JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 129 (2001) 37–62

www.elsevier.nl/locate/cam

Solving variational inequality problems via smoothing-nonsmooth reformulations ☆

Defeng Sun *, Liqun Qi

School of Mathematics, The University of New South Wales, Sydney 2052, Australia

Received 22 February 1999; received in revised form 23 September 1999

Abstract

It has long been known that variational inequality problems can be reformulated as nonsmooth equations. Recently, locally high-order convergent Newton methods for nonsmooth equations have been well established via the concept of semismoothness. When the constraint set of the variational inequality problem is a rectangle, several locally convergent Newton methods for the reformulated nonsmooth equations can also be globalized. In this paper, our main aim is to provide globally and locally high-order convergent Newton methods for solving variational inequality problems with general constraints. To achieve this, we first prove via convolution that these nonsmooth equations can be well approximated by smooth equations, which have desirable properties for the design of Newton methods. We then reformulate the variational inequality problems as equivalent smoothing-nonsmooth equations and apply Newton-type methods to solve the latter systems, and so the variational inequality problems. Stronger convergence results have been obtained. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Variational inequalities; Smoothing; Reformulation

1. Introduction

It has been a long history in mathematical programming field to construct smoothing functions to approximate nonsmooth functions. In this paper we will restrict our study to the smoothing functions of those nonsmooth functions arising from variational inequality problems. The variational inequality problem (VIP for abbreviation) is to find $x^* \in X$ such that

$$(x - x^*)^{\mathrm{T}} F(x^*) \geqslant 0 \quad \text{for all } x \in X, \tag{1.1}$$

☆ The main results of this paper were presented at the International Conference on Nonlinear Optimization and Variational Inequalities, Hong Kong, December 15–18, 1998. This work is supported by the Australian Research Council.

* Corresponding author.

E-mail addresses: sun@maths.unsw.edu.au (D. Sun), l.qi@maths.unsw.edu.au (L. Qi).

PII: S0377-0427(00)00541-0

where X is a nonempty closed convex subset of $\Re^n$ and $F : D \to \Re^n$ is continuously differentiable on some open set D, which contains X. When $X = \Re^n_+$, the VIP reduces to the nonlinear complementarity problem (NCP): Find $x^* \in \Re^n_+$ such that

$$F(x^*) \in \Re^n_+ \quad \text{and} \quad F(x^*)^{\mathrm{T}} x^* = 0. \tag{1.2}$$

It is well known (see, e.g., [5,14]) that solving (1.1) is equivalent to finding a root of the following equation:

$$W(x) := x - \Pi_X[x - F(x)] = 0, \tag{1.3}$$

where for any $x \in \Re^n$, $\Pi_X(x)$ is the Euclidean projection of x onto X. It is also well known that solving the VIP is equivalent to solving the following normal equation:

$$E(y) := F(\Pi_X(y)) + y - \Pi_X(y) = 0, \tag{1.4}$$

in the sense that if $y^* \in \Re^n$ is a solution of (1.4) then $x^* := \Pi_X(y^*)$ is a solution of (1.1), and conversely if x^* is a solution of (1.1) then $y^* := x^* - F(x^*)$ is a solution of (1.4) [34]. Both (1.3) and (1.4) are nonsmooth equations and have led to various generalized Newton's methods under semismoothness assumptions [29,26,23]. The difference between (1.3) and (1.4) is that W is only defined on D and may not have definition outside D if F is not well defined outside D while E is defined on $\Re^n$ even if F is only defined on X.

When X is a rectangle, several globally and locally superlinearly convergent Newton-type methods for solving (1.3) and/or (1.4) are available, see [31] for a review. It is interesting to know whether these results can be generalized to the case that X is not a rectangle. In this paper, we will address this issue and provide a way to solve (1.3) and (1.4) without assuming X to be a rectangle.

From the above discussions, we can see that the nonsmoothness of the reformulated functions inherits from the nonsmoothness of the projection operator Π_X. Due to its key role in our analysis, we will discuss its differentiability properties in details in Section 2. In particular, we show that all the generalized Jacobian of the projection operator Π_X are symmetric. This is actually a key result in analysing the nonsingularity of the generalized Jacobian of W and E and their smoothing counterparts. In Section 3, we discuss some smoothing functions of the projection operator Π_X. An algorithm for solving variational inequality problems and its convergence analysis are presented in Section 4. In particular, we provide some stronger convergence results for monotone variational inequality problems both globally and locally. We give some final remarks in Section 5.

2. Preliminaries

2.1. Generalized Jacobians

Suppose that $H : \Re^n \to \Re^m$ is locally Lipschitz continuous. Then by Rademacher's theorem, H is almost everywhere differentiable. There are many kinds of definitions about the generalized Jacobians of H. Among them, Clarke's definition ∂H [4] is most used in optimization. In this paper we also need other definitions of generalized Jacobians of H. Ioffe [16] and Hiriart-Urruty [15] defined the following so-called support bifunction:

$$H^0(x; u, v) := \limsup_{y \to x, t \downarrow 0} \frac{v^{\mathrm{T}}[H(y + tu) - H(y)]}{t}, \quad x, u \in \Re^n, \; v \in \Re^m. \tag{2.1}$$

For a point-set map $\mathscr{A} : \Re^n \to \Re^{m\times n}$, Sweetser [39] defines

$$\operatorname{plen} \mathscr{A} = \{A \in \Re^{m\times n} \mid Ab \in \mathscr{A}b \ \forall b \in \Re^n\}.$$

In [15], Hiriart-Urruty proved the following useful results.

Theorem 2.1. (i)

$$H^0(x; u, v) = \max_{V\in\partial H(x)} v^{\mathrm{T}} V u.$$

(ii)

$$\mathscr{C}H(x) = \operatorname{plen} \partial H(x),$$

where

$$\mathscr{C}H(x) := \{V \in \Re^{m\times n} \mid v^{\mathrm{T}} V u \leqslant H^0(x; u, v) \ \forall u \in \Re^n, \ v \in \Re^m\}.$$

2.2. Convolution

Let $\Phi : \Re^n \to \Re_+$ be a kernel function, i.e., Φ is integrable (in the sense of Lebesgue) and

$$\int_{\Re^n} \Phi(x)\,\mathrm{d}x = 1.$$

Define $\Theta : \Re_{++} \times \Re^n \to \Re_+$ by

$$\Theta(\varepsilon, x) = \varepsilon^{-n}\Phi(\varepsilon^{-1}x),$$

where $(\varepsilon, x) \in \Re_{++} \times \Re^n$. Then a smoothing approximation of H via convolution can be described by

$$G(\varepsilon, x) := \int_{\Re^n} H(x-y)\Theta(\varepsilon, y)\,\mathrm{d}y = \int_{\Re^n} H(x-\varepsilon y)\Phi(y)\,\mathrm{d}y = \int_{\Re^n} H(y)\Theta(\varepsilon, x-y)\,\mathrm{d}y, \tag{2.2}$$

where $(\varepsilon, x) \in \Re_{++} \times \Re^n$. In order to make (2.2) meaningful, we need some assumptions about the kernel function Φ, which will be addressed in the next section. When H is bounded and uniformly continuous, see [27] for some discussions about G.

For convenience of discussion, we always define

$$G(0, x) = H(x)$$

and for any $\varepsilon < 0$, let

$$G(\varepsilon, x) = G(-\varepsilon, x), \quad x \in \Re^n.$$

2.3. Jacobian characterizations of the projection operator Π_X

The following properties about the projection operator Π_X are well known [42].

Proposition 2.2. *For any $x, y \in \Re^n$,*

(i)
$$\|\Pi_X(y) - \Pi_X(x)\| \leqslant \|y - x\|.$$
(ii)
$$(y - x)^{\mathrm{T}}(\Pi_X(y) - \Pi_X(x)) \geqslant \|\Pi_X(y) - \Pi_X(x)\|^2.$$

Part (i) of Proposition 2.2 says that Π_X is a nonexpansive and globally Lipschitz continuous operator while part (ii) of Proposition 2.2 implies that Π_X is a monotone operator. Then Clarke's generalized Jacobian $\partial\Pi_X$ is well defined everywhere by Rademacher's theorem and for any $x \in \Re^n$, all $S \in \partial\Pi_X(x)$ are positive semidefinite [17]. The next theorem, which gives a partial answer to an open question posed by Hiriart-Urruty [15], summarizes some important properties of $\partial\Pi_X$.

Theorem 2.3. *For $x \in \Re^n$, all $S \in \partial\Pi_X(x)$ are symmetric, positive semidefinite and $\|S\| \leqslant 1$.*

Proof. We only need to prove that all $S \in \partial\Pi_X(x)$ are symmetric by considering of the arguments before this theorem. Define $\phi : \Re^n \to \Re$ by
$$\phi(y) = \tfrac{1}{2}(\|y\|^2 - \|y - \Pi_X(y)\|^2), \quad y \in \Re^n.$$
Then, by [42], ϕ is continuously differentiable with gradient given by
$$\nabla\phi(y) = \Pi_X(y).$$
Thus, if Π_X is differentiable at some point y, we have
$$\Pi_X'(y) = \nabla^2\phi(y),$$
which, according to [22, 3.3.4], proves that $\Pi_X'(y)$ is symmetric. This, by the definition of $\partial\Pi_X$, has in fact proved that all $S \in \partial\Pi_X(x)$ are symmetric. $\square$

We will see subsequently that the symmetric property of all $S \in \partial\Pi_X(x)$ plays an essential role in our analysis.

2.4. Quasi P_0-matrix and quasi P-matrix

A matrix $A \in \Re^{n\times n}$ is a called a P_0-matrix (P-matrix) if every of its principal minors is nonnegative (positive). Here, we will introduce some generalizations of P_0-matrix and P-matrix in order to exploit the properties of the generalized Jacobians of the projection operator Π_X.

Definition 2.4. A matrix $A \in \Re^{n\times n}$ is called a quasi P_0-matrix (P-matrix) if there exists an orthogonal matrix $U \in \Re^{n\times n}$ such that UAU^{T} is a P_0-matrix (P-matrix).

It is obvious that any P_0-matrix (P-matrix) is a quasi P_0-matrix (P-matrix). Any quasi P-matrix is a quasi P_0-matrix and any quasi P-matrix is nonsingular. If A is a quasi P_0-matrix, then for any

$\varepsilon > 0$, $B := A + \varepsilon I$ is a quasi P-matrix, where I is the identity matrix. We will see later that the concepts of quasi P_0-matrix and P-matrix are useful in the analysis of nonsingularity of generalized Jacobians considered in this paper.

