Research Article Generalized Newton Method for a Kind of Complementarity Problem

Shou-qiang Du

College of Mathematics, Qingdao University, Qingdao 266071, China

Correspondence should be addressed to Shou-qiang Du; dsq8933@sina.com

Received 20 February 2014; Accepted 6 April 2014; Published 27 April 2014

Academic Editor: Yisheng Song

Copyright © 2014 Shou-qiang Du. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A generalized Newton method for the solution of a kind of complementarity problem is given. The method is based on a nonsmooth equations reformulation of the problem by *F-B* function and on a generalized Newton method. The merit function used is a differentiable function. The global convergence and superlinear local convergence results are also given under suitable assumptions. Finally, some numerical results and discussions are presented.

1. Introduction

This paper considers a kind of complementarity problem:

$$F(x) \ge 0, \quad G(x) \ge 0, \quad \langle F(x), G(x) \rangle = 0, \quad (1)$$

where $F : \mathbb{R}^n \to \mathbb{R}^n, G : \mathbb{R}^n \to \mathbb{R}^n$ are two differentiable functions. Equation (1) is also denoted by GCP(F, G) in [1]. When G(x) = x, (1) reduces to the nonlinear complementarity problem, which is a general framework for optimality conditions of nonlinear optimization problems as well as some problems arising in different fields. In the past few years, several theoretical and computational results for complementarity have been established, such as [1–12]. The method is based on semismooth equations reformulation of (1). As we know, Newton-type methods are one of the fastest methods for the solution of equations and nonsmooth equations. But the methods are usually only locally convergent. In recent years, some research work has been devoted to techniques for globalizing the local methods. So, in this paper, we are interested in solving (1) by a Newton based method. The method is based on a line search and a semismooth equations reformulation of (1).

This paper is organized as follows. In the next section, (1) is converted into a semismooth equation. Then, a generalized Newton method for solving the reformulation is introduced. The global convergence and the local superlinear convergence

of the method are also presented. In the Section 3, numerical experimental results and some discussions are listed.

2. Preliminaries and Method

Firstly, we introduce some notations and propositions used in the paper (see for instance [1–5]). Given $H : \mathbb{R}^n \to \mathbb{R}^n$, we use H'(x) to denote its Jacobian at x. If H is locally Lipschitzian and D_H denote the set of differentiable points, the *B*-subdifferential and Clarke general Jacobian of H at xare defined as

 $\partial_B H(x)$

$$= \left\{ V \in \mathbb{R}^{m \times n} \mid \exists \left\{ x_k \right\} \subset D_H : \left\{ x_k \right\} \longrightarrow x, H'(x_k) \longrightarrow V \right\}; \\ \partial H(x)$$

$$= \operatorname{conv} \left\{ V \in \mathbb{R}^{m \times n} \mid \exists \left\{ x_k \right\} \subset D_H : \left\{ x_k \right\} \longrightarrow x, \\ H'(x_k) \longrightarrow V \right\}.$$
(2)

Let $H : \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz function. If the following limit,

$$\lim_{\substack{V \in \partial H(x+th') \\ h' \to h, t \to 0^+}} Vh',$$
(3)

exists for any $h \in \mathbb{R}^n$, H is said to be semismooth at x.

Proposition 1. If H(x) is semismooth at x, the following equations are equivalent.

(a) When
$$h \to 0$$
,

$$Vh - H'(x;h) = o(||h||).$$
 (4)

(b) Where
$$V \in \partial H(x+h)$$
,

$$\lim_{x+h\in D_{H},h\to 0}\frac{H'(x+h;h)-H'(x;h)}{\|h\|}=0.$$
 (5)

Definition 2. $M \in \mathbb{R}^{n \times n}$ is called a *P*-matrix if its principal minors are all positive and a P_0 -matrix if its principal minors are all nonnegative.

Definition 3. H is called SC^1 function if H is continuously differentiable and its gradient is semismooth.

Definition 4. H is BD-regular at x if $\forall V \in \partial_B H(x)$ are nonsingular.

