14	12.44	12	3.4600e-10	10.87	9	2.5746e-17	8.21
$(100, 100, 100, 100)^{\mathrm{T}}$	–	–	–	19	2.0228e-10	15.98	19	8.8180e-11	16.85
$(1000, 1000, 1000, 1000)^{\mathrm{T}}$	23	2.0187e-14	16.95	22	2.9940e-12	18.74	13	2.7798e-08	13.59

Table 3 Numerical results for Example 3.

Dim	Initial point	SLM			MSLM			Algorithm 1		
		Iter	$\|H(x)\|$	CPU (s)	Iter	$\|H(x)\|$	CPU (s)	Iter	$\|H(x)\|$	CPU (s)
4	$(1,0,0,1)^{\mathrm{T}}$	8	4.8183e-22	6.75	6	8.8448e-07	5.45	3	1.4860e-08	3.12
4	$(10,10,10,10)^{\mathrm{T}}$	14	7.4724e-16	10.22	11	1.4533e-08	9.29	7	3.2901e-08	6.80
5	$(1,2,\ldots,5)^{\mathrm{T}}$	13	1.8589e-13	14.76	9	2.4660e-08	10.28	7	1.1512e-17	9.05
5	$(10,10,\ldots,10)^{\mathrm{T}}$	15	3.3329e-13	17.17	11	7.2029e-07	12.31	7	2.5436e-10	8.95
8	$(10,10,\ldots,10)^{\mathrm{T}}$	14	3.7718e-08	32.14	12	4.8087e-16	25.30	8	4.3555e-16	21.44

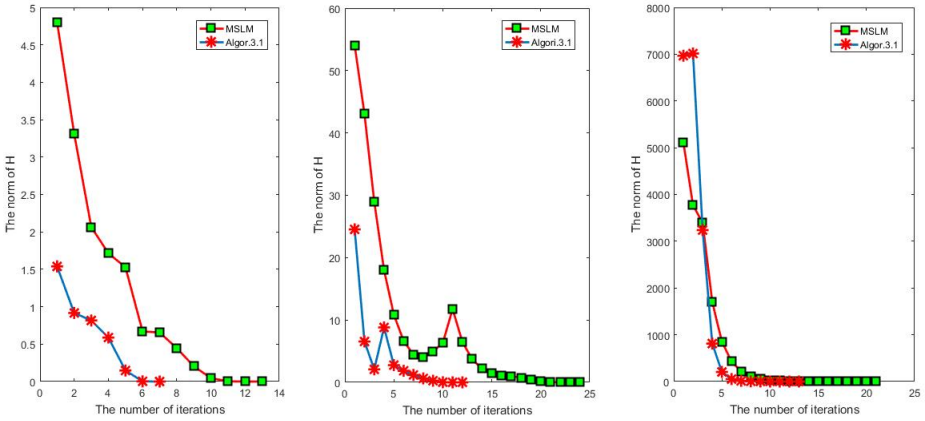

Fig. 1 The comparison of the number of iterations between MSLM and Algorithm 1.

Acknowledgements: The authors are grateful to the anonymous referees for their helpful comments and suggestions. This work was supported by the National Science Foundation of China (nos. 12101198 and 72071130).

REFERENCES

1. Wang HJ, Gao Y (2019) Real-time pricing method for smart grids based on complementarity problem. *J Mod Power Sys* **7**, 1280–1293.
2. Yang YF, Qi L (2005) Smoothing trust region methods for nonlinear complementarity problems with P0-functions. *Ann Oper Res* **133**, 99–117.
3. Qi LQ (1993) Convergence analysis of some algorithms for solving nonsmooth equations. *Math Oper Res* **18**, 227–244.
4. Qi HD, Liao LZ (2000) A smoothing Newton method for general nonlinear complementarity problems. *Comput Optim Appl* **17**, 231–253.
5. Cottle RW, Pang J, Stone R (1992) *The Linear Complementarity Problem*, Academic Press, Boston.
6. Fischer A (1992) A special Newton-type optimization method. *Optim* **24**, 269–284.
7. Song LS, Gao Y (2017) A nonsmooth Levenberg-Marquardt method for vertical complementarity problems. *Num Algor* **76**, 473–485.
8. Facchinei F, Kanzow C (1997) A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems. *Math Prog* **76**, 493–512.
9. Qi LY, Xiao XT, Zhang LW (2016) A parameter-self-adjusting Levenberg-Marquardt method for solving nonsmooth equations. *J Comput Math* **34**, 317–338.
10. Chu AJ, Du SQ, Su YX (2015) A new smoothing conjugate gradient method for solving nonlinear nonsmooth complementarity problems. *Algor* **8**, 1195–1209.
11. Yu HD, Pu DG (2011) Smoothing Levenberg-Marquardt method for general nonlinear complementarity problems under local error bound. *Appl Math Model* **35**, 1337–1348.
12. Song LS, Gao Y (2018) A smoothing Levenberg-Marquardt method for nonlinear complementarity problems. *Num Algor* **79**, 1305–1321.
13. Zheng H, Liu L (2018) A two-step modulus-based matrix splitting iteration method for solving nonlinear complementarity problems of H_+ matrices. *Comput Appl Math* **37**, 5410–5423.
14. Cao Y, Wang A (2019) Two-step modulus-based matrix splitting iteration methods for implicit complementarity problems. *Num Algor* **82**, 1377–1394.
15. Zheng H, Vong S (2021) A two-step modulus-based matrix splitting iteration method for horizontal linear complementarity problems. *Num Algor* **86**, 1791–1810.
16. Yamashita N, Fukushima M (2001) On the rate of convergence of the Levenberg-Marquardt method. *Comput Suppl* **15**, 239–249.
17. Fan JY, Pan J (2006) Convergence properties of a self-adaptive Levenberg-Marquardt algorithm under local error bound condition. *Comput Optim Appl* **34**, 47–62.
18. Fan JY (2012) The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence. *Math Comput* **81**, 447–466.
19. Amini K, Rostami F (2016) Three-steps modified Levenberg-Marquardt method with a new line search for systems of nonlinear equations. *J Comput Appl Math* **300**, 30–42.