Theorem 2.5. *Suppose that $S \in \partial\Pi_X[x - F(x)]$, $x \in \Re^n$. Then there exists an orthogonal matrix $U \in \Re^{n\times n}$ such that $\Sigma := USU^{\mathrm{T}}$ is a diagonal matrix with $0 \leqslant \Sigma_{ii} \leqslant 1$, $i \in \{1,2,\ldots,n\}$. Moreover, if $UF'(x)U^{\mathrm{T}}$ is a P_0-matrix (P-matrix), then $V := I - S(I - F'(x))$ is a quasi P_0-matrix (P-matrix).*

Proof. By Theorem 2.3 we know that there exists an orthogonal matrix U such that $\Sigma := USU^{\mathrm{T}}$ to be a diagonal matrix with $0 \leqslant \Sigma_{ii} \leqslant 1$, $i \in \{1,2,\ldots,n\}$. Then

$$UVU^{\mathrm{T}} = (I - \Sigma) + \Sigma(UF'(x)U^{\mathrm{T}}).$$

Since $0 \leqslant \Sigma_{ii} \leqslant 1$, $i \in \{1,2,\cdots,n\}$ and that $UF'(x)U^{\mathrm{T}}$ is a P_0-matrix (P-matrix), UVU^{T} is a P_0-matrix (P-matrix) as well. This, by Definition 2.4, completes our proof. □

In [7], Facchinei and Pang introduced a concept of the so-called generalized P_0-function. Suppose that X is the Cartesian product of m (with $m \geqslant 1$) lower-dimensional sets:

$$X := \prod_{j=1}^{m} X^j, \tag{2.3}$$

with each X^j being a nonempty closed convex subset of $\Re^{n_j}$ and $\sum_{j=1}^{m} n_j = n$. Correspondingly, suppose that both the variable x and the function $F(x)$ are partitioned in the following way:

$$x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^m \end{pmatrix} \quad \text{and} \quad F(x) = \begin{pmatrix} F^1(x) \\ F^2(x) \\ \vdots \\ F^m(x) \end{pmatrix}, \tag{2.4}$$

where for each j, both x^j and $F^j(x)$ belong to $\Re^{n_j}$. Let $L(X)$ denote all the sets in $\Re^n$ which have the same partitioned structure as X, i.e., $D \in L(X)$ if and only if D can be expressed as

$$D = \prod_{j=1}^{m} D^j$$

with $D^j \in \Re^{n_j}$. Then F is called a generalized P_0-function on $D \in L(X)$ if for every pair of distinct vectors x and y in D, there exists an index j_0 such that

$$x^{j_0} \neq y^{j_0} \quad \text{and} \quad (x^{j_0} - y^{j_0})^{\mathrm{T}}(F^{j_0}(x) - F^{j_0}(x)) \geqslant 0.$$

(Note that this definition is the one given in [7] if $D = X$. The above presentation is more accurate than that in [7]). If F is a generalized P_0-function on any $D \in L(X)$, we say that F is a generalized P_0-function on $L(X)$.

Corollary 2.6. *If F is a continuously differentiable generalized P_0-function on $L(X)$ and $x \in \Re^n$, then for any $S \in \partial\Pi_X[x - F(x)]$, $V := I - S[I - F'(x)]$ is a quasi P_0-matrix.*

Proof. According to Theorem 2.3, for any $S \in \partial\Pi_X[x-F(x)]$, there exist m matrices $S_i \in \partial\Pi_{X^i}[x^i - F^i(x)]$ such that

$$S = \operatorname{diag}(S_1, S_2, \ldots, S_m),$$

and for each $i \in \{1,2,\ldots,m\}$ there exists an orthogonal matrix U_i such that $\Sigma_i := U_i S U_i^{\mathrm{T}}$ is a diagonal matrix with $0 \leqslant (\Sigma_i)_{jj} \leqslant 1$, $j \in \{1,2,\ldots,n_i\}$. Define

$$U := \operatorname{diag}(U_1, U_2, \ldots, U_m) \quad \text{and} \quad \Sigma := \operatorname{diag}(\Sigma_1, \Sigma_2, \ldots, \Sigma_m).$$

Then U is an orthogonal matrix and

$$S = U^{\mathrm{T}} \Sigma U.$$

Since F is a generalized P_0-function on $L(X)$, it is easy to show that for any nonzero vector $h \in \Re^n$, there exists an index $i \in \{1,2,\ldots,m\}$ such that $h^i \neq 0$, $h^i \in \Re^{n_i}$ and $(h^i)^{\mathrm{T}}(F'(x)h)^i \geqslant 0$. Then for any $d \neq 0$, there exists an index $i \in \{1,2,\cdots,m\}$ such that $(U_i)^{\mathrm{T}} d^i \neq 0$, $d^i \in \Re^{n_i}$ and $[(U^i)^{\mathrm{T}} d^i]^{\mathrm{T}} (F'(x)U^{\mathrm{T}} d)^i \geqslant 0$. Therefore,

$$(d^i)^{\mathrm{T}}(Qd)^i \geqslant 0,$$

where $Q := UF'(x)U^{\mathrm{T}}$. Then there exists an index $j(i)$ such that $d^i_{j(i)} \neq 0$ and

$$d^i_{j(i)}(Qd)^i_{j(i)} \neq 0,$$

which implies that Q is a P_0-matrix. Thus, by Theorem 2.5, we get the desired result. □

Corollary 2.7. *If X is a rectangle and for $x \in \Re^n$, $F'(x)$ is a P_0-matrix (P-matrix) or if $F'(x)$ is positive semidefinite (positive definite), then for any $S \in \partial\Pi_X[x-F(x)]$, $V := I - S[I - F'(x)]$ is a quasi P_0-matrix (P-matrix).*

Actually, when X is a rectangle and F is a P_0-function (P-function), Ravindran and Gowda [33] have proved that W and E are P_0-functions (P-functions). See [12] for more discussions about this topic. One might expect that if F is monotone, then W or E must be monotone as well. This is, however, not true, and can be seen clearly by the following example. Consider the NCP with

$$F(x) = Mx, \quad M = \begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix}, \quad x \in \Re^2$$

and $X = \Re^2_+$. Then

$$V := \begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix} \in \partial W(0).$$

Clearly, V is not positive-semidefinite, and so, the function W defined in (1.3) is not monotone even F itself is strongly monotone. It is also noted that if X is not a rectangle and for some $x \in \Re^n$, $F'(x)$ is only a P_0-matrix, $V \in \partial W(x)$ may be not a quasi P_0-matrix, which can be demonstrated by the following example. Consider the VIP with

$$F(x) = Mx, \quad M = \begin{pmatrix} 1 & 0 \\ 20 & 1 \end{pmatrix}, \quad x \in \Re^2$$

and $X = \{x \in \Re^2 \mid x_1 + x_2 = 0\}$. Note that M is a P-matrix, but is not a positive semidefinite matrix. Then the function W defined in (1.3) is continuously differentiable with

$$W'(0,0) = \begin{pmatrix} -9 & 0 \\ 10 & 1 \end{pmatrix}.$$

The matrix $W'(0,0)$ is not a quasi P_0-matrix, because otherwise $T := W'(0,0) + \mathrm{diag}(1,1)$ should be a quasi P-matrix and then $\det(T) > 0$. The latter is not true since $\det(T) = -16 < 0$.

Next, we discuss a result related to level sets of the function W defined in (1.3).

Theorem 2.8. *Suppose that X is of structure* (2.3) *and that F is a continuous generalized P_0-function on $L(X)$. Suppose that $H : \Re^n \to \Re^n$ is continuous, H is partitioned in the way as F in* (2.4), *and there exists a constant $\gamma > 0$ such that for any $j \in \{1,2,\ldots,m\}$ and $y, x \in \Re^n$,*

$$(y^j - x^j)^{\mathrm{T}}(H^j(y) - H^j(x)) \geqslant \gamma \| y^j - x^j \|^2. \tag{2.5}$$

Define $Y : \Re^n \to \Re^n$ by

$$Y(x) := x - \Pi_X[x - (F(x) + H(x))], \quad x \in \Re^n.$$

Then for any $c \geqslant 0$, the following set:

$$L_c := \{x \in \Re^n \mid \|Y(x)\| \leqslant c\}$$

is bounded.

Proof. Suppose by contradiction that L_c is unbounded. Then there exists a sequence $\{x^k\}$ such that $x^k \in L_c$ and $\{x^k\}$ is unbounded. For each k, x^k can be partitioned as

$$x^k = ((x^k)^1, (x^k)^2, \ldots, (x^k)^m)$$

with $(x^k)^j \in \Re^{n_j}$, $j \in \{1,2,\ldots,m\}$. Let

$$I^\infty := \{j \mid (x^k)^j \text{ is unbounded}\} \quad \text{and} \quad I^{\mathrm{b}} := \{j \mid (x^k)^j \text{ is bounded}\}.$$

Then I^∞ is nonempty. Let $a = (a^1, a^2, \ldots, a^m)$ be an arbitrary vector in X with $a^j \in X^j$, $j \in \{1,2,\ldots,m\}$ and define a sequence $\{y^k\}$ by

$$(y^k)^j := \begin{cases} (x^k)^j & \text{if } j \in I^{\mathrm{b}}, \\ a^j + Y^j(x^k) & \text{if } j \in I^\infty. \end{cases}$$

Then $\{y^k\}$ is bounded. Since F is a generalized P_0-function on $L(X)$, there exists an index $j_k \in I^\infty$ (note that $(y^k)^j = (x^k)^j$ for all $j \in I^{\mathrm{b}}$) such that $(y^k)^{j_k} \neq (x^k)^{j_k}$ and

$$[(y^k)^{j_k} - (x^k)^{j_k}]^{\mathrm{T}}(F^{j_k}(y^k) - F^{j_k}(x^k)) \geqslant 0. \tag{2.6}$$

Without loss of generality, assume that $j_k \equiv j_0$. Since

$$(x^k)^{j_0} - Y^{j_0}(x^k) = \Pi_{X^{j_0}}[(x^k)^{j_0} - (F^{j_0}(x^k) + H^{j_0}(x^k))],$$

by (ii) of Proposition 2.1 with $y := a^{j_0}$ and $x := (x^k)^{j_0} - (F^{j_0}(x^k) + H^{j_0}(x^k))$, we have

$$(-Y^{j_0}(x^k) + F^{j_0}(x^k) + H^{j_0}(x^k))^{\mathrm{T}}[a^{j_0} - (x^k)^{j_0} + Y^{j_0}(x^k)] \geqslant 0,$$

which, together with (2.6), implies that

$$(H^{j_0}(x^k))^{\mathrm{T}}[(y^k)^{j_0}-(x^k)^{j_0}]\geqslant(Y^{j_0}(x^k)+F^{j_0}(y^k))^{\mathrm{T}}[(y^k)^{j_0}-(x^k)^{j_0}].$$

This means that there exists a constant $c_1>0$ such that for all k sufficiently large,

$$(H^{j_0}(x^k))^{\mathrm{T}}[(y^k)^{j_0}-(x^k)^{j_0}]\geqslant -c_1\|(y^k)^{j_0}-(x^k)^{j_0}\|.$$

Thus, for all k sufficiently large, we have

$$(H^{j_0}(x^k)-H^{j_0}(y^k))^{\mathrm{T}}[(y^k)^{j_0}-(x^k)^{j_0}]\geqslant -(c_1+\|H^{j_0}(y^k)\|)\|(y^k)^{j_0}-(x^k)^{j_0}\|,$$

which contradicts (2.5) (note that $\{y^k\}$ is bounded). This contradiction shows that L_c is bounded. □

When F itself is strongly monotone, from Theorem 2.8 we have the following result, which appeared in [24] with a lengthy proof.

