Journal of Computational Mathematics, Vol.20, No.5, 2002, 533–542.

A NEW SMOOTHING APPROXIMATION METHOD FOR SOLVING BOX CONSTRAINED VARIATIONAL INEQUALITIES*

Chang-feng Ma
(*Institute of Mathematics, Academy of Mathematics & System Sciences, CAS, Beijing, China*)
(*Dept. of Math. & Computer, Changsha University of Electric Power, Changsha, China*)
Guo-ping Liang Shao-peng Liu
(*Institute of Mathematics, Academy of Mathematics & System Sciences, CAS, Beijing, China*)

Abstract

In this paper, we first give a smoothing approximation function of nonsmooth system based on box constrained variational inequalities and then present a new smoothing approximation algorithm. Under suitable conditions,we show that the method is globally and superlinearly convergent. A few numerical results are also reported in the paper.

Key words: Box constrained variational inequalities, Smoothing approximation, Global convergence, Superlinear convergence.

1. Introduction

Consider the following box constrained variational inequalities of finding an $x \in X$, such that

$$f(x)^T(y-x) \geq 0, \quad \forall y \in X \tag{1.1}$$

where f is a smooth mapping from R^n into itself, set X has the following box form:

$$X = \{x \in R^n | a \leq x \leq b\} \tag{1.2}$$

with $a = (a_1, a_2, \cdots, a_n)^T,\ b = (b_1, b_2, \cdots, b_n)^T$ and $-\infty \leq a_i < b_i \leq +\infty, i = 1, \cdots, n$.

It is easily shown that problem (1.1) is equivalent to

$$H(x) = 0 \tag{1.3}$$

where $H(x) = (H_1(x), \cdots, H_n(x))^T,\ H_i(x)(i = 1, \cdots, n)$ defined by

$$H_i(x) = \begin{cases} x_i - a_i, & x_i - f_i(x) \leq a_i, \\ f_i(x), & a_i < x_i - f_i(x) < b_i, \\ x_i - b_i, & x_i - f_i(x) \geq b_i, \end{cases} \tag{1.4}$$

Generally, mapping H is nonsmooth. Qi and Chan(see[6]) established a so-called successive approximation method, which changes (1.3) into an equivalent nonsmooth equations of the following form:

$$H(x) \equiv f_k(x) + g_k(x) = 0 \tag{1.5}$$

where f_k and g_k are mappings from R^n into itself with f_k being F-differentiable and g_k not, but its norm relatively small. Global convergence of this method is obtained and numerical examples are given to illustrate its usefulness.

Another approach for solving problem (1.3) is the smoothing methods(see[7, 8, 9]). The feature of smoothing methods is to construct a smoothing approximation function $H(\cdot, \cdot):\ R^n \times R_{++} \to R^n$ of H such that for any $\epsilon > 0,\ H(\cdot, \epsilon)$ is continuously differentiable and

$$\|H(x) - H(x, \epsilon)\| \to 0 \ as\ \epsilon \to 0^+ \ for\ all\ x \in R^n$$

* Received June 23, 1999; Final revised November 30, 2000.

and then to find a solution of (1.3) by solving the following problem for a given sequence $\{\epsilon_k\}, k = 0, 1, \cdots,$

$$H(x, \epsilon_k) = 0. \tag{1.6}$$

In [7], Chen and Mangasarian introduced a class of smoothing approximation functions for nonlinear complementarity problems. And Gabriel and Moré extended Chen-Mangasarian's smoothing approximation functions to box constrained variational inequalities(see[9]).

In this paper, we will present a new smoothing approximation method for solving problem (1.1). Under mild conditions, we prove that the sequence $\{x_k\}$ generated by this method is bounded and each accumulation point is a solution of problem (1.3).

This paper is organized as follows. In the next section, we define the Jacobian consistency property and construct a new smoothing approximation function of H. And then we describe the algorithm in detail. In section 3, we show that the algorithm is globally and superlinearly convergent. Finally, some numerical examples are given in section 4.

We let $\|\cdot\|$ denote the Euclidean norm of R^n and let

$$R_+ = \{\epsilon | \epsilon \geq 0,\ \epsilon \in R\}, \qquad R_{++} = \{\epsilon | \epsilon > 0,\ \epsilon \in R\},$$

we denote the set of all nonnegative integers by N=$\{0,\ 1,\ 2, \cdots\}$.