In the following, we give the reformulation of (1) and the generalized Newton method. Based on the *F*-*B* function

$$\phi(a,b) = \sqrt{a^2 + b^2} - (a+b), \qquad (6)$$

we consider the following reformulation of (1):

$$\varphi(x) = \begin{pmatrix} \phi(F_1, G_1) \\ \vdots \\ \phi(F_n, G_n) \end{pmatrix} = 0.$$
(7)

x solves (1) \Leftrightarrow *x* solves (7). Denoting

$$\Omega = \left\{ x \in \mathbb{R}^{n} \mid (F_{i}(x))^{2} + (G_{i}(x))^{2} > 0, i = 1, \dots, n \right\},$$
(8)

we know that φ is locally Lipschitz on \mathbb{R}^n and differentiable on Ω . When $x \in \Omega$, we have

$$\nabla \varphi (x) = \operatorname{diag} (\gamma_i (x)) \nabla F (x) + \operatorname{diag} (\mu_i (x)) \nabla G (x), \quad (9)$$

where $\gamma_i(x) = F_i(x)(F_i^2(x) + G_i^2(x))^{1/2} - 1, \mu_i(x) = G_i(x)(F_i^2(x) + G_i^2(x))^{1/2} - 1, i = 1, ..., n$. We also know that

$$(\gamma_i(x) + 1)^2 + (\mu_i(x) + 1)^2 = 1, \quad i = 1, ..., n;$$

$$\mu_i(x) = -1, \quad \gamma_i(x) = 0 \iff F_i(x) > 0, \quad G_i(x) = 0,$$
$$i = 1, \dots, n; (10)$$

$$\mu_i(x) = 0, \quad \gamma_i(x) = -1 \iff F_i(x) = 0, \quad G_i(x) > 0,$$

 $i = 1, \dots, n.$

When *x* is not in Ω , for $V \in \partial \varphi(x)$, we have

$$V = \operatorname{diag}\left(\gamma_{i}\left(x\right)\right) \nabla F\left(x\right) + \operatorname{diag}\left(\mu_{i}\left(x\right)\right) \nabla G\left(x\right), \quad (11)$$

where $(\gamma_i(x) + 1)^2 + (\mu_i(x) + 1)^2 \le 1, i = 1, ..., n$. Then, for $\forall x$, we have

$$(\gamma_i(x)+1)^2 + (\mu_i(x)+1)^2 \le 1, \quad i=1,\ldots,n.$$
 (12)

Defining a merit function as

$$\Psi(x) = \frac{1}{2} \|\varphi(x)\|^2,$$
(13)

then solving (1) is equivalent to solving the following unconstrained minimization problem:

$$\min_{x \in \mathbb{R}^n} \Psi(x), \tag{14}$$

where Ψ is continuously differentiable. We know that $\nabla \Psi(x) = V^T \varphi(x)$, where $V \in \partial \varphi(x)$. Now, we give the following generalized Newton method for solving (1).

Method 1 (generalized Newton method). Given $x_0 \in \mathbb{R}^n$ and $\rho > 0, p > 2, \beta \in (0, 1/2), \epsilon \ge 0$.

Step 1. If $\|\nabla \Psi(x_k)\| \leq \epsilon$, stop.

Step 2. Choose $V_k \in \partial \varphi(x_k)$ and compute d_k of the equation

$$V_k d + \varphi\left(x_k\right) = 0. \tag{15}$$

If the condition

$$\nabla \Psi(x_k)^T d_k + \rho \|d_k\|^p \le 0 \tag{16}$$

is not satisfied, set $d_k = -(V_k)^T \varphi(x_k)$.

Step 3. Find the smallest integer i^k such that

$$\Psi(x_{k}) + (1 - \beta) 2^{-i^{k}} \nabla \Psi(x_{k})^{T} d_{k}$$

$$\leq \Psi\left(x_{k} + 2^{-i^{k}} d_{k}\right) \qquad (17)$$

$$\leq \Psi(x_{k}) + \beta 2^{-i^{k}} \nabla \Psi(x_{k})^{T} d_{k}.$$

Set $x_{k+1} = x_k + 2^{-i^k} d_k$, let k := k + 1, and go to Step 1. The direction obtained by solving (15) is a descent

The direction obtained by solving (15) is a descent direction for the function Ψ . This is a standard property of Newton method for the solution of a smooth equation. But it is no longer true for nonsmooth equations. In the above method, the direction obtained by solving (15) is a descent direction by the fact that $\nabla \Psi(x_k)^T d_k = -\|\varphi(x_k)\|^2$. On the other hand, as we know, d_k is not unique in (15) if V_k is not nonsingular. So, we will give the following proposition to ensure V_k is nonsingular.

