On the local convergence of a Levenberg-Marquardt method for nonsmooth nonlinear complementarity problems

Linsen Song\textsuperscript{a,b,*}, Yan Gao\textsuperscript{a}

\textsuperscript{a} School of Management, University of Shanghai for Science and Technology, Shanghai 200093, China
\textsuperscript{b} School of Mathematical Sciences, Henan Institute of Science and Technology, Xinxiang 453003, China

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

Received 14 Jul 2017
Accepted 9 Oct 2017

ABSTRACT: This paper is concerned with a nonlinear complementarity problem, the related functions of which are locally Lipschitzian. As is well known, the nonlinear complementarity problem is reformulated as a system of nonsmooth equations based on complementarity functions, and Levenberg-Marquardt methods are often used to solve it. However, an element of the nonlinear complementarity functions’ B-differential is required in these methods, which is always difficult or time consuming to obtain for a locally Lipschitzian function. By introducing a new subdifferential, rather than the B-differential of the nonlinear complementarity function, a modified Levenberg-Marquardt method is presented, and the local behaviour of this method under the local error bound condition, which is less strong than nonsingular, is shown. Finally, the numerical tests illustrate the effectiveness of the given algorithm.

KEYWORDS: nonsmooth equations, nonsmooth optimization, semismooth function

INTRODUCTION

The general nonlinear complementarity problem

\[ F(x) \geq 0, \ Z(x) \geq 0, \ F(x)^T Z(x) = 0, \]  

where \( F : \mathbb{R}^n \rightarrow \mathbb{R}^n \) and \( Z : \mathbb{R}^n \rightarrow \mathbb{R}^n \), is to find a solution \( x \in \mathbb{R}^n \) which satisfies (1). It plays an important role in the study of the nonlinear programming problems, the variational inequality, equilibrium problems, engineering mechanics and so on. In what follows, we denote \( F(x) = (f_1(x), \ldots, f_n(x)) \), \( Z(x) = (z_1(x), \ldots, z_n(x)) \) for convenience.

There has been extensive research on the nonlinear complementarity problems, especially for the case that \( F(x) \) and \( Z(x) \) in (1) are smooth.\textsuperscript{1–7} Others\textsuperscript{8–13} have considered the case that \( F(x) \) is locally Lipschitzian and \( Z(x) = x \). In Refs.\textsuperscript{8–10} they transformed (1) into an unconstrained optimization and then solved it by nonsmooth optimization methods. In Refs.\textsuperscript{11–13} they reformulated (1) as a nonsmooth equation and solved it by Newton-type methods. As far as we know, the case that the two functions \( F(x) \) and \( Z(x) \) are locally Lipschitzian and may be both nonsmooth has only been studied in Ref.\textsuperscript{14}.

As is well known, the problem (1) is equivalent to the nonsmooth equation

\[ G(x) := (\varphi(f_1(x), z_1(x)), \ldots, \varphi(f_n(x), z_n(x))) = 0, \]  

where \( \varphi : \mathbb{R}^2 \rightarrow \mathbb{R} \) is a nonlinear complementarity function, such as \( \varphi_{\text{min}}(a, b) = \min\{a, b\} \)\textsuperscript{15}, or the Fischer-Burmeister function \( \varphi_{FB}(a, b) = \sqrt{a^2 + b^2} - a - b \)\textsuperscript{16}. Newton-type methods are popular ones to solve it. However, at least one nonsingular element in the subdifferential of \( G \) in the Newton method and an available element in the B-subdifferential of \( G \) in the Levenberg-Marquardt methods are required at each iteration point, which are difficult or time consuming to obtain for locally Lipschitzian functions. With these in mind, by introducing another subdifferential of \( G \), rather than the B-subdifferential, we propose a modified Levenberg-Marquardt method and focus on its local behaviour under the local error bound, which is less strong than nonsingular. The problem we consider here is the case that the two functions \( F(x) \) and \( Z(x) \) are locally Lipschitzian.