2. The Algorithm

Let $F : R^n \to R^m$ be locally Lipschitz continuous. According to Rademacher's theorem, H is differentiable almost everywhere. Let D_F be the set where F is differentiable. The B-derivative of F is defined by(see[3])

$$\partial_B F(x) = \{ \lim_{x_k (\in D_F) \to x} F'(x_k) \}.$$

The generalized Jacobian of F at x in the sense of Clarke(see[1]) is

$$\partial F = conv \partial_B F(x).$$

In this paper, for the function F, we use a kind of generalized Jacobian, denote by $\partial_C F$ and defined as(see[10])

$$\partial F = \partial F_1(x) \times \partial F_2(x) \times \cdots \partial F_n(x).$$

Next we give the definition of the Jacobian consistency property.

Definition 2.1. Let $H(\cdot)$ be a Lipschitz function in R^n. We call $H(\cdot, \epsilon) : R^n \times R_{++} \to R^n$ a smoothing approximation function of $H(\cdot)$ if $H(\cdot, \epsilon)$ is continuously differentiable with respect to variable x and there exists a constant $c > 0$ such that for any $x \in R^n$ and $\epsilon \in R_{++}$,

$$\|H(x, \epsilon) - H(x)\| \leq c\epsilon \tag{2.1}$$

Further, if for any $x \in R^n$,

$$\lim_{\epsilon \downarrow 0} dist(H'_x(x, \epsilon), \partial_C H(x)) = 0 \tag{2.2}$$

then we say $H(\cdot, \epsilon)$ satisfies the Jacobian consistency property.

We now construct a new smoothing approxiomation function $H(x, \epsilon) = (H_i(x, \epsilon))$ of $H(x)$ defined by (1.4) as follows

$$H_i(x, \epsilon) = \begin{cases} x_i - a_i - \frac{(x_i - f_i(x) - a_i + \epsilon)^2}{4\epsilon}, & i \in \alpha(x), \\ x_i - b_i + \frac{(x_i - f_i(x) - b_i - \epsilon)^2}{4\epsilon}, & i \in \beta(x), \\ H_i(x), & otherwise \end{cases} \tag{2.3}$$

where $\alpha(x) = \{i : |x_i - f_i(x) - a_i| < \epsilon\}$, $\beta(x) = \{i : |x_i - f_i(x) - b_i| < \epsilon\}$. It isn't difficult to find that $H(x, \epsilon)$ is continuously differential while $\epsilon \leq \min\limits_{1 \leq i \leq n} \{\frac{1}{2}(b_i - a_i)\}$. And Jacobian $H'_x(x, \epsilon)$ is of following form

$$[H'_x(x,\epsilon)]_i = \begin{cases} e_i - \frac{x_i - f_i(x) - a_i + \epsilon}{2\epsilon}(e_i - f'_i(x)), & i \in \alpha(x), \\ e_i + \frac{x_i - f_i(x) - b_i - \epsilon}{2\epsilon}(e_i - f'_i(x)), & i \in \beta(x), \\ H'_i(x), & otherwise \end{cases} \tag{2.4}$$

Now we show that $H(x,\epsilon)$ has the Jacobian consistency property. Indeed, for $i = 1,\ 2, \cdots, n$, it is easy to get

$$\partial_B H_i(x) = \begin{cases} \{e_i\}, & if\ x_i - f_i(x) \notin [a_i,\ b_i], \\ \{f'_i(x)\}, & if\ x_i - f_i(x) \in (a_i,\ b_i), \\ \{e_i,\ f'_i(x)\}, & if\ x_i - f_i(x) = a_i\ or\ x_i - f_i(x) = b_i \end{cases} \tag{2.5}$$

Let

$$\gamma(x) = \min_{1 \le i,\ j \le n} \left\{ \begin{array}{ll} |x_i - f_i(x) - a_i|, & |x_j - f_j(x) - b_j| : \\ |x_i - f_i(x)| \neq a_i, & |x_j - f_j(x) \neq b_j \end{array} \right\} \tag{2.6}$$

and

$$\bar{\epsilon} = \min_{1 \le i \le n} \{\frac{1}{2}(b_i - a_i)\} \tag{2.7}$$

By (2.4), for any $\epsilon \in [0,\ \epsilon(x)]$, we have

$$\lim_{\epsilon \downarrow 0} [H'_x(x,\epsilon)]_i = \begin{cases} \{e_i\}, & if\ x_i - f_i(x) \notin [a_i,\ b_i], \\ \{f'_i(x)\}, & if\ x_i - f_i(x) \in (a_i,\ b_i), \\ \frac{1}{2}(e_i + f'_i(x)), & if\ x_i - f_i(x) = a_i\ or\ x_i - f_i(x) = b_i \end{cases} \tag{2.8}$$

where $\epsilon(x) = \min\{\bar{\epsilon},\ \gamma(x)\}$. By (2.5), we obtain

$$\lim_{\epsilon \downarrow 0} dist(H'_x(x,\epsilon), \quad \partial_C H(x)) = 0$$

i.e., function $H(x,\epsilon)$ satisfies the Jacobian consistency property. We next describe the algorithm in detail and denote