Proposition 5. Suppose that F, G given by (1) satisfy $d_i \nabla F_i(x) d \ge \overline{\gamma} ||d||^2, d_i \nabla G_i(x) d \ge \overline{\gamma} ||d||^2$, for any $x \in \mathbb{R}^n, d \in \mathbb{R}^n, \overline{\gamma} > 0$ and i = 1, ..., n; then $\forall V \in \partial \varphi(x)$ is nonsingular.

Proof. Let $\forall V \in \partial \varphi(x)$ and let *d* solve Vd = 0. By the above analysis of *V*, for constants $\gamma_i, \mu_i, i = 1, ..., n$, we have

$$\operatorname{diag}(\gamma_{i}(x)) \nabla F(x) d + \operatorname{diag}(\mu_{i}(x)) \nabla G(x) d = 0,$$

$$i = 1, \dots, n,$$
(18)

and so we have

$$\mu_1 \nabla G_1(x) d + \gamma_1 \nabla F_1(x) d = 0,$$

$$\vdots \qquad (19)$$

$$\mu_n \nabla G_n(x) d + \gamma_n \nabla F_n(x) d = 0.$$

Multiplying the *i*th equation in the above equations by d_i , we get

$$d_{1}\mu_{1}\nabla G_{1}(x) d + d_{1}\gamma_{1}\nabla F_{1}(x) d = 0,$$

$$\vdots$$
$$d_{n}\mu_{n}\nabla G_{n}(x) d + d_{n}\gamma_{n}\nabla F_{n}(x) d = 0.$$
(20)

If
$$\gamma_i = 0, i = 1, ..., n$$
, by (12), we have $1 + (\mu_i + 1)^2 \le 1$. So,

$$0 \le (\mu_i + 1)^2 \le 0,$$
 (21)

and then

$$\mu_i = -1, \quad i = 1, \dots, n.$$
 (22)

We have

0

$$-d_i \nabla G_i(x) d = 0, \quad i = 1, \dots, n,$$

$$= -d_i \nabla G_i(x) d \le -\overline{\gamma} \|d\|^2 \le 0, \quad i = 1, \dots, n.$$
(23)

Then d = 0. On the other hand, if $\gamma_i \neq 0$, $i = 1, \dots, n$, we have

$$d_i \nabla F_i(x) d = -\frac{\mu_i}{\gamma_i} \left(d_i \nabla G_i(x) d \right) \le -\frac{\mu_i}{\gamma_i} \overline{\gamma} ||d||^2 \le 0,$$

$$i = 1, \dots, n.$$
(24)

By $d_i \nabla F_i(x) d \ge \overline{\gamma} ||d||^2 \ge 0$, we get

$$0 \le \overline{\gamma} \|d\|^2 \le 0. \tag{25}$$

Then d = 0 and the nonsingularity of *V* follows. This completes the proof.

 Ψ may have a global minimum x^* with $\Psi(x^*) > 0$ when (1) has no solution. Another point which is worth concern is that the conditions of a stationary point of Ψ are a solution of (1). Similar to [1], we give the conditions, which guarantee that every stationary point of Ψ is a solution of (1).

Proposition 6. If (1) has a nonempty solution set, then x^* is a solution of (1) $\Leftrightarrow \Psi(x^*) = 0$.

Proposition 7. If x^* is a stationary point of Ψ such that $G'(x^*)$ is nonsingular and $F'(x^*)G'(x^*)^{-1}$ is a P_0 matrix, then x^* is a solution of (1).

Now, we give the global convergence theorem for Method

Theorem 8. Suppose that the sequence $\{x_k\}$ is generated by Method 1. Then each accumulation point of the sequence $\{x_k\}$ is a stationary point of Ψ .

Proof. The following proof is given by two parts.

Part I. If, for an infinite set of indices *K*, we have $d_k = -(V_k)^T \varphi(x_k)$, then, by Proposition 1.16 in [7], we know that any limit point of $\{x_k\}$ is a stationary point of Ψ .