Corollary 2.9. *Suppose that F is continuous and strongly monotone on $\Re^n$ and that W is defined in* (1.3). *Then for any $c\geqslant 0$, the following set*:

$$\{x\in\Re^n \mid \|W(x)\|\leqslant c\}$$

is bounded.

All the results hold for W in this section can be parallelized to E. Here we omit the details.

3. Smoothing approximations

In this section, unless otherwise stated, we suppose that $H:\Re^n\to\Re^m$ is locally Lipschitz continuous. Suppose that G is defined by (2.2). Define

$$\mathrm{supp}(\Phi)=\{y\in\Re^n \mid \Phi(y)>0\}.$$

The kernel function Φ in supp(Φ) will be omitted if it is obvious from the context.

3.1. Supp(Φ) is bounded

In this section, we suppose that supp(Φ) is bounded. The first smoothing function we want to discuss is the Steklov averaged function. Define

$$\Phi(y)=\begin{cases}1 & \text{if } \max_i|y_i|\leqslant 0.5,\\ 0 & \text{otherwise.}\end{cases} \tag{3.1}$$

Then, for any $\varepsilon>0$, the function $G(\varepsilon,\cdot)$ defined in (2.2), according to [6], is the Steklov averaged function of H with

$$G(\varepsilon,x)=\frac{1}{\varepsilon^n}\int_{x_1-\varepsilon/2}^{x_1+\varepsilon/2}\cdots\int_{x_n-\varepsilon/2}^{x_n+\varepsilon/2}H(y)\,\mathrm{d}y,\quad x\in\Re^n. \tag{3.2}$$

The following result regarding of the continuous differentiability of the Steklov averaged function was first obtained by Gupal [13]. See also Mayne and Polak [21] and Xu and Chang [41].

Proposition 3.1. *Suppose that $H:\Re^n\to\Re^m$ is continuous and that Φ is given by* (3.1). *Then for any $\varepsilon>0$, $G(\varepsilon,\cdot)$ is continuously differentiable on $\Re^n$ and*

$$\begin{aligned}(G_j)'_x(\varepsilon,x) &= \sum_{i=1}^n e_i^{\mathrm T}\frac{1}{\varepsilon^{n-1}}\int_{x_1-\varepsilon/2}^{x_1+\varepsilon/2}\mathrm dy_1\cdots\int_{x_{i-1}-\varepsilon/2}^{x_{i-1}+\varepsilon/2}\mathrm dy_{i-1}\int_{x_{i+1}-\varepsilon/2}^{x_{i+1}+\varepsilon/2}\mathrm dy_{i+1}\cdots\int_{x_n-\varepsilon/2}^{x_n+\varepsilon/2}\mathrm dy_n\\ &\quad\times\frac{1}{\varepsilon}[H_j(y_1,\ldots,y_{i-1},x_i-\varepsilon/2,y_{i+1},\ldots,y_n)-H_j(y_1,\ldots,y_{i-1},x_i+\varepsilon/2,y_{i+1},\ldots,y_n)],\end{aligned}\tag{3.3}$$

where $j=1,2,\ldots,m$ and e_i is the ith unit coordinate vector, $i=1,2,\ldots,n$.

Proposition 3.2 (Sobolev [37], Schwartz [35]). *If* $\operatorname{supp}(\Phi)$ *is bounded and Φ is continuously differentiable, then for any $\varepsilon>0$, $G(\varepsilon,\cdot)$ is continuously differentiable and for $x\in\Re^n$*

$$G'_x(\varepsilon,x)=\int_{\Re^n}H(y)\Theta'_x(\varepsilon,x-y)\,\mathrm dy.\tag{3.4}$$

Recall that a locally Lipschitz continuous function $H:\Re^n\to\Re^m$ is said to be semismooth at $x\in\Re^n$, if

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

exists for any $h\in\Re^n$ [29]. H is said to be strongly semismooth at x if H is semismooth at x and for any $V\in\partial H(x+h)$, $h\to 0$,

$$Vh-H'(x;h)=\mathrm O(\|h\|^2).$$

See [29,26] for details about the discussion of semismoothness and strong semismoothness.

Theorem 3.3. *If Φ is defined by* (3.1) *or if* $\operatorname{supp}(\Phi)$ *is bounded and Φ is continuously differentiable, then we have*

(i) *$G(\cdot,\cdot)$ is continuously differentiable on $\Re_{++}\times\Re^n$.*

(ii) *G is locally Lipschitz continuous on $\Re^{n+1}$. If H itself is globally Lipschitz continuous on $\Re^n$ with Lipschitz constant L, then G is Lipschitz continuous on $\Re^{n+1}$ and for any fixed $\varepsilon>0$ and $x\in\Re^n$,*

$$\|G'_x(\varepsilon,x)\|\leqslant L.\tag{3.5}$$

(iii)

$$\lim_{z\to x,\ \varepsilon\downarrow 0}G'_x(\varepsilon,z)\subseteq\operatorname{plen}\partial H(x).$$

(iv) *For any $x\in\Re^n$,*

$$\pi_x\partial G(0,x)\subseteq\operatorname{plen}\partial H(x)\subseteq\partial H_1(x)\times\partial H_2(x)\times\cdots\times\partial H_m(x),\tag{3.6}$$

where

$$\pi_x \partial G(0,x) := \{A \in \Re^{m\times n} \mid \text{there exists } a \in \Re^m \text{ such that } (a\ A) \in \partial G(0,x)\}.$$

(v) *For any* $\varepsilon > 0$ *and* $x \in \Re^n$, $G'_x(\varepsilon,x)$ *is positive semidefinite (positive definite) if* H *is monotone (strongly monotone) on* $\Re^n$.

(vi) *If* $m = n$ *and there exists a continuously differentiable function* $f : \Re^n \to \Re$ *such that* $H(y) = \nabla f(y)$ *for all* $y \in \Re^n$, *then* $G'_x(\varepsilon,x)$ *is symmetric at any* $(\varepsilon,x) \in \Re_{++} \times \Re^n$.

(vii) *If* H *is semismooth at* x, *then for any* $(\varepsilon,d) \in \Re_{++} \times \Re^n$ with $(\varepsilon,d) \to 0$ *we have*

$$G(\varepsilon,x+d) - G(0,x) - G'(\varepsilon,x+d)\begin{pmatrix}\varepsilon\\ d\end{pmatrix} = \mathrm{o}(\|(\varepsilon,d)\|) \tag{3.7}$$

and

$$G(\varepsilon,x+d) - G(0,x) - G'(\varepsilon,x+d)\begin{pmatrix}\varepsilon\\ d\end{pmatrix} = \mathrm{O}(\|(\varepsilon,d)\|^2) \tag{3.8}$$

if H *is strongly semismooth at* x.

Proof. (i) Similar to the proof of Propositions 3.1 and 3.2, we can see that for any fixed $x \in \Re^n$, $G(\cdot,x)$ is continuously differentiable on $\Re_+$. Moreover, by (3.3) and (3.4), we can see that for any $\varepsilon > 0$ and $x \in \Re^n$ we have

$$\lim_{\tau\to\varepsilon,\, z\to x} G'_x(\tau,z) = G'_x(\varepsilon,x).$$

Then by the definition we can prove that $G'(\varepsilon,x)$ exists for any $(\varepsilon,x) \in \Re_{++} \times \Re^n$. The continuity of G' follows from the continuous differentiability of $G(\cdot,x)$ and $G(\varepsilon,\cdot)$.

(ii) Let (ε,x), (τ,z) be any two points in $\Re^{n+1}$. Then

$$\begin{aligned}\|G(\varepsilon,x) - G(\tau,z)\| &= \left\| \int_{\Re^n} H(x - |\varepsilon| y)\Phi(y)\,\mathrm{d}y - H(z - |\tau| y)\Phi(y)\,\mathrm{d}y \right\| \\ &\leqslant \left\| \int_{\Re^n} \|H(x - |\varepsilon| y) - H(z - |\tau| y) \right\| \Phi(y)\,\mathrm{d}y \\ &= \int_{\mathrm{supp}} \|H(x - |\varepsilon| y) - H(z - |\tau| y)\|\Phi(y)\,\mathrm{d}y \\ &\leqslant \sup_{y\in\mathrm{supp}} \|H(x - |\varepsilon| y) - H(z - |\tau| y)\|,\end{aligned}$$

which, together with the Lipschitz continuity of H, implies the Lipschitz continuity of G. By the above arguments we can see that if H is globally Lipschitz continuous with Lipschitz constant L, then

$$\|G(\varepsilon,y) - G(\varepsilon,z)\| \leqslant L\|y - z\| \quad \forall y,z \in \Re^n,$$

which implies (3.5).

(iii) First, from the proof in [16, Theorem 10.4], for any $v \in \Re^m$ and $u \in \Re^n$, we have

$$v^{\mathrm{T}} G'_x(\varepsilon,z)u = \lim_{t\downarrow 0} \frac{v^{\mathrm{T}}[G(\varepsilon,z+tu) - G(\varepsilon,z)]}{t}$$

$$= \lim_{t\downarrow 0} \int_{\rm supp} \frac{v^{\rm T}[H(z+tu-\varepsilon y)-H(z-\varepsilon y)]}{t} \Phi(y)\,{\rm d}y$$

$$\leqslant \sup_{y\in{\rm supp}} H^0(z-\varepsilon y; u, v).$$

Therefore,

$$v^{\rm T}G_x'(\varepsilon,z)u \leqslant \sup_{y\in{\rm supp}} \max_{V\in\partial H(z-\varepsilon y)} v^{\rm T}Vu.$$

Since ∂H is upper semicontinuous, we have

$$\limsup_{z\to x,\,\varepsilon\downarrow 0} v^{\rm T}G_x'(\varepsilon,z)u \leqslant \max_{V\in\partial H(x)} v^{\rm T}Vu.$$

Hence, by Theorem 2.1, for any

$$A = \lim_{z\to x,\,\varepsilon\downarrow 0} G_x'(\varepsilon,z),$$

we have

$$v^{\rm T}Au \leqslant H^0(x;u,v),$$

which, by Theorem 2.1, implies $A \in {\rm plen}\,\partial H(x)$.