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

and

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

Algorithm 2.1.

step 0. Given $\rho,\ \alpha,\ \eta \in (0,\ 1)$, $\gamma \in (0,\ +\infty)$ and $x_0 \in R^n$. Choose $\sigma \in (0,\ \frac{3}{4}(1-\alpha))$ and $c > 0$ satisfying (2.1). Let $\beta_0 = \|H(x_0)\|$ and $\epsilon_0 = \frac{\alpha}{2c}\beta_0$. Set $k := 0$.

step 1. Solve linear equations of d:

$$H(x_k) + H'_x(x_k,\epsilon_k)d = 0. \tag{2.9}$$

Let d_k be the solution of (2.9).

step 2. Let m_k be the smallest nonnegative integer m such that

$$\theta(x_k + \rho^m d_k,\ \epsilon_k) - \theta(x_k,\ \epsilon_k) \le -2\sigma\rho^m\theta(x_k). \tag{2.10}$$

Set $t_k = \rho^{m_k}$ and $x_{k+1} = x_k + t_k d_k$.

step 3. If $\|H(x_{k+1})\| = 0$, stop; otherwise, go to step 4.

step 4. If $\|H(x_{k+1})\| > 0$ and

$$\|H(x_{k+1})\| \le \max\{\eta\beta_k,\ \alpha^{-1}\|H(x_{k+1}) - H(x_{k+1},\ \epsilon_k)\|\} \tag{2.11}$$

we let

$$\beta_{k+1} = \|H(x_{k+1})\|$$

and choose an ϵ_{k+1} satisfying

$$0 < \epsilon_{k+1} \leq \min\{\frac{\alpha}{2c}\beta_{k+1},\ \frac{\epsilon_k}{2}\} \tag{2.12}$$

and

$$dist(H'_x(x_{k+1},\ \epsilon_{k+1}),\ \partial_C H(x_{k+1})) \leq \gamma\beta_{k+1}. \tag{2.13}$$

If $\|H(x_{k+1})\| > 0$, but (2.11) does not hold, we let $\beta_{k+1} := \beta_k$ and $\epsilon_{k+1} := \epsilon_k$.

step 5. Set $k := k+1$, go to step 1.

Remark. The constant c in step 0 of Algorithm 2.1 can be chosen properly on interval [0.5, 1]. Indeed,by (1.4) and (2.3),it is not difficult to get

$$\begin{aligned}
&|H_i(x,\epsilon) - H_i(x)| \\
\leq &\max\{(\frac{(x_i - f_i(x) - a_i + \epsilon)^2}{4\epsilon})_{i\in\alpha(x)}, (\frac{(x_i - f_i(x) - b_i - \epsilon)^2}{4\epsilon})_{i\in\beta(x)}\} \\
\leq &\frac{(\epsilon+\epsilon)^2}{4\epsilon} = \epsilon.
\end{aligned}$$

3. Convergence Analysis

In this section, we prove that Algorithm 2.1 is globally and superlinearly convergent. Without loss of generality, we assume that $\|H(x_k)\| > 0$ for all k in the following convergence analysis.

Lemma 3.1. *Suppose that $H'_x(x_k,\ \epsilon_k)$ is nonsingular. Then there exists a finite nonnegative integer m_k such that (2.10) holds.*

Proof. From the construction of Algorithm 2.1, we have

$$\|H(x_k) - H(x_k,\epsilon_k)\| \leq \alpha\|H(x_k)\|. \tag{3.1}$$

The continuous differentiability of $H(\cdot,\epsilon_k)$ implies that $\theta(\cdot,\epsilon_k)$ is continuously differentiable and $\theta'(x_k,\epsilon_k) = H'_x(x_k,\epsilon_k)^T H(x_k,\epsilon_k)$. By (2.9), $H'_x(x_k,\epsilon_k)d_k = -H(x_k)$. Then, from (3.1), we get

$$\begin{aligned}
&\theta(x_k + td_k,\epsilon_k) - \theta(x_k,\epsilon_k) \\
=&t\theta'(x_k,\epsilon_k)d_k + o(t) = -tH(x_k)^T H(x_k,\epsilon_k) + o(t) \\
=&-2t\theta(x_k) + tH(x_k)^T[H(x_k) - H(x_k,\epsilon_k)] + o(t) \\
\leq&-2t\theta(x_k) + 2t\alpha\theta(x_k) + o(t) \\
=&-2t(1-\alpha)\theta(x_k) + o(t).
\end{aligned}$$

Since $\sigma \leq \frac{3}{4}(1-\alpha) < 1-\alpha$, there is a finite nonnegative integer m_k such that (2.10) holds.