Part II. We assume that the direction is always computed by (15). Suppose that $\{x_k\} \rightarrow x^*$ and $\nabla \Psi(x^*) \neq 0$. By (15), we have

$$\|\varphi(x_k)\| = \|V_k d_k\| \le \|V_k\| \|d_k\|.$$
 (26)

From (26), we get

1.

$$\|d_k\| \ge \frac{\|\varphi(x_k)\|}{\|V_k\|}.$$
 (27)

For m > 0 and M > 0, we know that

$$0 < m \le \left\| d_k \right\| \le M. \tag{28}$$

If $\{||d_k||\}_{K_1} \to 0$ for k in K_1 . For V_k which is bounded and (27), we get $\Psi(x^*) = 0$, which is contradicting the assumption. On the other hand, taking into account that $\nabla \Psi(x_k)$ is bounded and (16), $||d_k||$ cannot be unbounded. Because of (17) and the fact that Ψ is a continuously differentiable function, we have

$$\left\{2^{-i^{k}}\nabla\Psi(x_{k})^{T}d_{k}\right\}\longrightarrow0,\quad(k\longrightarrow\infty)\,.$$
(29)

In the following, we will prove that 2^{-i^k} is bounded away from 0. Supposing the contrary,

$$\frac{\Psi\left(x_{k}+2^{-(i^{k}-1)}d_{k}\right)-\Psi\left(x_{k}\right)}{2^{-(i^{k}-1)}}>\beta\nabla\Psi\left(x_{k}\right)^{T}d_{k}.$$
(30)

By (28), we can assume that $\{d_k\} \to \overline{d}, \overline{d} \neq 0$. Taking limit of both sides of (30), we have

$$\nabla \Psi(x^*)^T \overline{d} \ge \beta \nabla \Psi(x^*)^T \overline{d}.$$
 (31)

From (16), we know $\nabla \Psi(x^*)^T \overline{d} < 0$, which contradicts (31). So, we know that 2^{-i^k} is bounded away from 0. On the other hand, (16) and (29) imply that $\{d_k\} \to 0$ and this contradicts (28). This completes the proof.

In the following part of this section, we will discuss the local superlinear convergence of Method 1.

Remark 9. Suppose that the sequence $\{x_k\}$ is generated by Method 1. If one of the limit points of the sequence $\{x_k\}$ is x^* , which is a solution of (1), and x^* is a BD-regular solution

of $\varphi(x) = 0$, then, we can prove locally that the direction is always the solution of (15) and eventually the step size of one satisfies (17). So the method eventually reduces to the local method $x_{k+1} = x_k + (V_k)^{-1}\varphi(x_k), V_k \in \partial_B\varphi(x_k)$.

Theorem 10. Let the BD-regular condition hold at an accumulation point x^* of the sequence $\{x_k\}$ generated by Method 1. Then $\{x_k\}$ converges to x^* Q-superlinearly. The rate of convergence is Q-quadratic if F and G are SC¹ functions.

Proof. Since x^* is an accumulation point of $\{x_k\}$, there exists a subsequence $\{x_k\}_{k\in K}$ such that $x_k \to x^*$, $k \in K$. We know $x_{k+1} = x_k + (V_k)^{-1}\varphi(x_k)$ for all k sufficiently large. Therefore, the results of this theorem are guaranteed by Theorem 3.1 in [8].

Remark 11. In Method 1, (17) can be replaced by the following line search:

$$\Psi\left(x_{k}+2^{-i^{k}}d_{k}\right) \leq \max_{0\leq j\leq m(k)}\Psi\left(x_{k-j}\right)+\beta 2^{-i^{k}}\nabla\Psi\left(x_{k}\right)^{T}d_{k},$$
(32)

where m(0) = 0 and $m(k) = \min\{m(k - 1) + 1, M_0\}$, and $M_0 > 0$ is a integer.