(iv) Define

$$\mathscr{B}(x) := \{B \in \Re^{m\times(n+1)} \mid B = \lim G'(0,y),\, y\to x,\, G'(0,y) \text{ exists}\}.$$

Since, if G is differentiable at any $(0,y) \in \Re \times \Re^n$, then H must be differentiable at y as well and $G_x'(0,y) = H'(y)$, for any $B \in \mathscr{B}(x)$, there exist a vector $a \in \Re^m$ and a matrix $A \in \partial H(x)$ such that

$$(a\ A) = B \in \mathscr{B}(x). \tag{3.9}$$

Then, (iii) of this theorem, (3.9) and the convexity of ${\rm plen}\,\partial H(x)$ imply that

$$\pi_x \partial G(0,x) \subseteq {\rm plen}\,\partial H(x).$$

Since for each i, $\partial H_i(x)$ is compact and convex,

$${\rm plen}\,\partial H(x) \subseteq \partial H_1(x) \times \partial H_2(x) \times \cdots \times \partial H_m(x).$$

So, (3.6) is proved.

(v) First, from our assumptions there exists a nonnegative number α ($\alpha > 0$ if H is strongly monotone) such that

$$(x-z)^{\rm T}[H(x)-H(z)] \geqslant \alpha\|x-z\|^2, \quad \forall x,z \in \Re^n.$$

Then for any $h \in \Re^n$ and any fixed $\varepsilon > 0$, we have

$$h^{\rm T}G_x'(\varepsilon,x)h = \lim_{t\downarrow 0} \frac{h^{\rm T}[G(\varepsilon,x+th)-G(\varepsilon,x)]}{t}$$

$$= \lim_{t\downarrow 0} \int_{\rm supp} \frac{h^{\rm T}[H(x+th-\varepsilon y)-H(x-\varepsilon y)]}{t} \Phi(y)\,{\rm d}y$$

$$\geqslant \lim_{t\downarrow 0}\int_{\rm supp} \alpha h^{\rm T} h \Phi(y)\,{\rm d}y$$

$$= \alpha h^{\rm T} h,$$

which proves (v).

(vi) For any $\varepsilon > 0$, define $\gamma_\varepsilon : \Re^n \to \Re$ by

$$\gamma_\varepsilon(x) = \int_{\Re^n} f(x-\varepsilon y)\Phi(y)\,{\rm d}y, \quad x\in\Re^n.$$

Then, by (i) of this theorem, γ_ε is continuously differentiable. By direct computations, we have

$$\begin{aligned}\nabla\gamma_\varepsilon(x) &= \int_{\Re^n} \nabla f(x-\varepsilon y)\Phi(y)\,{\rm d}y\\ &= \int_{\Re^n} H(x-\varepsilon y)\Phi(y)\,{\rm d}y\\ &= G(\varepsilon,x), \quad x\in\Re^n.\end{aligned}$$

This, together with (i), means that γ_ε is twice continuously differentiable on $\Re^n$ and

$$G'_x(\varepsilon,x) = \nabla^2\gamma_\varepsilon(x), \quad x\in\Re^n,$$

which means that $G'_x(\varepsilon,x)$ is symmetric by [22, 3.3.4].

(vii) According to the definition, for any $(\varepsilon,d)\in\Re_{++}\times\Re^n$ with $(\varepsilon,d)\to 0$ we have

$$\begin{aligned}&G(\varepsilon,x+d) - G(0,x) - G'(\varepsilon,x+d)\begin{pmatrix}\varepsilon\\ d\end{pmatrix}\\ &\quad= \lim_{t\downarrow 0}\int_{\rm supp}\left[H(x+d-\varepsilon y) - H(x) - \frac{H(x+d+td-(\varepsilon+t\varepsilon)y) - H(x+d-\varepsilon y)}{t}\right]\Phi(y)\,{\rm d}y\\ &\quad= \lim_{t\downarrow 0}\int_{\rm supp}[H(x+d-\varepsilon y) - H(x) - V_t(d-\varepsilon y)]\Phi(y)\,{\rm d}y,\end{aligned}$$

where the last equality follows from the mean value theorem [4, Proposition 2.6.5] and V_t is an element of the convex hull of

$$\partial H([x+d-\varepsilon y, x+d-\varepsilon y + t(d-\varepsilon y)]).$$

By Carathéodory theorem, there exists at most $mn+1$ elements

$$W_{t,i}\in\partial H(x+d-\varepsilon y + t\theta_i(d-\varepsilon))$$

such that

$$V_t = \sum_{i=1}^{mn+1}\lambda_i W_{t,i},$$

where $\theta_i,\ \lambda_i \in [0,1]$ and $\sum_{i=1}^{mn+1} \lambda_i = 1$. Then

$$G(\varepsilon,x+d) - G(0,x) - G_x'(\varepsilon,x+d)\begin{pmatrix}\varepsilon\\ d\end{pmatrix}$$

$$= \lim_{t\downarrow 0} \int_{\text{supp}} \sum_{i=1}^{mn+1} \lambda_i [H(x+d-\varepsilon y) - H(x) - W_{t,i}(d-\varepsilon y)]\Phi(y)\,\mathrm{d}y$$

$$= \lim_{t\downarrow 0} \int_{\text{supp}} \sum_{i=1}^{mn+1} \lambda_i [H(x+d-\varepsilon y + t\theta_i(d-\varepsilon y)) - H(x)$$

$$- W_{t,i}(d-\varepsilon y + t\theta_i(d-\varepsilon y)) + \mathrm{O}(t)]\Phi(y)\,\mathrm{d}y.$$

Then, if H is semismooth at x, we have

$$\left\| G(\varepsilon,x+d) - G(\varepsilon,x) - G_x'(\varepsilon,x+d)\begin{pmatrix}\varepsilon\\ d\end{pmatrix}\right\| = \sup_{y\in\text{supp}} \mathrm{o}(\|d-\varepsilon y\|) = \mathrm{o}(\|(\varepsilon,d)\|).$$

This proves (3.7). If H is strongly semismooth at x, by following the above arguments, we can prove (3.8). □

In proving part (iii) of Theorem 3.3, we used an idea from Ioffe [16]. For the Steklov averaged function, Xu and Chang [41] proved a similar result to part (iii) of Theorem 3.3. Recently, by assuming that $H : \Re^n \to \Re^n$ is globally Lipschitz continuous, Xu [40] applied this result for the Steklov-averaged function of H to modify an algorithm proposed in [3].

3.2. Supp(Φ) is infinite

Assumption 3.4. (i) H is globally Lipschitz continuous with Lipschitz constant L.

(ii)

$$\kappa := \int_{\Re^n} \|y\|\Phi(y)\,\mathrm{d}y < \infty.$$

(iii) Φ is continuously differentiable and for any $\varepsilon > 0$ and $x \in \Re^n$, the following integral:

$$\int_{\Re^n} H(y)\Theta_x'(\varepsilon,x-y)\,\mathrm{d}y$$

exists.

(iv) For any $\varepsilon > 0$, $x \in \Re^n$ and $h \to 0$ it holds that

$$\sup_{t\in[0,1]} \int_{\Re^n} \|y\| \|[\Theta_x'(\varepsilon,x+th-y) - \Theta_x'(\varepsilon,x-y)]h\|\,\mathrm{d}y = \mathrm{o}(\|h\|).$$

Proposition 3.5. *Suppose that Assumption* 3.4 *holds. Then for any* $\varepsilon > 0$, $G(\varepsilon,\cdot)$ *is continuously differentiable with Jacobian given by*

$$G_x'(\varepsilon,x) = \int_{\Re^n} H(y)\Theta_x'(\varepsilon,x-y)\,\mathrm{d}y, \quad x \in \Re^n. \tag{3.10}$$

Proof. By (i) and (ii) of Assumption 3.4, for any $(\varepsilon, x) \in \Re_{++} \times \Re^n$, G is well defined. By (iii) and (iv) of Assumption 3.4, for any $h \in \Re^n$ with $h \to 0$ we have

$$\begin{aligned}
&\left\| G(\varepsilon, x+h) - G(\varepsilon, x) - \int_{\Re^n} H(y)\Theta'_x(\varepsilon, x-y)h \, \mathrm{d}y \right\| \\
&\quad = \left\| \int_{\Re^n} H(y)[\Theta(\varepsilon, x+h-y) - \Theta(\varepsilon, x-y) - \Theta'_x(\varepsilon, x-y)h] \, \mathrm{d}y \right\| \\
&\quad \leqslant \int_{\Re^n} \|H(y)\| \max_{t\in[0,1]} \|[\Theta'_x(\varepsilon, x+th-y) - \Theta'_x(\varepsilon, x-y)]h\| \, \mathrm{d}y \\
&\quad \leqslant L \sup_{t\in[0,1]} \int_{\Re^n} \|y\| \|[\Theta'_x(\varepsilon, x+th-y) - \Theta'_x(\varepsilon, x-y)]h\| \, \mathrm{d}y \\
&\qquad + (\|H(y_0)\| + L\|y_0\|) \sup_{t\in[0,1]} \int_{\Re^n} \|[\Theta'_x(\varepsilon, x+th-y) - \Theta'_x(\varepsilon, x-y)]h\| \, \mathrm{d}y \\
&\quad = \mathrm{o}(\|h\|),
\end{aligned}$$

where $y_0 \in \Re^n$ is an arbitrary point. Then we have proved that $G'_x(\varepsilon, x)$ exists and (3.10) holds. The continuity of $G'(\varepsilon, \cdot)$ follows from (iv) of Assumption 3.4. $\square$

Assumption 3.6. (i) For any $\varepsilon > 0$ and $x \in \Re^n$, the following integral:

$$\int_{\Re^n} H(y)\Theta'_\varepsilon(\varepsilon, x-y) \, \mathrm{d}y \tag{3.11}$$

exists.