Condition A. (1) The level set

$$\Omega = \{x \in R^n | \theta(x) \leq (1+\alpha)^2\theta(x_0)\}$$

is bounded;

(2) For any $\epsilon \in R_{++}$ and $x \in \Omega$, $H'_x(x,\epsilon)$ is nonsingular.

Theorem 3.1. *Suppose that Condition A holds. And the sequence $\{x_k\}$ is generated by Algorithm 2.1. Then*

(i) Algorithm 2.1 is well defined and

$$\{x_k\} \subset \Omega. \tag{3.2}$$

(ii)

$$\lim_{k\to\infty} H(x_k) = 0. \tag{3.3}$$

Proof. Denote

$$K=\{0\}\cup\{k:\|H(x_k)\|\le\max\{\eta\beta_{k-1},\ \alpha^{-1}\|H(x_k)-H(x_k,\epsilon_{k-1})\|\},\ k\in N\},\tag{3.4}$$

and

$$K_1=\{k\in K|\eta\beta_{k-1}\ge\alpha^{-1}\|H(x_k)-H(x_k,\epsilon_{k-1})\|\},$$
$$K_2=\{k\in K|\eta\beta_{k-1}<\alpha^{-1}\|H(x_k)-H(x_k,\epsilon_{k-1})\|\}.$$

Clearly, $K=K_1\cup K_2\cup\{0\}$. Assume that $K=\{k_0,\ k_1\ ,k_2,\cdots\}$ with $0=k_0<k_1<k_2<\cdots$. Let k be an arbitrary nonnegative integer and k_j be the largest number in K such that $k_j\le k$. Then

$$\epsilon_k=\epsilon_{k_j}\quad and\quad \beta_k=\beta_{k_j}.$$

By (2.10), notice that $H(x,\epsilon)$ is a smoothing approximation function,

$$\|H(x_k,\epsilon_{k_j})\|\le\|H(x_{k_j},\epsilon_{k_j})\|.$$

Then by (2.1), for any $j\ge 0$,

$$\begin{aligned}&\|H(x_k)\|\le\|H(x_k,\epsilon_k)\|+\|H(x_k)-H(x_k,\epsilon_k)\|\\&=\|H(x_k,\epsilon_{k_j})\|+\|H(x_k)-H(x_k,\epsilon_{k_j})\|\\&\le\|H(x_{k_j},\epsilon_{k_j})\|+c\epsilon_{k_j}\le\|H(x_{k_j})\|+c\epsilon_{k_j}+c\epsilon_{k_j}\\&=\beta_{k_j}+2c\epsilon_{k_j}.\end{aligned}\tag{3.5}$$

If $j=0,\ \beta_{k_j}=\beta_0,\ \epsilon_{k_j}=\epsilon_0$ and

$$\|H(x_k)\|\le\beta_0+2c\epsilon_0\le(1+\alpha)\|H(x_0)\|.$$

If $j\ge 1$, by step 4 of Algorithm 2.1,

$$\epsilon_{k_j}\le\frac{1}{2}\epsilon_{k_j-1}=\frac{1}{2}\epsilon_{k_{j-1}},$$

and

$$\beta_{k_j}\le\eta\beta_{k_j-1}=\eta\beta_{k_{j-1}},\quad if\ k_j\in K_1.$$

or

$$\beta_{k_j}\le\alpha^{-1}\|H(x_{k_j},\epsilon_{k_j-1})-H(x_{k_j-1})\|\le c\alpha^{-1}\epsilon_{k_j-1}=c\alpha^{-1}\epsilon_{k_{j-1}},\ if\ k_j\in K_2$$

Let

$$r=\max\{\frac{1}{2},\ \eta\}.$$

Then for $j\ge 1$, by the definition of ϵ_0 and β_0,

$$\epsilon_{k_j}\le\frac{1}{2^{j-1}}\epsilon_0=\frac{1}{2^j}\cdot\frac{\alpha}{c}\|H(x_0)\|,\tag{3.6}$$

and

$$\beta_{k_j}\le r^{j-1}\beta_0=r^{j-1}\|H(x_0)\|.\tag{3.7}$$

Therefore by (3.5), for $j\ge 1$,

$$\|H(x_k)\|\le(r^{j-1}+\frac{\alpha}{2^{j-1}})\|H(x_0)\|\le r^{j-1}(1+\alpha)\|H(x_0)\|.\tag{3.8}$$

Hence in two cases,

$$\|H(x_k)\|\le(1+\alpha)\|H(x_0)\|.$$

This implies that (3.2) holds.