3. Numerical Results and Discussions

In this section, we present some numerical results and also give some discussions about a method for calculating a generalized Jacobian of φ . For solving the systems of (7), we can take ∂_{\star} as a tool instead of the Clarke generalized Jacobian and *B*-differential. We give the following $\partial_{\star} \varphi(x)$ for φ in (7):

$$\partial_{\star}\varphi(x) = \left\{ \left(\partial\varphi_{1}(x), \dots, \partial\varphi_{n}(x) \right)^{T}, x \in \mathbb{R}^{n} \right\}, \quad (33)$$

where $\partial \varphi_i(x) = \nabla F_i(x)$, if $F_i(x) < G_i(x)$, $\partial \varphi_i(x) = \nabla F_i(x)$ or $\partial \varphi_i(x) = \nabla G_i(x)$ if $F_i(x) = G_i(x)$, and $\partial \varphi_i(x) = \nabla G_i(x)$, if $F_i(x) > G_i(x)$.

Proposition 12. Suppose that $\varphi(x)$ and $\partial_*\varphi(x)$ are defined by (7) and by (33); then all $V \in \partial_*\varphi(x)$ are nonsingular.

Example 13. We consider the generalized complementarity problem (1), where the functions

$$F(x_1, x_2) = (x_1^2, x_2^2)^T,$$

$$G(x_1, x_2) = (x_1^2 + 10, x_2^2 + 1)^T.$$
(34)

Both F and G are $R^2 \rightarrow R^2$ continuously differentiable functions.

We use Method 1 to compute Example 13. Results for Example 13 with initial points $x_0 = (10, 1)^T$, $x_0 = (100, 1000)^T$, $x_0 = (1000, 10000)^T$, $x_0 = (1000000, 1000000)^T$ and $x_0 = (10000000, 10000000)^T$, $\epsilon = 10^{-4}$ are presented in Table 1.

In Method 1, if we replace (17) by (32), we also can use this method to compute Example 13. Results for

TABLE 1

Step	$\Psi(x)$		
$x_0 = (10, 1)^T$			
16	9.797290497304267e - 005		
$x_0 = (100, 100)^T$			
22	8.832183027429608e - 005		
$x_0 = (1000, \ 10000)^T$			
48	9.248483401180785e - 005		
$x_0 = (100000, 100000)^T$			
32	8.742500968894618e - 005		
	$x_0 = (10000000, 10000000)^T$		
37	9.944195523003963e - 005		

TABLE 2

$\Psi(x)$		
$\Gamma(\lambda)$		
$x_0 = (10, 1)^T$		
9.172858713357712e-005		
$x_0 = (1000, 10000)^T$		
9.291914775655386e – 005		
$x_0 = (100, 100)^T$		
9.773252141220823e - 005		
$x_0 = (100000, 100000)^T$		
9.700029527337506e - 005		
(1000000, 1000000) ^T		
8.943237635050672e - 005		

Example 13 with initial points $x_0 = (10, 1)^T$, $x_0 = (1000, 10000)^T$, $x_0 = (100, 100)^T$, $x_0 = (1000000, 1000000)^T$ and $x_0 = (10000000, 10000000)^T$, $\epsilon = 10^{-4}$ are presented in Table 2.

Discussion 1. From the numerical results for Method 1 in Tables 1 and 2, we can see that (17) work as well as (32). So we can use (17) or (32) in practice.

In Method 1, we also can replace (15) by (35) or (36). Then Method 1 becomes the Levenberg-Marquardt method and the modified Levenberg-Marquardt method, which were given in [4, 10]. d_k was computed by

$$\left(\left(V_{k}\right)^{T}V_{k}+\sigma_{k}I\right)d=-\left(V_{k}\right)^{T}\varphi\left(x_{k}\right), \quad V_{k}\in\partial\varphi\left(x\right) \quad (35)$$

or

$$\left(\left(V_{k}\right)^{T}V_{k} + \operatorname{diag}\left(\lambda_{i}^{(k)}\varphi_{i}\left(x_{k}\right)\right)\right)d$$

$$= -\left(V_{k}\right)^{T}\varphi\left(x_{k}\right), \quad V_{k} \in \partial\varphi\left(x\right).$$
(36)

We use Method 1 (d_k was computed by (15)), the Levenberg-Marquardt method (d_k was computed by (35)), and the modified Levenberg-Marquardt method (d_k was computed by (36)) to compute Example 13. Results for Example 13 with initial points $x_0 = (1, 1)^T$ and $x_0 = (10, 10)^T$ are presented in Table 3.