PRELIMINARIES

In this section, we recall some basic concepts and propositions in nonsmooth analysis, and give a brief
review of the Levenberg-Marquardt method.

In nonsmooth analysis, there are various generalized differentials with different forms such as the B-differential \(^17\), the Clarke generalized Jacobian \(^18\), and the C-differential \(^19\). Let \(H : \mathbb{R}^n \rightarrow \mathbb{R}^m\) be locally Lipschitzian at \(x \in \mathbb{R}^n\) and let \(\Omega_d\) be a set where \(H\) is differentiable. The B-differential of \(H\) at \(x\) is defined by

\[
\partial_{B}H(x) = \{ \lim JH(x') : x' \rightarrow x, x' \in \Omega_d \},
\]

where \(J\) denotes the Jacobian. The Clarke generalized Jacobian of \(H\) at \(x\) is defined by

\[
\partial_{C}H(x) = \text{conv} \partial_{B}H(x),
\]

where \(\text{conv}\) denotes the convex hull. The C-subdifferential of \(H\) at \(x\) is defined by

\[
\partial_{C}H(x) = \partial_{C}H_{1}(x) \times \cdots \times \partial_{C}H_{m}(x),
\]

where \(H_{i}\) is the \(i\)th component of \(H\).

The function \(H : \mathbb{R}^n \rightarrow \mathbb{R}\) is said to be sub-differentially regular at \(x\) if it is locally Lipschitz continuous at \(x\) and for all \(d \in \mathbb{R}^n\), the classical directional derivative \(H'(x; d)\) exists and

\[
H'(x; d) = H'(x; d),
\]

where \(H'(x; d)\) is the generalized derivative.

Let \(H_{i} : \mathbb{R}^n \rightarrow \mathbb{R}\) be locally Lipschitz continuous and subdifferentially regular at \(x\) for all \(i = 1, \ldots, m\). Then the function

\[
H(x) := \min\{ H_{i}(x) : i = 1, \ldots, m \}
\]

is also subdifferentially regular at \(x\) and

\[
\partial_{C}H(x) = \text{conv} \{ \partial_{C}H_{i} : i \in I(x) \},
\]

where

\[
I(x) := \{ i \in \{1, \ldots, m\} : H_{i}(x) = H(x) \}.
\]

Let \(H : \mathbb{R}^n \rightarrow \mathbb{R}^m\) be locally Lipschitzian at \(x \in \mathbb{R}^n\). It is semismooth if the following limit exists:

\[
\lim_{V \in \partial_{C}H(x+t'd') \atop d' \rightarrow d, t \rightarrow 0^+} V d'.
\]

The class of semismooth functions is very broad and includes many functions, such as smooth functions, convex functions, and maximum functions. In addition, the sums, differences and composites of semismooth functions are also semismooth.

Let \(H : \mathbb{R}^n \rightarrow \mathbb{R}^m\) be locally Lipschitzian and semismooth at \(x \in \mathbb{R}^n\).

(i) There exists \(V \in \partial_{C}H(x)\) such that

\[
H'(x; d) = V d.
\]

(ii) \(H(x + h) - H(x) - H'(x; d) = o(\|h\|)\).

In what follows, we take \(\varphi(a, b) = \min\{a, b\}\) in (2). The Levenberg-Marquardt (LM) method \(^20, 21\) is a classical and popular approach for solving nonlinear equations. It is used to solve the nonsmooth equations as follows \(^22, 23\)

\[
x_{k+1} = x_k - (\xi_k^T \xi_k + \lambda_k I)^{-1} \xi_k^T G(x_k),
\]

where \(\xi_k := \xi_k(x_k) \in \partial_{B}G(x_k), \lambda_k > 0\) is called the LM parameter. Let

\[
\Phi(x) = \frac{1}{2} \|G(x)\|^2,
\]

\(\Phi(x)\) is semismooth when \(F(x)\) and \(Z(x)\) are semismooth.

**ALGORITHM AND CONVERGENCE ANALYSIS**

In this section, we present a modified LM method and show its local convergence result under local error bound.