Now we prove that the second part of the theorem. If K is infinite, by (3.8),

$$\lim_{k\to\infty}\|H(x_k)\|\le\lim_{j\to\infty}r^{j-1}(1+\alpha)\|H(x_0)\|=0.$$

This proves (3.3). Therefore, to prove (3.3), it suffices to prove that set K is infinite. Suppose that K is finite, this means that both K_1 and K_2 are finite. Let $\hat{k}$ be the largest number in K. Then for all $k > \hat{k}$,

$$\epsilon_k \equiv \epsilon_{\hat{k}}, \quad \beta_k \equiv \beta_{\hat{k}} = \|H(x_{\hat{k}})\|. \tag{3.9}$$

Thus, we have

$$\|H(x_k)\| > \eta\beta_k = \eta\|H(x_{\hat{k}})\| > 0, \tag{3.10}$$

and

$$\alpha\|H(x_k)\| > \|H(x_k, \epsilon_{\hat{k}}) - H(x_k)\|. \tag{3.11}$$

By (3.10), for all $k > \hat{k}$,

$$\theta(x_k) \geq \eta^2\theta(x_{\hat{k}}). \tag{3.12}$$

Let

$$\hat{\epsilon} = \epsilon_{\hat{k}}, \qquad \hat{H}(x) = H(x, \hat{\epsilon}),$$

and

$$\hat{\theta}(x) = \frac{1}{2}\|\hat{H}(x)\|^2.$$

Notice that for all $k > \hat{k}$,

$$H(x_k, \epsilon_k) = \hat{H}(x_k) \ \ \textit{and} \ \ \theta(x_k, \epsilon_k) = \hat{\theta}(x_k)$$

By Condition A, there exists an $M > 0$ such that for all $x \in \Omega$, $\|H'_x(x, \hat{\epsilon})^{-1}\| \leq M$. Then for all $k > \hat{k}$,

$$\|d_k\| = \|H'_x(x_k, \hat{\epsilon})^{-1} \cdot H(x_k)\| \leq M\|H(x_k)\| \leq M(1+\alpha)\|H(x_0)\| =: L$$

If $\inf_k \rho^{m_k} = t^* > 0$, then from (3.12) and (2.10), for all $k \geq 0$,

$$\hat{\theta}(x_{k+1}) - \hat{\theta}(x_k) \leq -2\sigma\rho^{m_k}\theta(x_k) \leq -2\sigma t^*\eta^2\theta(x_{\hat{k}}) < 0.$$

This, together with the monotonicity of $\{\hat{\theta}(x_k) : \ k \geq \hat{k}\}$, implies that $\hat{\theta}(x_k) \to -\infty$ as $k \to \infty$. This contradicts the fact that $\hat{\theta}(x_k) \geq 0$ for all $k \geq 0$. Therefore K can not be finite. Thus (3.3) holds.

We now consider the case $t^* = 0$. Let K_0 be a subsequence of N such that $\{t_k : k \in K_0\}$ converges to zero. Since $\{x_k\}$ is bounded, we assume that $\{x_k : k \in K_0\}$ converges to x^*. By (2.10), for all $k > \hat{k}$,

$$\hat{\theta}(x_k + \rho^{m_k - 1}d_k) - \hat{\theta}(x_k) > -2\sigma\rho^{m_k-1}\theta(x_k). \tag{3.13}$$

That is,

$$\begin{aligned}
-2\sigma\theta(x_k) < & \frac{\hat{\theta}(x_k + \rho^{m_k-1}d_k) - \hat{\theta}(x_k)}{\rho^{m_k-1}} \\
&= \hat{\theta}'(x_k)d_k + o(1) = -H(x_k)^T\hat{H}(x_k) + o(1) \\
&= -2\theta(x_k) + H(x_k)[\hat{H}(x_k) - H(x_k)] + o(1) \\
&\leq -2\theta(x_k) + 2\alpha\theta(x_k) + o(1) \\
&= -2(1-\alpha)\theta(x_k) + o(1)
\end{aligned}$$

By taking the limit in the above inequality on the subsequence $k \in K_0$, we get

$$-2\sigma\theta(x^*) \leq -2(1-\alpha)\theta(x^*).$$

This implies $\sigma > 1 - \alpha$, which contradicts the fact that $\sigma \leq \frac{3}{4}(1-\alpha)$. Hence K cannot be finite. Thus (3.3) holds.

Next we discuss convergence rate of Alorithm 2.1. From the construction of the algorithm, for $k \in K,\ k \geq 1$,

$$dist(H'_x(x_k, \epsilon_k),\ \partial_C H(x_k)) \leq \gamma \|H(x_k)\|, \tag{3.14}$$

where K is defined in (3.4). To verify that superlinear convergence ratre, we first give the following lemma.