(ii) For any $\varepsilon > 0$, $x \in \Re^n$ and $\tau \in \Re$ with $\tau \to 0$ we have

$$\sup_{t\in[0,1]} \int_{\Re^n} \|y\| |[\Theta'_\varepsilon(\varepsilon + t\tau,\ x-y) - \Theta'_\varepsilon(\varepsilon + \tau,\ x-y)]\tau| \, \mathrm{d}y = \mathrm{o}(\tau). \tag{3.12}$$

(iii) For any $\varepsilon > 0$ and $x \in \Re^n$ we have

$$\lim_{\tau\to\varepsilon,\, z\to x} \int_{\Re^n} \|y\| |\Theta'_\varepsilon(\tau, z-y) - \Theta'_\varepsilon(\varepsilon, x-y)| \, \mathrm{d}y = 0. \tag{3.13}$$

If H is globally Lipschitz continuous, there are plenty of kernel functions satisfying Assumptions 3.4 and 3.6, e.g., the Weierstrass kernel function

$$\Phi(y) = \frac{1}{(\sqrt{\pi})^n} \mathrm{e}^{-\|y\|^2}.$$

Analogously to Theorem 3.3 we have the following theorem.

Theorem 3.7. *Suppose that* $\operatorname{supp}(\Phi)$ *is infinite and* Φ *is continuously differentiable. Suppose that Assumption* 3.4 *holds. Then we have*

(i) *If Assumption* 3.6 *holds, then* $G(\cdot, \cdot)$ *is continuously differentiable on* $\Re_{++} \times \Re^n$.

(ii) *G is globally Lipschitz continuous on* $\Re^{n+1}$ *and for any fixed* $\varepsilon > 0$ *and* $x \in \Re^n$,

$$\|G'_x(\varepsilon, x)\| \leqslant L. \tag{3.14}$$

(iii)

$$\lim_{z \to x, \varepsilon \downarrow 0} G'_x(\varepsilon, z) \subseteq \text{plen}\, \partial H(x).$$

(iv) *For any* $x \in \Re^n$,

$$\pi_x \partial G(0, x) \subseteq \text{plen}\, \partial H(x) \subseteq \partial H_1(x) \times \partial H_2(x) \times \cdots \times \partial H_m(x). \tag{3.15}$$

(v) *For any* $\varepsilon > 0$ *and* $x \in \Re^n$, $G'_x(\varepsilon, x)$ *is positive semidefinite* (*positive definite*) *if* H *is monotone* (*strongly monotone*) *on* $\Re^n$.

(vi) *Suppose that* $m = n$. *If there exists a continuously differentiable function* $f : \Re^n \to \Re$ *such that* $H(y) = \nabla f(y)$ *for all* $y \in \Re^n$ *and for all* $\varepsilon > 0$ *and* $x \in \Re^n$,

$$\gamma_\varepsilon(x) := \int_{\Re^n} f(x - \varepsilon y)\Phi(y)\, \mathrm{d}y$$

is well defined, then $G'_x(\varepsilon, x)$ *is symmetric at any* $(\varepsilon, x) \in \Re_{++} \times \Re^n$.

(vii) *If* H *is semismooth at* x, *then for any* $(\varepsilon, d) \in \Re_{++} \times \Re^n$ *with* $(\varepsilon, d) \to 0$ *we have*

$$G(\varepsilon, x + d) - G(0, x) - G'(\varepsilon, x + d)\begin{pmatrix} \varepsilon \\ d \end{pmatrix} = \mathrm{o}(\|(\varepsilon, d)\|). \tag{3.16}$$

Proof. (i) From parts (i) and (ii) of Assumption 3.6 we can see that for any $(\varepsilon, x) \in \Re_{++} \times \Re^n$, $G'_\varepsilon(\varepsilon, x)$ exists and $G'_\varepsilon(\cdot, x)$ is continuous. Part (iii) of Assumption 3.6 guarantees that

$$\lim_{\tau \to \varepsilon, z \to x} G'_\varepsilon(\tau, z) = G'_\varepsilon(\varepsilon, x),$$

which, together with the continuity of $G'_\varepsilon(\cdot, x)$ and $G'_x(\varepsilon, \cdot)$ (Proposition 3.5), implies that G is continuously differentiable on $\Re_{++} \times \Re^n$.

(ii) Let (ε, x), (τ, z) be any two points in $\Re^{n+1}$. Then

$$\begin{aligned}
&\|G(\varepsilon, x) - G(\tau, z)\| \\
&= \left\| \int_{\Re^n} H(x - |\varepsilon| y)\Phi(y)\, \mathrm{d}y - H(z - |\tau| y)\Phi(y)\, \mathrm{d}y \right\| \\
&\leqslant \int_{\Re^n} \|H(x - |\varepsilon| y) - H(z - |\tau| y)\| \Phi(y)\, \mathrm{d}y \\
&= L \int_{\Re^n} \|(x - |\varepsilon| y) - (z - |\tau| y)\| \Phi(y)\, \mathrm{d}y \\
&\leqslant L[\|x - z\| + |\varepsilon - \tau| \kappa] \\
&\leqslant L\sqrt{2} \max\{1, \kappa\} \|(\varepsilon, x) - (\tau, z)\|,
\end{aligned}$$

which proves the global Lipschitz continuity of G. By the above arguments, we also have

$$\|G(\varepsilon, y) - G(\varepsilon, z)\| \leqslant L\|y - z\| \quad \forall y,\ z \in \Re^n.$$

This implies (3.14).

(iii) For any given $\delta > 0$, there exists a number $M > 0$ such that

$$\int_{\|y\| \geqslant M} \Phi(y)\,\mathrm{d}y \leqslant \frac{\delta}{2}.$$

Let

$$D_M := \{ y \in \Re^n \mid \|y\| \leqslant M \}.$$

Suppose that $v \in \Re^m$ and $u \in \Re^n$ are two arbitrarily chosen points. Since ∂H is upper semicontinuous, there exists a number $\tau > 0$ such that for all $\varepsilon \in [0, \tau]$ and all z with $\|z - x\| \leqslant \tau$ we have

$$\sup_{y \in D_M} \max_{V \in \partial H(z - \varepsilon y)} v^{\mathrm{T}} V u \leqslant \max_{V \in \partial H(x)} v^{\mathrm{T}} V u + \frac{\delta}{2} \|v\| \|u\|. \tag{3.17}$$

Since H is globally Lipschitz continuous, by the definition, we have

$$\begin{aligned} v^{\mathrm{T}} G_x'(\varepsilon, z) u &= \lim_{t \downarrow 0} \frac{v^{\mathrm{T}} [G(\varepsilon, z + tu) - G(\varepsilon, z)]}{t} \\ &= \lim_{t \downarrow 0} \left\{ \int_{D_M} \frac{v^{\mathrm{T}} [H(z + tu - \varepsilon y) - H(z - \varepsilon y)]}{t} \Phi(y)\,\mathrm{d}y \right. \\ &\qquad \left. + \int_{\|y\| \geqslant M} \frac{v^{\mathrm{T}} [H(z + tu - \varepsilon y) - H(z - \varepsilon y)]}{t} \Phi(y)\,\mathrm{d}y \right\} \\ &\leqslant \sup_{y \in D_M} F^0(z - \varepsilon y; u, v) + L \|v\| \|u\| \frac{\delta}{2}, \end{aligned}$$

which, together with (3.17), implies that for any $\delta > 0$, there exists a number $\tau > 0$ such that for all $\varepsilon \in [0, \tau]$ and all z with $\|z - x\| \leqslant \tau$ we have

$$v^{\mathrm{T}} G_x'(\varepsilon, z) u \leqslant \sup_{y \in D_M} \max_{V \in \partial H(z - \varepsilon y)} v^{\mathrm{T}} V u + L \|v\| \|u\| \frac{\delta}{2} \leqslant \max_{V \in \partial H(x)} v^{\mathrm{T}} V u + (L + 1) \|v\| \|u\| \frac{\delta}{2}.$$

This means that

$$\limsup_{z \to x, \varepsilon \downarrow 0} v^{\mathrm{T}} G_x'(\varepsilon, z) u \leqslant \max_{V \in \partial H(x)} v^{\mathrm{T}} V u.$$

Hence, by Theorem 2.1, for any

$$A = \lim_{z \to x, \varepsilon \downarrow 0} G_x'(\varepsilon, z),$$

we have

$$v^{\mathrm{T}} A u \leqslant F^0(x; u, v),$$

which, by Theorem 2.1, implies $A \in \operatorname{plen} \partial H(x)$.

(iv) Similar to the proof of (iv) in Theorem 3.3.

(v) Similar to the proof of (v) in Theorem 3.3.

(vi) By the definition,

$$\gamma_\varepsilon(x+h)-\gamma_\varepsilon(x)-\int_{\Re^n} H(x-\varepsilon y)^{\mathrm{T}} h\Phi(y)\,\mathrm{d}y$$

$$=\int_{\Re^n}[f(x+h-\varepsilon y)-f(x-\varepsilon y)-f'(x-\varepsilon y)h]\Phi(y)\,\mathrm{d}y$$

$$=\int_{\Re^n}\int_0^1[f'(x+th-\varepsilon y)-f'(x-\varepsilon y)]h\,\mathrm{d}t\Phi(y)\,\mathrm{d}y.$$

Since $f'=H$ is globally Lipschitz continuous, for any $\delta>0$ there exists a number $M>0$ such that

$$\int_{\|y\|\geqslant M}\int_0^1\|[f'(x+th-\varepsilon y)-f'(x-\varepsilon y)]h\|\,\mathrm{d}t\Phi(y)\,\mathrm{d}y\leqslant\frac{\delta}{2}\|h\|^2.$$

For the given M, there exists a number $\tau>0$ such that for all $\varepsilon\in[0,\tau]$ and all h with $\|h\|\leqslant\tau$ we have

$$\int_{\|y\|\leqslant M}\int_0^1\|[f'(x+th-\varepsilon y)-f'(x-\varepsilon y)]h\|\,\mathrm{d}t\Phi(y)\,\mathrm{d}y\leqslant\frac{\delta}{2}\|h\|$$

because f' is uniformly continuous on any bounded set. Therefore, for any $\delta>0$ there exists a number $\tau\in(0,1]$ such that for all $\varepsilon\in[0,\tau]$ and all h with $\|h\|\leqslant\tau$ we have

$$\int_{\Re^n}\int_0^1\|[f'(x+th-\varepsilon y)-f'(x-\varepsilon y)]h\|\,\mathrm{d}t\Phi(y)\,\mathrm{d}y\leqslant\delta\|h\|.$$

This implies that

$$\gamma_\varepsilon(x+h)-\gamma_\varepsilon(x)-\int_{\Re^n} H(x-\varepsilon y)^{\mathrm{T}} h\Phi(y)\,\mathrm{d}y=\mathrm{o}(\|h\|),$$

which means that γ_ε is differentiable and

$$\nabla\gamma_\varepsilon(x)=\int_{\Re^n} H(x-\varepsilon y)\Phi(y)\,\mathrm{d}y=G(\varepsilon,x).$$

Since $G(\varepsilon,\cdot)$ is continuous, γ_ε is continuously differentiable. By Proposition 3.5, $G(\varepsilon,\cdot)$ is continuously differentiable, and so $\gamma_\varepsilon(\cdot)$ is twice continuously differentiable with

$$\nabla^2\gamma_\varepsilon(x)=G'_x(\varepsilon,x),\quad x\in\Re^n.$$

This, together with the symmetric property of $\nabla^2\gamma_\varepsilon(x)$, implies that $G'_x(\varepsilon,x)$ is symmetric.