As mentioned in Refs. \(^22, 23\), the LM method can be used to solve (2) directly. Since it is necessary to calculate an element in the B-differential of \(G\) or \(\varphi\) at each iteration, which is difficult or time consuming in general, inspired by a new subdifferential proposed by Gao \(^24\), we set a set-valued mapping \(V \rightarrow V(x)\) from \(\mathbb{R}^n \rightarrow 2^{m \times n}\) as follows:

\[
V(x) = V_1(x) \times \cdots \times V_n(x),
\]

where

\[
V_i(x) = \begin{cases} \partial_{B}f_i(x), & i \in \alpha(x) \cup \beta(x), \\ \partial_{B}z_i(x), & i \in \gamma(x) \end{cases},
\]

and \(\alpha(x) = \{ i : f_i(x) < z_i(x) \}\), \(\beta(x) = \{ i : f_i(x) = z_i(x) \}\), \(\gamma(x) = \{ i : f_i(x) > z_i(x) \}\).

Evidently, \(V(x)\) defined above can be calculated by determining the index sets \(\alpha(x), \beta(x), \gamma(x)\), and evaluating one element in the B-differential of \(f_i(x)\) or \(z_i(x)\). In particular, if \(f_i(x)\) or \(z_i(x)\) is smooth, we can take \(V_i(x)\) in (4) correspondingly as follows:

\[
V_i(x) = \begin{cases} \nabla f_i(x), & i \in \alpha(x) \cup \beta(x), \\ \nabla z_i(x), & i \in \gamma(x) \end{cases},
\]

or

\[
V_i(x) = \begin{cases} \partial_{B}f_i(x), & i \in \alpha(x) \cup \beta(x), \\ \partial_{B}z_i(x), & i \in \gamma(x) \end{cases}.
\]
However, it should be mentioned that $V$ may be not the B-differential of $G$ and not even upper-semismooth as a set-valued mapping. In what follows, we take the differential $V$ as a tool instead of the B-differential of $G$.

Let the LM parameter $\lambda_k = \mu_k\|G(x_k)\|^2$, where $\mu_k$ is adjusted by the trust region technology. We now present the LM method for solving (2) as follows:

$$x_{k+1} = x_k - (\xi_k^T \xi_k + \lambda_k I)^{-1} \xi_k^T G(x_k),$$

where $\xi_k \in V(x_k)$.

**Algorithm 1**

**Step 1**: Give an initial point $x_0 \in \mathbb{R}^n$ and parameters $\epsilon > 0, \mu_0 > 0, 0 < m < 1, 0 < M < \infty, k := 0$.

**Step 2**: If $\|G(x_k)\| \leq \epsilon$, stop.

**Step 3**: Set $\lambda_k = \mu_k\|G(x_k)\|^2$. Solve the following nonlinear equations to obtain $d_k$,

$$(\xi_k^T \xi_k + \lambda_k I) d_k = -\xi_k^T G(x_k),$$

where $\xi_k \in V(x_k)$.

**Step 4**: Compute $r_k = A/P$, where

$$A := \Phi(x_k + d_k) - \Phi(x_k),$$

$$P := G(x_k)^T \xi_k d_k + \frac{1}{2} d_k^T \xi_k^T \xi_k d_k.$$ 

**Step 5**: Update the parameter $\mu_k$ using

$$\mu_{k+1} = \begin{cases} 
\min \{4\mu_k, M\}, & r_k < 0.25, \\
\mu_k, & r_k \in [0.25, 0.75], \\
\max \{0.25\mu_k, m\}, & r_k > 0.75. 
\end{cases}$$

(6)

**Step 6**: Set $x_{k+1} = x_k + d_k, k := k + 1$. Go to Step 2.

**Remark 1** The trust region strategy for the parameter $\mu_k$ is useful in Algorithm 1 as it can improve the numerical results. On the other hand, the lower bound $m$ and the upper bound $M$ is necessary in the next convergence analysis.