Lemma 3.2. *If there is a scalar*

$$\lambda \in [\frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2},\ \frac{1}{2}], \qquad 3.15)$$

such that for some $k \in K$,

$$\theta(x) - \theta(x_k) \leq -2\lambda\theta(x_k), \tag{3.16}$$

then the following relation holds

$$\theta(x, \epsilon_k) - \theta(x_k, \epsilon_k) \leq -2\sigma\theta(x_k). \tag{3.17}$$

Proof. By (3.4), we have

$$0 < \epsilon_k \leq \frac{\alpha}{2c}\|H(x_k)\|, \quad k \in K.$$

Therefore, from (2.1), for any $x \in R^n,\ k \in K$,

$$\|H(x,\ \epsilon_k)\| \leq \|H(x)\| + \frac{\alpha}{2}\|H(x_k)\|,$$

and

$$\|H(x_k,\ \epsilon_k)\| \geq \|H(x_k)\| - \frac{\alpha}{2}\|H(x_k)\|.$$

Using (4.3) and the above two inequalities, we get

$$\begin{aligned}
&\theta(x, \epsilon_k) - \theta(x_k, \epsilon_k) = \frac{1}{2}\|H(x,\epsilon)\|^2 - \frac{1}{2}\|H(x_k, \epsilon)\|^2 \\
\leq& \frac{1}{2}(\|H(x)\| + \frac{\alpha}{2}\|H(x_k)\|)^2 - \frac{1}{2}(1-\frac{\alpha}{2})^2\|H(x_k)\|^2 \\
=&\theta(x) + \frac{1}{2}\alpha\|H(x)\| \cdot \|H(x_k)\| + \frac{\alpha^2}{4}\theta(x_k) - (1-\alpha+\frac{\alpha^2}{4})\theta(x_k) \\
\leq&\theta(x) + \frac{1}{2}\alpha\sqrt{1-2\lambda}\|H(x_k)\| \cdot \|H(x_k)\| - (1-\alpha)\theta(x_k) \\
\leq&(1-2\lambda)\theta(x_k) + \alpha\sqrt{1-2\lambda}\theta(x_k) - (1-\alpha)\theta(x_k) \\
=&(\alpha + \alpha\sqrt{1-2\lambda} - 2\lambda)\theta(x_k).
\end{aligned} \tag{3.18}$$

Denote

$$\varphi(\lambda) = \alpha + \alpha\sqrt{1-2\lambda} - 2\lambda,$$

and

$$\bar{\lambda} = \frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2}.$$

It is not difficult to find that $\varphi(\lambda)$ is monotone decreasing in $[\bar{\lambda},\ \frac{1}{2}]$. Therefore, for any $\lambda \in [\bar{\lambda},\ \frac{1}{2}]$,

$$\begin{aligned}
&\varphi(\lambda) \leq \varphi(\bar{\lambda}) = \alpha + \alpha\sqrt{1 - 2[\frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2}]} - 2[\frac{1}{2} - \frac{(1-\alpha-2\sigma)^2}{2(2+\alpha)^2}] \\
=&\alpha + \alpha\frac{1-\alpha-2\sigma}{2+\alpha} - 1 + \frac{(1-\alpha-2\sigma)^2}{(2+\alpha)^2} \leq \alpha + \alpha\frac{1-\alpha-2\sigma}{2+\alpha} - 1 + 2\frac{1-\alpha-2\sigma}{2+\alpha} \\
=& - 2\sigma.
\end{aligned}$$

Thus, it follows from (3.17) and the above inequality that (3.18).

Theorem 3.2. *Suppose that Condition A holds. If for an accumulation x^* of the sequence $\{x_k\}$ generated by Algorithm 2.1, all $V \in \partial_C H(x^*)$ are nonsingular, then x^* is a solution of $H(x) = 0$ and $\{x_k\}$ converges to x^* superlinearly.*

Proof. By Theorem 3.1, x^* is a solution of $H(x) = 0$ and the set K defined by (3.4) is infinite. Hence, there exists a subsequence $\{x_k : k \in K_0 \subset K\}$ converges to x^*.