(vii) First, for any given $\delta>0$, under the assumptions, we know that there exists a number $M>0$ such that

$$\max\left\{\int_{\|y\|\geqslant M}\Phi(y)\,\mathrm{d}y,\int_{\|y\|\geqslant M}\|y\|\Phi(y)\,\mathrm{d}y\right\}\leqslant\frac{\delta}{4}.$$

Then for any $d\in\Re^n$ and $\varepsilon>0$, we have

$$\left\|\int_{\|y\|\geqslant M}\left[H(x+d-\varepsilon y)-H(x)-\frac{H(x+d+td-(\varepsilon+t\varepsilon)y)-H(x+d-\varepsilon y)}{t}\right]\Phi(y)\,\mathrm{d}y\right\|$$

$$\leqslant(L\|d\|+\varepsilon)\frac{\delta}{2}.$$

Next, according to the proof in part (vii) of Theorem 3.3, we know that there exists a number $\tau > 0$ such that for all $(\varepsilon, d) \in \Re^{n+1}$ with $\varepsilon > 0$ and $\|(\varepsilon, d)\| \leqslant \tau$ we have

$$\limsup_{t\downarrow 0} \left\| \int_{\|y\|\leqslant M} \left[H(x+d-\varepsilon y) - H(x) - \frac{H(x+d+td-(\varepsilon+t\varepsilon)y) - H(x+d-\varepsilon y)}{t} \right] \Phi(y)\,\mathrm{d}y \right\| \leqslant \frac{\delta}{2}\|(\varepsilon, d)\|.$$

Therefore, for any given $\delta > 0$, there exists a number $\tau > 0$ such that for all $(\varepsilon, d) \in \Re^{n+1}$ with $\varepsilon > 0$ and $\|(\varepsilon, d)\| \leqslant \tau$ we have

$$\begin{aligned}
&\left\| G(\varepsilon, x+d) - G(0,x) - G'_x(\varepsilon, x+d)\begin{pmatrix}\varepsilon\\ d\end{pmatrix} \right\| \\
&= \left\| \lim_{t\downarrow 0} \int_{\Re^n} \left[H(x+d-\varepsilon y) - H(x) - \frac{H(x+d+td-(\varepsilon+t\varepsilon)y) - H(x+d-\varepsilon y)}{t} \right] \Phi(y)\,\mathrm{d}y \right\| \\
&\leqslant \limsup_{t\downarrow 0} \left\| \int_{\|y\|\leqslant M} \left[H(x+d-\varepsilon y) - H(x) - \frac{H(x+d+td-(\varepsilon+t\varepsilon)y) - H(x+d-\varepsilon y)}{t} \right] \Phi(y)\,\mathrm{d}y \right\| \\
&\quad + \limsup_{t\downarrow 0} \left\| \int_{\|y\|\geqslant M} \left[H(x+d-\varepsilon y) - H(x) - \frac{H(x+d+td-(\varepsilon+t\varepsilon)y) - H(x+d-\varepsilon y)}{t} \right] \Phi(y)\,\mathrm{d}y \right\| \\
&\leqslant \frac{\delta}{2}\|(\varepsilon, d)\| + (L+1)\frac{\delta}{2}\|(\varepsilon, d)\| \\
&\leqslant (L+1)\delta\|(\varepsilon, d)\|,
\end{aligned}$$

which proves (3.16). □

Contrary to Theorem 3.3, in Theorem 3.7 we could not prove a result similar to (3.8). When X is a rectangle, such a result holds for several smoothing functions with supp(Φ) being infinite [32].

3.3. Smoothing-nonsmooth reformulations

If in (2.2), H is replaced by Π_X, a smoothing approximation of Π_X via convolution can be described by

$$G(\varepsilon,x) = \int_{\Re^n} \Pi_X(x-y)\Theta(\varepsilon,y)\,\mathrm{d}y, \tag{3.18}$$

where $(\varepsilon,x) \in \Re_{++} \times \Re^n$. Since $\int_{\Re^n} \Theta(\varepsilon,y)\,\mathrm{d}y = 1$, for any $\varepsilon > 0$ and $x \in \Re^n$, we have

$$G(\varepsilon,x) \in X.$$

As we stated early, we always define

$$G(0,x) := \Pi_X(x) \quad \text{and} \quad G(-|\varepsilon|,x) := G(|\varepsilon|,x), \quad x \in \Re^n.$$

Suppose that Φ is chosen such that G is continuous on $\Re^{n+1}$ and is continuously differentiable on $\Re_{++} \times \Re^n$. Then to solve the VIP is equivalent to solve

$$Q(\varepsilon,x) := \begin{pmatrix} \varepsilon \\ P(\varepsilon,x) \end{pmatrix} = 0, \tag{3.19}$$

where

$$P(\varepsilon,x) := x - G(\varepsilon, x - F(x)).$$

Since Q is continuously differentiable on $\Re_{++} \times \Re^n$ and $\Re_{--} \times \Re^n$ and may be nonsmooth on $0 \times \Re^n$, we can see (3.19) as a smoothing-nonsmooth reformulation of the VIP. Another smoothing-nonsmooth reformulation of the VIP is

$$R(\varepsilon,x) := \begin{pmatrix} \varepsilon \\ S(\varepsilon,x) \end{pmatrix} = 0, \tag{3.20}$$

where

$$S(\varepsilon,x) := F(G(\varepsilon,x)) + x - G(\varepsilon,x).$$

Theorem 3.8. *Suppose that Φ is defined by* (3.1) *or suppose that* $\operatorname{supp}(\Phi)$ *is bounded and Φ is continuously differentiable. Then Q is continuously differentiable on $\Re_{++} \times \Re^n$, and for any $\varepsilon > 0$ there exists an orthogonal matrix U such that $\sum := UG'_x(\varepsilon, x - F(x))U^{\mathrm{T}}$ is a diagonal matrix with $0 \leqslant \sum_{ii} \leqslant 1, i \in \{1,2,\ldots,n\}$. Moreover, we have*
(i) *if $UF'(x)U^{\mathrm{T}}$ is a P_0-matrix (P-matrix), then $P'_x(\varepsilon,x)$ is a quasi P_0-matrix (P-matrix);*
(ii) *if Π_X is semismooth at $x - F(x)$, then for any $(\varepsilon,d) \in \Re_{++} \times \Re^n$ with $(\varepsilon,d) \to 0$ we have*

$$Q(\varepsilon,x+d) - Q(0,x) - Q'(\varepsilon,x+d)\begin{pmatrix} \varepsilon \\ d \end{pmatrix} = \mathrm{o}(\|(\varepsilon,d)\|) \tag{3.21}$$

and

$$Q(\varepsilon,x+d) - G(0,x) - G'(\varepsilon,x+d)\begin{pmatrix} \varepsilon \\ d \end{pmatrix} = \mathrm{O}(\|(\varepsilon,d)\|^2) \tag{3.22}$$

if Π_X is strongly semismooth at $x - F(x)$ and F' is Lipschitz continuous at x.

Proof. First, from Theorem 3.3, P is continuously differentiable on $\Re_{++} \times \Re^n$. Then, since Π_X is monotone and globally Lipschitz continuous with Lipschitz constant 1, from Theorem 3.3 and the proof of Theorem 2.3, we know that for any $\varepsilon > 0$, $G_x'(\varepsilon,x)$ is symmetric, positive semidefinite and $\|G_x'(\varepsilon,x)\| \leqslant 1$. Thus, there exists an orthogonal matrix U such that $\sum := UG_x'(\varepsilon,x - F(x))U^{\mathrm{T}}$ is a diagonal matrix with $0 \leqslant \sum_{ii} \leqslant 1$, $i \in \{1,2,\dots,n\}$.

Part (i) can be proved similarly to that in Theorem 2.5 and part (ii) follows from (vii) in Theorem 3.3. □

Theorem 3.9. *Suppose that* supp(Φ) *is infinite and* Φ *is continuously differentiable. Suppose that Assumptions* 3.4 and 3.6 *hold and that*

$$\int_{\Re^n} \|y\|^2 \Phi(y)\,\mathrm{d}y < \infty.$$

Then Q *is continuously differentiable on* $\Re_{++} \times \Re^n$, and for any $\varepsilon > 0$ *there exists an orthogonal matrix* U *such that* $\sum := UG_x'(\varepsilon,x-F(x))U^{\mathrm{T}}$ *is a diagonal matrix with* $0 \leqslant \sum_{ii} \leqslant 1, i \in \{1,2,\dots,n\}$. *Moreover, we have*

(i) *if* $UF'(x)U^{\mathrm{T}}$ is a P_0-matrix (P-matrix), then $P_x'(\varepsilon,x)$ is a quasi P_0-matrix (P-matrix);

(ii) *if* Π_X *is semismooth at* $x - F(x)$, *then for any* $(\varepsilon,d) \in \Re_{++} \times \Re^n$ *with* $(\varepsilon,d) \to 0$ *we have*

$$Q(\varepsilon,x+d) - G(0,x) - Q'(\varepsilon,x+d)\begin{pmatrix}\varepsilon\\ d\end{pmatrix} = \mathrm{o}(\|(\varepsilon,d)\|). \tag{3.23}$$

Proof. Similar to the proof of Theorem 3.8, we can prove this theorem by Theorem 3.7. We omit the details. □

In this subsection, the results hold for P hold for R as well. We omit the details here.