TABLE 3	
$x_0 = (1, 1)$	$)^T$
Step (d_k was computed by (12))	$\Psi(x)$
16	9.441173879087828e-005
$x_0 = (1, 1)$) ^T
Step (d_k was computed by (35))	$\Psi(x)$
22	9.162651686955847 <i>e</i> - 005
$x_0 = (1, 1)$) ^T
Step (d_k was computed by (36))	$\Psi(x)$
42	9.397247497289207 <i>e</i> - 005
$x_0 = (10, 1)$	0) ^T
Step (d_k was computed by (12))	$\Psi(x)$
18	8.575708886859879 <i>e</i> - 005
$x_0 = (10, 1)$	0) ^{<i>T</i>}
Step (d_k was computed by (35))	$\Psi(x)$
21	9.313859056403473 <i>e</i> - 005
$x_0 = (10, 1)$	$\left(0\right)^{T}$
Step (d_k was computed by (36))	$\Psi(x)$
24	9.637003553941703e - 005

TABLE	Т		
$x_0 = (1, 1)$	$1)^T$		
Step (d_k was computed by (12))	$\Psi(x)$		
14	9.696887098398221e - 005		
$x_0 = (1, 1)$	$1)^T$		
Step (d_k was computed by (35))	$\Psi(x)$		
21	9.401294082644895e - 005		
$x_0 = (1, 1)^T$			
Step (d_k was computed by (36))	$\Psi(x)$		
22	9.162651686955847e - 005		

TABLE 4

Discussion 2. From the numerical results for the Method 1 (d_k was computed by (15)), the Levenberg-Marquardt method (d_k was computed by (35)), and the modified Levenberg-Marquardt method (d_k was computed by (36)) in Table 3, we can see that our method in this paper works quite better than the methods in [4, 10].

From Discussion 1, we can see that (17) work as well as (32). When we use line search (32) to replace (17), we give the following numerical results for Method 1 (d_k was computed by (15)), the Levenberg-Marquardt method (d_k was computed by (35)), and the modified Levenberg-Marquardt method (d_k was computed by (36)) to compute Example 13. The numerical results are given in Table 4.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by National Science Foundation of China (11101231).

References

- C. Kanzow and M. Fukushima, "Equivalence of the generalized complementarity problem to differentiable unconstrained minimization," *Journal of Optimization Theory and Applications*, vol. 90, no. 3, pp. 581–603, 1996.
- [2] H. Jiang, M. Fukushima, L. Qi, and D. Sun, "A trust region method for solving generalized complementarity problems," *SIAM Journal on Optimization*, vol. 8, no. 1, pp. 140–157, 1998.
- [3] T. de Luca, F. Facchinei, and C. Kanzow, "A semismooth equation approach to the solution of nonlinear complementarity problems," *Mathematical Programming*, vol. 75, no. 3, pp. 407– 439, 1996.
- [4] F. Facchinei and C. Kanzow, "A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems," *Mathematical Programming*, vol. 76, no. 3, pp. 493–512, 1997.
- [5] Y. Gao, "Newton methods for solving two classes of nonsmooth equations," *Applications of Mathematics*, vol. 46, no. 3, pp. 215– 229, 2001.
- [6] L. Qi and P. Tseng, "On almost smooth functions and piecewise smooth functions," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 67, no. 3, pp. 773–794, 2007.
- [7] D. P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, New York, NY, USA, 1982.
- [8] L. Q. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations," *Mathematics of Operations Research*, vol. 18, no. 1, pp. 227–244, 1993.
- [9] C. Wang, Q. Liu, and C. Ma, "Smoothing SQP algorithm for semismooth equations with box constraints," *Computational Optimization and Applications*, vol. 55, no. 2, pp. 399–425, 2013.
- [10] S.-Q. Du and Y. Gao, "The Levenberg-Marquardt-type methods for a kind of vertical complementarity problem," *Journal of Applied Mathematics*, vol. 2011, Article ID 161853, 12 pages, 2011.
- [11] X. Chen, "Smoothing methods for nonsmooth, nonconvex minimization," *Mathematical Programming*, vol. 134, no. 1, pp. 71–99, 2012.
- [12] X. Chen, L. Qi, and D. Sun, "Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities," *Mathematics* of Computation, vol. 67, no. 222, pp. 519–540, 1998.