**Remark 2** Actually, we can replace $\lambda_k = \mu_k\|G(x_k)\|^2$ by $\lambda_k = \mu_k\|G(x_k)\|^2$ in Algorithm 1 to improve the numerical results, where

$$\delta_k = \begin{cases} 
\frac{1}{\|G(x_k)\|}, & \|G(x_k)\| \geq 1, \\
1 + \frac{1}{k}, & \text{otherwise.}
\end{cases}$$

In what follows, we investigate the local convergence of the given algorithm under the local error bound condition. We define the following assumptions.

(A1) The solution set $X$ of problem 2 is nonempty.

(A2) The functions $f_i(x)$ and $z_i(x)$ are semismooth and regular.

(A3) $\|G\|$ provides a local error bound on some neighbourhood of $x^* \in X$. That is, there exist constants $r > 0, c > 0$ such that

$$\|G(x)\| \geq c \cdot \text{dist}(x, X), \quad \forall x \in N(x^*, r).$$

**Lemma 1** Let $A_3$ hold. Then, there exist constants $c_1 > 0, p > 0$ such that

$$G(x) - G(x^*) - \xi(x - x^*) \leq c_1\|x - x^*\|^{1+p},$$

for any $x, x' \in \mathbb{R}^n$, where $\xi \in V(x)$.

**Proof**: Denote $g_i(x) = \min\{f_i(x), z_i(x)\}$. We start by proving $V(x) \subseteq \partial_{G} g(x)$. Actually, since $A_2$ holds and

$$\partial_{G} g(x) = \mathcal{V}(x) \cup \ldots \cup \mathcal{V}(x),$$

there is $V(x) \subseteq \mathcal{V}(x) \subseteq \partial_{G} g(x)$. Hence $V(x) \subseteq \partial_{G} g(x)$ holds clearly.

For $i = 1, \ldots, n$, there exists a constant $p_i > 0$ such that

$$g_i(x) - g_i(x') - \xi_i(x - x') = O(\|x - x'\|^{1+p_i}),$$

where $\xi_i = \mathcal{V}(x) \subseteq \partial_{G} g_i(x)$. Hence one has that

$$G(x) - G(x') - \xi(x - x') = O(\|x - x'\|^{1+p}),$$

where $\xi = (\xi_1, \ldots, \xi_n) \in V(x), p = \min\{p_1, \ldots, p_n\}$, which implies the statement.

**Lemma 2** Let $A_1$, $A_2$, and $A_3$ hold. If $\{x_k\}$ is generated by Algorithm 1 then $\text{dist}(x_k, X) = \|x_k - \bar{x}_k\|$, where $\bar{x}_k \in X$. Then, we have

$$\|d_k\| \leq O(\|x_k - \bar{x}_k\|^p),$$

where $\alpha = \min\{p, 1\}$.

**Proof**: Let $d_k$ be the optimal solution of the problem as follows:

$$\min q(d) = \|G(x_k) + \xi_k d\|^2 + \lambda_k \|d\|^2, \quad \forall \xi_k \in V(x_k).$$

(9)
We obtain
\[ ||d_k||^2 \leq \frac{1}{\lambda_k} q(d_k) \leq \frac{1}{\lambda_k} q(\bar{x}_k - x_k) \]
\[ = \frac{\|G(x_k) + \xi_k \bar{x}_k - x_k\|^2}{\lambda_k} + \|\bar{x}_k - x_k\|^2. \]

Then, from (6) and (8),
\[ ||d_k||^2 \leq \frac{c_1^{2(1+p)}}{\mu_k} \|\bar{x}_k - x_k\|^2 + \|\bar{x}_k - x_k\|^2 \]
\[ \leq m^{-1} c^{-1} c_1^{2(1+p)} \|\bar{x}_k - x_k\|^2 (1 + p) \]
\[ + \|\bar{x}_k - x_k\|^2, \]
which implies the statement.