3.4. Computable smoothing functions of Π_X

The smoothing approximation function G via convolution is not "computable" if X is not a rectangle since in this case a multivariate integral is involved. When X is a rectangle, see [2,9] for various concrete forms of G. In [30], the authors discussed a way to get an approximate smoothing function of $\Pi_X(\cdot)$ when X is explicitly expressed as

$$X := \{x \in \Re^n \mid g_i(x) \leqslant 0, i = 1,2,\dots,m\}, \tag{3.24}$$

where for each i, g_i is a twice continuously differentiable convex function. Suppose that the Slater constraint qualification holds, i.e., there exists a point x^0 such that $g_i(x^0) < 0$ for all $i \in \{1,2,\cdots,m\}$. Then for any $x \in \Re^n$, there exists a vector $\lambda \in \Re^m_+$ such that

$$\begin{aligned} &y - x + \sum_{i=1}^m \lambda_i \nabla g_i(y) = 0,\\ &\lambda - \max(0,\lambda + g(y)) = 0. \end{aligned} \tag{3.25}$$

Suppose that $a(\varepsilon, t)$ is the Chen–Harker–Kanzow–Smale (CHKS) smoothing function [1,18,36] of $\max(0, t)$, $t \in \Re$, which is given by

$$a(\varepsilon, t) := \frac{\sqrt{4\varepsilon^2 + t^2} + t}{2}, \quad t \in \Re.$$

(We can use other smoothing functions of $\max(0, t)$, $t \in \Re$. Here, we choose the CHKS smoothing function for the ease of discussion.) Define $A : \Re \times \Re^m \to \Re^m$ by

$$A_i(\varepsilon, z) := a(\varepsilon, z_i), \quad i \in \{1, 2, \dots, m\}.$$

Consider the perturbed system of (3.25)

$$D((y, \lambda), (\varepsilon, x)) := \begin{pmatrix} y - x + \sum_{i=1}^{m} \lambda_i \nabla g_i(y) \\ \lambda - A(\varepsilon, \lambda + g(y)) \end{pmatrix} = 0, \tag{3.26}$$

where ε and x are treated as parameters and y and λ as variables. For any $(\varepsilon, x) \in \Re_{++} \times \Re^n$, the system (3.26) has a unique solution $(y(\varepsilon, x), \lambda(\varepsilon, x))$.

Proposition 3.10 (Qi and Sun [30]). *Suppose that the Slater constraint qualification holds. Then*
(i) *$y(\cdot)$ is continuously differentiable on $\Re_{++} \times \Re^n$ and for any $(\varepsilon, x) \in \Re_{++} \times \Re^n$, $y'_x(\varepsilon, x)$ is symmetric, positive semidefinite and*

$$\|y'_x(\varepsilon, x)\| \leqslant 1.$$

(ii) For any $x^0 \in \Re^n$,

$$\lim_{\varepsilon \downarrow 0,\, x \to x^0} y(\varepsilon, x) = \Pi_X(x^0).$$

See [30] for some important differentiability properties of the above-defined smoothing function. Note that for any $\varepsilon > 0$ and $x \in \Re^n$, to compute $y(\varepsilon, x)$ is no more difficult than to compute $\Pi_X(x)$ since $D((\cdot,\cdot), (\varepsilon, x))$ is continuously differentiable everywhere when $\varepsilon \neq 0$.

Define

$$y(0, x) := \Pi_X(x) \quad \text{and} \quad y(-|\varepsilon|, x) := y(|\varepsilon|, x), \quad (\varepsilon, x) \in \Re \times \Re^n.$$

Then we can define a smoothing function P of W as

$$P(\varepsilon, x) := x - y(\varepsilon, x - F(x)),$$

where $(\varepsilon, x) \in \Re \times \Re^n$. By considering the above proposition, P is continuously differentiable everywhere except $(0, x)$, $x \in \Re^n$ and P is continuous at $(0, x)$, $x \in \Re^n$. Similarly, we can define a smoothing function of E as

$$S(\varepsilon, x) := F(y(\varepsilon, x)) + x - y(\varepsilon, x), \quad (\varepsilon, x) \in \Re \times \Re^n.$$

It is noted that for any $\varepsilon \neq 0$, $y(\varepsilon, x) \in \operatorname{int} X$. This guarantees that S is well defined even if F is only defined on X.

4. Algorithm and its convergence

Suppose that $Q : \Re^{n+1} \to \Re^{n+1}$ is defined by (3.19). In this section we will discuss a method to find a solution of $Q(z) = 0$, where $z := (\varepsilon, x) \in \Re \times \Re^n$. If for the VIP, F is not well defined outside X, we use R defined by (3.20) to replace Q.

Assumption 4.1. (i) Q is continuous on $\Re^{n+1}$ and is continuously differentiable on $\Re_{++} \times \Re^n$.
(ii) For any $\varepsilon > 0$ and $x \in \Re^n$, $Q'(\varepsilon, x)$ is nonsingular.

Remark. According to Theorems 3.8 and 3.9 and Corollary 2.6, we know that if F is a continuously differentiable generalized P_0-function on $L(X)$, then for any $\varepsilon > 0$ and $x \in \Re^n$, $P'_x(\varepsilon, x)$ is a quasi P_0-matrix. This implies that if we redefine

$$P(\varepsilon, x) := P(\varepsilon, x) + \varepsilon x$$

or

$$P(\varepsilon, x) := x - G(\varepsilon, x - (F(x) + \varepsilon x)),$$

then $P'_x(\varepsilon, x)$ becomes a quasi P-matrix, and thus, a nonsingular matrix. Therefore, if F is a generalized P_0-function on $L(X)$, then, by redefining P if necessary, part (ii) in Assumption 4.1 always holds.

Choose $\bar{\varepsilon} \in \Re_{++}$ and $\gamma \in (0, 1)$ such that $\gamma\bar{\varepsilon} < 1$. Let $\bar{z} := (\bar{\varepsilon}, 0) \in \Re \times \Re^n$. Define the merit function $\psi : \Re^{n+1} \to \Re_+$ by

$$\psi(z) := \|Q(z)\|^2$$

and define $\beta : \Re^{n+1} \to \Re_+$ by

$$\beta(z) := \gamma \min\{1, \psi(z)\}.$$

Let

$$\Omega := \{z := (\varepsilon, x) \in \Re \times \Re^n \mid \varepsilon \geq \beta(z)\bar{\varepsilon}\}.$$

Then, because for any $z \in \Re^{n+1}$, $\beta(z) \leqslant \gamma < 1$, it follows that for any $x \in \Re^n$,

$$(\bar{\varepsilon}, x) \in \Omega.$$

Algorithm 4.2 (Squared smoothing Newton method [32]). *Step* 0. *Choose constants* $\delta \in (0, 1)$ *and* $\sigma \in (0, 1/2)$. *Let* $\varepsilon^0 := \bar{\varepsilon}$, $x^0 \in \Re^n$ *be an arbitrary point and* $k := 0$.
Step 1. *If* $Q(z^k) = 0$ *then stop. Otherwise, let* $\beta_k := \beta(z^k)$.
Step 2. *Compute* $\Delta z^k := (\Delta\varepsilon^k, \Delta x^k) \in \Re \times \Re^n$ *by*

$$Q(z^k) + Q'(z^k)\Delta z^k = \beta_k \bar{z}. \tag{4.1}$$

Step 3. Let l_k *be the smallest nonnegative integer* l *satisfying*

$$\psi(z^k + \delta^l \Delta z^k) \leqslant [1 - 2\sigma(1 - \gamma\bar{\varepsilon})\delta^l]\psi(z^k). \tag{4.2}$$

Define $z^{k+1} := z^k + \delta^{l_k}\Delta z^k$.
Step 4. *Replace* k *by* $k + 1$ *and go to Step* 1.

The following global convergence result is proved in Qi et al. [32].

Theorem 4.3. *Suppose that Assumption* 4.1 *is satisfied. Then an infinite sequence* $\{z^k\}$ *is generated by Algorithm* 4.2 *with* $\{z^k\} \in \Omega$ *and each accumulation point* $\tilde{z}$ *of* $\{z^k\}$ *is a solution of* $Q(z) = 0$.

When F is a generalized P_0-function on $L(X)$, we have the following stronger result.

Theorem 4.4. *Suppose that* Q *is defined by* (3.19) *with* P *given by*

$$P(\varepsilon,x) := x - G(\varepsilon, x - (F(x) + \varepsilon x)), \quad (\varepsilon,x) \in \Re^{n+1},$$

where G *is continuously differentiable on* $\Re_{++} \times \Re^n$ *and for any* $(\varepsilon,x) \in \Re_{++} \times \Re^n$, $G'_x(\varepsilon,x)$ *is symmetric, positive semidefinite and* $\|G'_x(\varepsilon,x)\| \leqslant 1$. *Suppose that* X *is of the structure* (2.3) *and that* F *is a continuously differentiable generalized* P_0*-function on* $L(X)$. *If the solution set of the VIP is nonempty and bounded, then a bounded infinite sequence* $\{z^k\}$ *is generated by Algorithm* 4.2 *and each accumulation point* $\tilde{z}$ *of* $\{z^k\}$ *is a solution of* $Q(z) = 0$.

Proof. By the proof of Corollary 2.6 we can prove that for any $\varepsilon > 0$ and $x \in \Re^n$, $P'_x(\varepsilon,x)$ is a quasi P-matrix, and so, $Q'(\varepsilon,x)$ is nonsingular. Then, by Theorem 4.3, an infinite $\{z^k\}$ is generated by Algorithm 4.2 and $\{z^k\} \in \Omega$. Since $\psi(z^k)$ is a decreasing sequence, $\lim_{k\to\infty} \psi(z^k)$ exists. If

$$\lim_{k\to\infty} \psi(z^k) > 0, \tag{4.3}$$

then there exists a $\varepsilon' > 0$ such that $\varepsilon^k > \varepsilon'$. By the proof of Theorem 2.8, we can prove that for any $\delta > 0$ the following set

$$\{x \in \Re^n \mid \|M(\varepsilon,x)\| \leqslant \delta,\ \varepsilon \in [\varepsilon', \bar{\varepsilon}]\}$$

is bounded, where

$$M(\varepsilon,x) := x - \Pi_X[x - (F(x) + \varepsilon x)].$$

Since G is globally Lipschitz continuous, for any $\delta > 0$, the following set:

$$\{x \in \Re^n \mid \|P(\varepsilon,x)\| \leqslant \delta,\ \varepsilon \in [\varepsilon', \bar{\varepsilon}]\}$$

is also bounded. This means that $\{z^k\}$ is bounded. From Theorem 4.2, $\{z^k\}$ has at least one accumulation point which is a solution of $H(z) = 0$. This contradicts (4.3). Therefore, we have

$$\lim_{k\to\infty} \psi(z^k) = 0. \tag{4.4}$$

Since F is a generalized P_0-function, it is not difficult to show that $Q : \Re^n \to \Re^n$ is a weakly univalent function [11]. By the assumption that the solution set of the VIP is nonempty and bounded, the inverse image $Q^{-1}(0)$ is nonempty and bounded. Then, by [33, Theorem 1], we know from (4.4) that $\{z^k\}$ is bounded and any accumulation point of $\{z^k\}$ is a solution of $Q(z) = 0$. □

In [32], Qi et al. provided a superlinear (quadratic) convergence result by assuming that Q is semismooth at a solution point. Here we cannot prove that Q is semismooth under our assumptions. Nevertheless, we still can provide a superlinear (quadratic) convergence result. Its proof can be drawn similarly from that in [32]. This shows that semismoothess, together with some kind of nondegeneracy, is sufficient but not necessary for superlinear (quadratic) convergence. This fact has been first observed and exploited by Kummer [19,20] without using semismoothness and later by several authors, among them are Sun et al. [38], Fischer [8], Qi et al. [28], Pu and Zhang [25] and Gowda and Ravindran [10].