It is not difficult to verify that the $H(x)$ defined by (1.4) is semismooth at x^* from the definition of semismooth function(see[4]). Notice that for any $x \in R^n$, $\partial_C H(x)$ is a compact set. Let $V_k \in \partial_C H(x_k)$ be such that

$$dist(H'_x(x_k, \epsilon_k),\ \partial_C H(x_k)) = \|H'_x(x_k, \epsilon_k) - V_k\|.$$

By Algorithm 2.1,

$$\|H'_x(x_k, \epsilon_k) - V_k\| \leq \gamma\beta_k, \quad k \in K_0.$$

By Theorem 3.1, $\beta_k \to 0 (k \to \infty)$. This implies that there exist $M > 0$ and $\hat{k} \geq 0$ such that for all $k \geq \hat{k}$ and $k \in K_0$, $\|H'_x(x_k, \epsilon_k)^{-1}\| \leq M$. Thus, we have

$$\begin{aligned}
&\|x_k + d_k - x^*\| = \|x_k - x^* - H'_x(x_k, \epsilon_k)^{-1} H(x_k)\| \\
\leq&\|H'_x(x_k, \epsilon_k)^{-1}\| \|H'_x(x_k, \epsilon_k)(x_k - x^*) - H(x_k) + H(x^*)\| \qquad (3.19)\\
\leq&\|H'_x(x_k, \epsilon_k)^{-1}\| (\|(H'_x(x_k, \epsilon_k) - V_k)(x_k - x^*)\| + \|V_k(x_k - x^*) - H(x_k) + H(x^*)\|) \\
\leq&M(\gamma\beta_k \|x_k - x^*\| + \|H(x_k) - H(x^*) - V_k(x_k - x^*)\|)
\end{aligned}$$

Since H is semismooth at x^* if and only if each H_i is semismooth at x^*(see[4]), we have

$$\begin{aligned}
&\|H(x_k) - H(x^*) - V_k(x_k - x^*)\| \\
\leq&\sqrt{\sum_{i=1}^{n} \|H_i(x_k) - H_i(x^*) - V_k^i(x_k - x^*)\|} \qquad (3.20)\\
=&o(\|x_k - x^*\|), \quad (k \to \infty,\ k \in K_0),
\end{aligned}$$

where V_k^i denotes the ith row of V_k. Hence, by (3.19),

$$\|x_k + d_k - x^*\| = o(\|x_k - x^*\|), \quad (k \to \infty,\ k \in K_0). \qquad (3.21)$$

Furthermore, we have(see[3]),

$$\|H(x_k + d_k)\| = o(\|H(x_k)\|),\ (k \to \infty,\ k \in K_0). \qquad (3.22)$$

Let

$$\lambda = \max\{\frac{1 - \eta^2}{2},\ \frac{1}{2} - \frac{(1 - \alpha - 2\sigma)^2}{2(2 + \alpha)^2}\}.$$

Then (3.22) implies that there exist $\bar{k} \geq \hat{k}$ such that $\bar{k} \in K_0$ and for any $k \geq \bar{k}$ and $k \in K_0$,

$$\theta(x_k + d_k) - \theta(x_k) \leq -2\lambda\theta(x_k). \qquad (3.23)$$

By Lemma 3.2, for any $k \geq \bar{k}$ and $k \in K_0$,

$$\theta(x_k + d_k, \epsilon_k) - \theta(x_k, \epsilon_k) \leq -2\sigma\theta(x_k),$$

that is, for all $k \geq \bar{k}$ and $k \in K_0$, $t_k \equiv 1$ and $x_{k+1} = x_k + d_k$. From (3.23), we get

$$\|H(x_{\bar{k}+1})\| \leq \sqrt{1 - 2\lambda}\|H(x_{\bar{k}})\| \leq \eta\|H(x_{\bar{k}})\| = \eta\beta_{\bar{k}},$$

which implies $\bar{k} + 1 \in K_0$. Repeating the above process, we may prove that for all $k \geq \bar{k}$,

$$k \in K_0 \quad and \quad x_{k+1} = x_k + d_k.$$

Then by using (3.21), we have proved that $\{x_k\}$ converges to x^* superlinearly.

4. Numerical Results

In this section, we give some numerical results. Throughout the computational experiments, the parameters used in Algorithm 2.1 were $\alpha = 0.4,\ \sigma = 0.25,\ \rho = 0.98,\ \eta = 0.5,\ c = 0.9,\ \gamma = 0.6,\ \eta = 0.8$. The stopping criterion is: $\|H(x_k)\| \le 10^{-6}$.

Example 1[11]**.** Test function is as follows:

$$f(x) = Mx + q$$

where

$$\begin{aligned}
&[M]_{ii} = 4(i-1)+1,\ i = 1,\ 2,\ \cdots, n,\\
&[M]_{ij} = [M]_{ii}+1,\ i = 1, \cdots, n-1,\ j = i+1, \cdots, n,\\
&[M]_{ij} = [M]_{jj}+1,\ j = 1, \cdots, n-1,\ i = j+1, \cdots, n.
\end{aligned}$$

and $q = (-1,\ -1, \cdots, -1)^T$, $X = \{x \in R^n | 0 \le x_i \le 50,\ i = 1, \cdots, n\}$. The solution is $x^* = (1,\ 0, \cdots, 0)^T$ and the starting point $x_0 = (1,\ 1, \cdots, 1)^T$. Table 1 lists the numerical results for different dimensions n.