**Theorem 1** Let A₁, A₂, and A₃ hold. If \( x^* \) is a solution of the problem (2), the iteration (5) generates the sequence \{\( x_k \)\} converging to \( x^* \) with the convergence order \( q \), where \( q = p^* (1 + p) \) when \( p < 1 \), and \( q = 2 \) when \( p \geq 1 \). In particular, if \( f_i(x) \) and \( z_i(x) \) are strongly semismooth, the sequence \{\( x_k \)\} converges to \( x^* \) quadratically.

**Proof:** Let \( \text{dist}(x_k, X) = \|x_k - \bar{x}_k\| \). Since \( G \) is locally Lipschitzian, there exists a constant \( L > 0 \) such that
\[ \|G(x_k)\| \leq L \|x_k - \bar{x}_k\|. \]
(10)

According to A₁ and Lemma 1,
\[ c \|x_{k+1} - \bar{x}_{k+1}\| \]
\[ \leq L \|x_k - \bar{x}_k\| + O(||d_k||^2) \]
\[ \leq \|G(x_k) + \xi_k d_k\| + O(||d_k||^{1+p}) \]
\[ \leq \|G(x_k) + \xi_k d_k\| + L_1 ||d_k||^{1+p}, \]
where \( L_1 > 0 \) is a constant, \( \xi_k \in V(x) \). From (8) and (9),
\[ \|G(x_k) + \xi_k d_k\|^2 \]
\[ \leq \|G(x_k) + \xi_k (\bar{x}_k - x_k)\|^2 + \lambda_k \|\bar{x}_k - x_k\|^2 \]
\[ = \|G(x_k) - G(\bar{x}_k) - \xi_k (x_k - \bar{x}_k)\|^2 \]
\[ + \lambda_k \|\bar{x}_k - x_k\|^2 \]
\[ \leq c_1^2 \|x_k - \bar{x}_k\|^2 (1 + p) + ML^2 \|x_k - \bar{x}_k\|^4, \]
which means that, for \( k \) large enough,
\[ \|G(x_k) + \xi_k d_k\| \leq (c_1 + \sqrt{ML}) \|x_k - \bar{x}_k\|^p, \]
where \( \beta = \min\{1, p\} \). Hence, for \( k \) large enough,
\[ \|x_{k+1} - \bar{x}_{k+1}\| \leq O(||x_k - \bar{x}_k||^{\beta}), \]
(11)

**Table 1** Numerical results for Example 1.

<table>
<thead>
<tr>
<th>Initial point</th>
<th>( G(x_1) )</th>
<th>( k )</th>
<th>CPU (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.6927 \times 10^{-13}</td>
<td>3</td>
<td>0.95</td>
</tr>
<tr>
<td>2</td>
<td>7.0541 \times 10^{-9}</td>
<td>3</td>
<td>1.03</td>
</tr>
<tr>
<td>3</td>
<td>6.2892 \times 10^{-20}</td>
<td>4</td>
<td>1.02</td>
</tr>
<tr>
<td>4</td>
<td>4.6613 \times 10^{-12}</td>
<td>4</td>
<td>1.05</td>
</tr>
<tr>
<td>5</td>
<td>5.5108 \times 10^{-18}</td>
<td>5</td>
<td>1.03</td>
</tr>
</tbody>
</table>

where \( \gamma = \min\{\beta, \alpha (1 + p)\} \), i.e.,
\[ \gamma = \begin{cases} p (1 + p), & p < 1, \\ 2, & p \geq 1. \end{cases} \]

On the other hand, it is clear that
\[ \|\bar{x}_k - x_k\| \leq ||\bar{x}_{k+1} - x_{k+1}|| + ||d_k||. \]
(12)
The above inequality and Lemma 2 imply that, for \( k \) large enough,
\[ \|\bar{x}_k - x_k\| \leq 2 ||d_k||. \]
Hence, from (11), (12), and Lemma 2,
\[ ||d_{k+1}|| \leq O(||d_k||^q), \]
where
\[ q = \alpha \gamma = \begin{cases} p^2 (1 + p), & p < 1, \\ 2, & p \geq 1, \end{cases} \]
which implies the statement.