Theorem 4.5. *Suppose that Assumption* 4.1 *is satisfied and that* $z^* := (0, x^*)$ *is an accumulation point of the infinite sequence* $\{z^k\}$ *generated by Algorithm* 4.2. *Suppose that for any* $\varepsilon > 0$ *and* $d \in \Re^n$ *with* $(\varepsilon, d) \to 0$ *we have*

$$P(\varepsilon, x^* + d) - P(0, x^*) - P'(\varepsilon, x^* + d)\begin{pmatrix} \varepsilon \\ d \end{pmatrix} = \mathrm{o}(\|(\varepsilon, d)\|)$$

and that all $V \in \partial Q(z^*)$ *are nonsingular. Then the whole sequence* $\{z^k\}$ *converges to* z^*,

$$\|z^{k+1} - z^*\| = \mathrm{o}(\|z^k - z^*\|)$$

and

$$\varepsilon^{k+1} = \mathrm{o}(\varepsilon^k).$$

Furthermore, if for any $\varepsilon > 0$ *and* $d \in \Re^n$ *with* $(\varepsilon, d) \to 0$ *we have*

$$P(\varepsilon, x^* + d) - P(0, x^*) - P'(\varepsilon, x^* + d)\begin{pmatrix} \varepsilon \\ d \end{pmatrix} = \mathrm{O}(\|(\varepsilon, d)\|^2),$$

then

$$\|z^{k+1} - z^*\| = \mathrm{O}(\|z^k - z^*\|^2)$$

and

$$\varepsilon^{k+1} = \mathrm{O}(\varepsilon^k)^2.$$

In Theorem 4.5 we assume that all $V \in \partial Q(z^*)$ are nonsingular in order to get a high order convergent result. According to Theorems 3.3 and 3.7, this assumption is satisfied by assuming that all

$$T \in I - \operatorname{plen} \partial \Pi_X[x^* - F(x^*)][I - F'(x^*)]$$

are nonsingular. By the definition of plen, the latter is equivalent to say that all

$$T \in I - \partial \Pi_X[x^* - F(x^*)][I - F'(x^*)]$$

are nonsingular. See Section 2 for the conditions to ensure the nonsingularity of the above matrices. In particular, if $F'(x^*)$ is positive definite, all

$$T \in I - \partial \Pi_X(x^* - F(x^*))[I - F'(x^*)]$$

are nonsingular.

5. Concluding remarks

In this paper, by using convolution, we reformulated the VIP equivalently as smoothing-nonsmooth equations, which have some desirable properties. A globally and locally high-order convergent Newton method has been applied for solving the smoothing-nonsmooth equations, and so the VIP. There is no specific assumption on the structure of the constraint set X. Due to a multivariate integral involved, it may not be easy to compute the smoothing functions of the projection operator Π_X via convolution. However, it may lead to the discovery of some effective ways to compute smoothing functions of Π_X when the structure of X can be used. In fact, based on this observation, an effective way to compute the smoothing functions of Π_X is proposed in [30] when X can be expressed as the set defined by several twice continuously differentiable convex functions. We believe that the research done in this paper can deepen the understanding of smoothing functions of Π_X when X is not a rectangle.

Acknowledgements

The authors would like to thank Houduo Qi for his comments and suggestions while writing this paper.

References

[1] B. Chen, P.T. Harker, A non-interior-point continuation method for linear complementarity problems, SIAM J. Matrix Anal. Appl. 14 (1993) 1168–1190.

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

[3] X. Chen, L. Qi, D. Sun, Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities, Math. Comput. 67 (1998) 519–540.

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

[5] B.C. Eaves, On the basic theorem of complementarity, Math. Programming 1 (1971) 68–75.

[6] Y.M. Ermoliev, V.I. Norkin, R.J.-B. Wets, The minimization of semicontinuous functions: mollifier subgradients, SIAM J. Control Optim. 33 (1995) 149–167.

[7] F. Facchinei, J.S. Pang, Total stability of variational inequalities, Preprint, Department of Mathematical Sciences, Whiting School of Engineering, The Johns Hopkins University, Baltimore, Maryland 21218-2682, USA, 1998.

[8] A. Fischer, Solution of monotone complementarity problems with locally Lipschitzian functions, Math. Programming 76 (1997) 513–532.

[9] S.A. Gabriel, J.J. Moré, Smoothing of mixed complementarity problems, in: M.C. Ferris, J.S. Pang (Eds.), Complementarity and Variational Problems: State of the Art, SIAM, Philadelphia, PA, 1997, pp. 105–116.

[10] M.S. Gowda, G. Ravindran, Algebraic univalence theorems for nonsmooth functions, Research Report, Department of Mathematics and Statistics, University of Maryland, Baltimore Country, Baltimore, Maryland 21250, 1998.

[11] M.S. Gowda, R. Sznajder, Weak univalence and connectedness of inverse images of continuous functions, Math. Oper. Res. 24 (1999) 255–261.

[12] M.S. Gowda, M.A. Tawhid, Existence and limiting behavior of associated with P_0-equations, Comput. Optim. Appl. 12 (1999) 229–251.

[13] A.M. Gupal, On a method for the minimization of almost differentiable functions, Kibernetika (1977) 114–116.

[14] P.T. Harker, J.-S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications, Math. Programming 48 (1990) 161–220.

[15] J.-B. Hiriart-Urruty, Characterizations of the plenary hull of the generalized Jacobian matrix, Math. Programming Study 17 (1982) 1–12.

[16] A.D. Ioffe, Nonsmooth analysis: differentiable calculus of nondifferentiable mappings, Trans. Amer. Math. Soc. 266 (1981) 1–56.

[17] H. Jiang, L. Qi, Local uniqueness and convergence of iterative methods for nonsmooth variational inequalities, J. Math. Anal. Appl. 196 (1995) 314–331.

[18] C. Kanzow, Some noninterior continuation methods for linear complementarity problems, SIAM J. Matrix Anal. Appl. 17 (1996) 851–868.
[19] B. Kummer, Newton's method for nondifferentiable functions, in: J. Guddat et al. (Eds.), Advances in Mathematical Optimization, Akademie-Verlag, Berlin, 1988, pp. 114–125.
[20] B. Kummer, Newton's method based on generalized derivatives for nonsmooth functions: convergence analysis, in: W. Oettli and D. Pallaschke (Eds.), Advances in Optimization, Springer, Berlin, 1992, pp. 171–194.
[21] D.Q. Mayne, E. Polak, Nondifferential optimization via adaptive smoothing, J. Optim. Theory Appl. 43 (1984) 19–30.
[22] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.
[23] J.-S. Pang, L. Qi, Nonsmooth equations: motivation and algorithms, SIAM J. Optim. 3 (1993) 443–465.
[24] J.M. Peng, M. Fukushima, A hybrid Newton method for solving the variational inequality problem via the D-gap function, Math. Programming, 86 (1999) 367–386.
[25] D. Pu, J. Zhang, Inexact generalized Newton methods for second-order C-differentiable optimization, J. Comput. Appl. Math. 93 (1998) 107–122.
[26] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Math. Oper. Res. 18 (1993) 227–244.
[27] L. Qi, X. Chen, A globally convergent successive approximation method for severely nonsmooth equations, SIAM J. Control Optim. 33 (1995) 402–418.
[28] L. Qi, D. Ralph, G. Zhou, Semiderivative functions and reformulation methods for solving complementarity and variational problems, in: G. Di Pillo, F. Giannessi (Eds.), Nonlinear Optimization and Applications 2, Kluwer Academic Publishers B.V., Dordrecht, to appear.
[29] L. Qi, J. Sun, A nonsmooth version of Newton's method, Math. Programming 58 (1993) 353–367.
[30] L. Qi, D. Sun, Smoothing functions and a smoothing Newton method for complementarity and variational inequality problems, AMR 98/24, Applied Mathematics Report, School of Mathematics, the University of New South Wales, October 1998.
[31] L. Qi, D. Sun, A survey of some nonsmooth equations and smoothing Newton methods, in: A. Eberhard, B. Glover, R. Hill, D. Ralph (Eds.), Progress in Optimization: Contributions from Australasia, Kluwer Academic, Dordrecht, 1999.
[32] L. Qi, D. Sun, G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, AMR 97/13, Applied Mathematics Report, School of Mathematics, the University of New South Wales, Sydney 2052, Australia, June 1997. Revised August 1998.
[33] G. Ravindran, M.S. Gowda, Regularization of P_0-functions in box variational inequality problems, SIAM J. Optim., to appear.
[34] S.M. Robinson, Normal maps induced by linear transformation, Math. Oper. Res. 17 (1992) 691–714.
[35] L. Schwartz, Théorie des Distributions, Hermann, Paris, 1966.
[36] S. Smale, Algorithms for solving equations, Proceedings of the International Congress of Mathematicians, Berkeley, CA, 1986, pp. 172–195.
[37] S.L. Sobolev, Some Applications of Functional Analysis in Mathematical Physics, 3rd Edition, Nauka, Moscow, 1988.
[38] D. Sun, M. Fukushima, L. Qi, A computable generalized Hessian of the D-gap function and Newton-type methods for variational inequality problem, in: M.C. Ferris, J.-S. Pang (Eds.), Complementarity and Variational Problems – State of the Art, SIAM Publications, Philadelphia, 1997, pp. 452–473.
[39] T.H. Sweetser, A set-valued strong derivative for vector-valued Lipschitz functions, J. Optim. Theory Appl. 23 (1977) 549–562.
[40] H. Xu, Deterministically computable generalized Jacobians and Newton's methods, Preprint, School of Information Technology and Mathematics Sciences, University of Ballarat, Victoria, VIC 3353, Australia, 1998.
[41] H. Xu, X. Chang, Approximate Newton methods for nonsmooth equations, J. Optim. Theory Appl. 93 (1997) 373–394.
[42] E.H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory, in: E.H. Zarantonello (Ed.), Contributions to Nonlinear Functional Analysis, Academic Press, New York, 1971, pp. 237–424.