Table 1

Dimensions	Num. of iter.	Final θ_value
5	3	$3.10136e-15$
10	3	$9.65366e-16$
15	3	$2.32960e-15$
20	3	$7.16800e-16$
25	2	$3.20000e-13$
30	3	$3.08229e-13$
35	3	$3.04883e-13$
40	3	$2.80314e-13$
45	3	$2.82765e-13$

Example 2[12]**.** This example was tested by Kanzow with five variables defined by

$$f_i(x) = 2(x_i - i + 2)\exp\{\sum_{i=1}^{5}(x_i - i + 2)^2\},\ 1 \le i \le 5,$$

where constraint set $X = \{x \in R^5 | 0 \le x_i \le 10,\ i = 1, \cdots, 5\}$. This example has one degenerate solution $x^* = (0,\ 0,\ 1,\ 2,\ 3)^T$. The numerical results are given in Table 2 using different staring points.

Table 2

Starting points	Num. of iter.	Final θ_value
(0, 0, 0, 0, 0)	8	$8.34940e-16$
(1, 1, 1, 1, 1)	10	$3.97507e-15$
(−1, −1, −1, −1, −1)	13	$4.07091e-14$
(10, 10, 10, 10, 10)	9	$1.05879e-15$
(3, 2, 1, 2, 3)	2	$1.66400e-13$
(1, 0, 1, 3, 5)	5	$6.48225e-15$

Example 3[13]**.** Consider that optimization problem with simple bonded constraint,

$$\begin{aligned}
&\textit{minimize}\ f(x) = 100(x_2 - x_1^2)^2 + (1-x_1)^2 + 90(x_4 - x_3^2)^2 + (1-x_3)^2\\
&s.t.\ -10 \le x_i \le 10, \qquad + 10.1[(x_2-1)^2 + (x_4-1)^2] + 19.8(x_2-1)(x_4-1)\\
&\quad i = 1, 2, 3, 4.
\end{aligned}$$

This example has one solution $x^* = (1,\ 1,\ 1,\ 1)^T$. Using the KKT condition, it can be converted into a box constrained variational inequeality problem. The numerical results are listed in Table 3 using different staring points.

Table 3

Starting points	Num. of iter.	Final θ_value
$(0,\ 0,\ 0,\ 0)$	7	$3.39176e-13$
$(1,\ 0,\ 1,\ 0)$	9	$2.55705e-13$
$(-3,-1,-3,-1)$	9	$9.66910e-13$
$(0,\ 2,\ 0,\ 2)$	8	$5.11948e-14$
$(-5,-5,-5,-5)$	18	$5.41301e-15$
$(-10,-10,-10,-10)$	34	$7.71102e-14$

References

[1] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[2] P.T. Harker and J.S. Pang, Finite_dimensional variational inequality and nonliner complementarity problems, a survey review of theory, algorithms and applications, *Math. Prog.* **48** (1990), 161–220.

[3] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.,* **18** (1993), 227–244.

[4] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Math. Prog.* **58** (1993), 353–367.

[5] J.S. Pang, Complementarity problems, in:R. Horst and P. Pardalos, eds., "Handbook of Global Optimization", Kluwer Academic Publishers, Boston. pp271–338, 1995.

[6] L. Qi and X. Chen, A globally convergent successive approximation method for severely nonsmooth equtions, *SIAM J. Control & Optim.,* **33** (1995), 420–418.

[7] C. Chen and O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, *Comp. Optim. Appl.,* **5** (1996), 97–138.

[8] B. Chen and P.T. Harker, Smooth approximation to nonliear complementarity problems, *SIAM J. Optim.,* **7** (1997), 403–420.

[9] S.A. Gabriel and J.J. Moré, Smoothing of mixed complementarity problems, in: M.C. Ferris and J.S. Pang, eds., "Complementarity and variational problems: State of the Art", SIAM, Philadelphia, Pennsylvania, 1997, 105–116.

[10] 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, *Math. of Computation,* **67** (1998), 519–540.

[11] A. Friedlander, J.M. Martinez and S.A. Stantos, A new strategy for solving variational inequalities in bounded polytopes, *Numer. Funct. Anal. And Optimiz.,* **16** (5&6)653–668(1995).

[12] C. Kanzow, Some equation_based methods for the nonlinear complementarity problems, *Optim. Methods Software,* **3** (1994), 327–340.

[13] B. Xiao and P.T. Harker, A nonsmooth Newton method for the variational inequalities, II: numerical results, *Math. Prog.* **65** (1994), 195–216.