In particular, if \( f_i(x) \) and \( z_i(x) \) are strongly semismooth, which means that \( p \geq 1 \), the sequence \{\( x_k \)\} converges to \( x^* \) quadratically.

**NUMERICAL TESTS**

In this section, we present some numerical results to show the performance of the LM method. We coded the algorithm in MATLAB R2010A, and set the parameters \( \mu_0 = 1 \), \( m = 1.0 \times 10^{-6} \), \( M = 1.0 \times 10^{12} \). The stop criteria is \( ||G(x)|| \leq 1.0 \times 10^{-6} \).

**Example 1** Consider the nonsmooth nonlinear complementarity problem (1), where
\[ F(x) = \max\{x - 2, 2x - 5\}, \quad Z(x) = \max\{x, \frac{1}{2}x - \frac{3}{4}\}. \]
There are two exact solutions, namely, 0 and 2.

**Example 2** Consider the nonsmooth nonlinear complementarity problem (1), where
\[ F(x) = \begin{cases} |2x_1 - 1| \frac{2x_1 - 1}{|4x_2 + x_1 - 1|}, \\ \max\{x_1, x_1 - 6\} \frac{\max\{x_1, x_1 - 6\}}{\max\{x_2, x_2 - \frac{1}{2}x_1^2\}}. \end{cases} \]
There are three exact solutions: \( \left( \frac{1}{2}, 0 \right), \left( 0, \frac{1}{8} \right), \left( 0, 0 \right). \)
Example 3 Consider the nonsmooth nonlinear complementarity problem (1), where

\[ F(x) = \begin{pmatrix} |2x_1 - x_2 + 3x_3 + 2x_4 - 6| \\ 3x_1 - 3x_2 + 3x_3 + 2x_4 - 5 \\ 3x_1 - x_2 + 3x_3 + 2x_4 - 3 \\ 3x_1 - x_2 + 3x_3 - 4 \end{pmatrix}, \]

\[ Z(x) = (x_1, x_2, x_3, x_4). \]

The exact solutions of this problem are \((1, 1, 1, 1), \left(\frac{1}{2}, 0, 0, \frac{1}{2}\right), (0, 0, \frac{11}{5}, \frac{13}{5}), (3, 0, 0, 0).\)

Example 4 Consider the nonsmooth nonlinear complementarity problem (1), where

\[ F(x) = (f_1(x), f_2(x), f_3(x), f_4(x)), \]

\[ Z(x) = (x_1, x_2, x_3, x_4) \text{ with } f_i(x) = \sum_{j=1}^{3} \max(-x_j - x_{j+1} - x_j - x_{j+1} + x_j^2 + x_{j+1}^2 - 1), \]

\[ i = 1, 2, 3, 4. \]

The exact solution of this problem is \((0, 0, 0, 0).\)

Example 5 Consider the nonsmooth nonlinear complementarity problem (1), where

\[ F(x) = (f_1(x), \ldots, f_m(x)), Z(x) = (x_1, \ldots, x_m) \]

with \(f_i(x) = \max\{x_1^2, \ldots, x_m^2\}, i = 1, \ldots, m.\) The exact solution of this problem is \((0, \ldots, 0).\)

**CONCLUSIONS**

In this paper, a Levenberg-Marquardt method is presented and its local behaviour is shown for general nonlinear nonsmooth complementarity problem. In each iteration of this method, an element of the B-differential of the nonsmooth function \(f_i(x)\) or \(z_i(x)\), rather than the B-differential of \(G\) is needed, which is always difficult or time consuming to obtain in most cases. Since the nonlinear complementarity problems here includes many types, such that the smooth cases and the convex cases, the corresponding local convergence rate is obtained based on the present subdifferential.

**Acknowledgements:** This work was supported by National Science Foundation of China (No. 11171221), Research Fund for the Doctoral Programme of Higher Education of China (No. 20123120110004), Natural Science Foundation of Shanghai (No. 14ZR1429200), Innovation Programme of Shanghai Municipal Education Commission (No. 15ZZ073).

**REFERENCES**