Journal of Global Optimization **27:** 215–232, 2003.

Semismooth Newton Methods for Solving Semi-Infinite Programming Problems

LIQUN QI[1], SOON-YI WU[2] and GUANGLU ZHOU[3]
[1]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: maqilq@polyu.edu.hk);* [2]*Institute of Applied Mathematics, National Cheng-Kung University, Tainan 700, Taiwan (e-mail: soonyi@mail.ncku.edu.tw);* [3]*Department of Mathematics, National University of Singapore, Singapore (e-mail: smazgl@nus.edu.sg).*

Abstract. In this paper we present some semismooth Newton methods for solving the semi-infinite programming problem. We first reformulate the equations and nonlinear complementarity conditions derived from the problem into a system of semismooth equations by using NCP functions. Under some conditions a solution of the system of semismooth equations is a solution of the problem. Then some semismooth Newton methods are proposed for solving this system of semismooth equations. These methods are globally and superlinearly convergent. Numerical results are also given.

Key words: Semi-infinite programming, semismooth equations, semismooth Newton method

1. Introduction

Consider the semi-infinite programming (SIP) problem in the following form:

$$\min\{f(x) : x \in X\} \tag{1}$$

where $X = \{x \in \Re^n : g(x, v) \leqslant 0,\ \forall v \in V\}$, V is a nonempty compact subset of $\Re^m$, defined by $V = \{v \in \Re^m : c(v) \leqslant 0\}$, $f : \Re^n \to \Re$, $g : \Re^n \times \Re^m \to \Re$ and $c : \Re^m \to \Re^q$ are twice continuously differentiable functions. For any $x \in \Re^n$, let

$$V(x) = \{v \in V : g(x, v) = 0\},$$

$$T(x) = \{\nabla_x g(x, v) : v \in V(x)\}$$

and

$$r = r(x) = \text{rank}\,\{T(x)\},$$

where rank $T(x)$ means the cardinality of a maximum independent subset of $T(x)$. Then

$$0 \leqslant r(x) \leqslant n.$$

Such a SIP problem has wide applications [16, 24] and many algorithms have been designed to solve this problem, see [3, 4, 13, 14, 16, 17, 27, 28, 29, 30, 31]. It is well-known [26] that if x is a local minimum of the SIP problem (1) and the extended Mangasarian-Fromovitz constraint qualification (EMFCQ) holds, i.e., there is a vector $h \in \Re^n$ such that

$$(\nabla_x g(x, v))^T h < 0$$

for all $v \in V(x)$, then there are p positive numbers u_i and p vectors $v^i \in V(x)$ such that

$$\nabla f(x) + \sum_{i=1}^{p} u_i \nabla_x g(x, v^i) = 0, \tag{2}$$

with $p \leqslant n$. If the EMFCQ does not hold, the optimality condition (2) may not hold. The following is such an example: $n = m = 1$, $V = [-1, 1]$, $f(x) = x$, $g(x, v) = -(x + v^2)(x + 2v^2)$. Then $X = (-\infty, -2] \cup [0, \infty)$. Thus, $x = 0$ is a local minimum. $V(0) = \{0\}$. But $\nabla f(0) = 1$ and $\nabla_x g(0, 0) = 0$. Then the left hand side of (2) is equal to 1, i.e., (2) cannot hold for this example. However, it is also clear that the EMFCQ does not hold at $x = 0$ since $\nabla_x g(0, 0) = 0$.

By the theory of basic feasible solutions of linear programming, we may always find adequate $u_i > 0$ and $v^i \in V(x)$ for (2) such that $\{\nabla_x g(x, v^i) : i = 1, \dots, p\}$ is linearly independent and $p \leqslant r$.

On the other hand, if $x \in X$ with p positive numbers u_i and p vectors $v^i \in V(x)$ satisfies (2), we call x a **stationary point** of the SIP problem, and call $u \in \Re^p$ and v^i for $i = 1, \dots, p$ its **Lagrange multiplier** and **attainers** respectively. If $\{\nabla_x g(x, v^i) : i = 1, \dots, p\}$ is linearly independent, we say that the Lagrange multiplier $u \in \Re^p$ is **regular**.

Naturally, we may think to find a stationary point x for the SIP problem. To this purpose, we may relax the condition that $u_i > 0$ to $u_i \geqslant 0$ for constructing algorithms. Finally, we may always drop those v^i for which $u_i = 0$.

We may also explicitly write out the conditions $x \in X$ and $v^i \in V(x)$ as

$$g(x, v) \leqslant 0, \ \forall v \in V, \tag{3}$$

and for $i = 1, \dots, p$,

$$u_i \geqslant 0, \ g(x, v^i) = 0. \tag{4}$$

Since $v^i \in V(x)$ and $x \in X$, v^i for $i = 1, \dots, p$ are global minima of the nonlinear programming problem

$$\min\{-g(x, v) : c(v) \leqslant 0\}. \tag{5}$$

Thus, if a constraint qualification for the nonlinear programming problem (5) holds, then there are p **auxiliary Lagrange multipliers** $w^i \in \Re^q$ for $i = 1, \dots, p$ such that for $i = 1, \ldots, p$,

$$\begin{aligned}
&\nabla_v g(x, v^i) - \sum_{j=1}^{q} w_j^i \nabla c_j(v^i) = 0, \\
&w_j^i \geqslant 0, \ c_j(v^i) \leqslant 0, \\
&w_j^i c_j(v^i) = 0, \ \text{for } j = 1, \ldots, q.
\end{aligned} \tag{6}$$

Well-known constraint qualifications for nonlinear programming include the linear independence constraint qualification [12], the Slater constraint qualification [12], the Mangasarian-Fromovitz constraint qualification [12], the constant rank constraint qualification [8], [23], etc.

We call an $x \in \Re^n$ with $u \in \Re^p$, $v^i \in \Re^m$ and $w^i \in \Re^q$, for $i = 1, \ldots, p$, $p \leqslant n$, satisfying (2), (3), (4) and (6) a **substationary point** of the SIP problem.

If some second-order sufficient conditions hold for (5) at v^i for $i = 1, \ldots, p$, then a substationary point x is a stationary point of the SIP problem. If some second-order sufficient conditions hold for (1) at x, then a stationary point x is a local minimum of the SIP problem. It is thus desirable to find a substationary point of the SIP problem. Note that there are only a finite number of equalities and inequalities for x, u, v^i and w^i in (2), (4) and (6). But (3) involves an infinite number of inequalities for x. In most applications, the following assumption holds [17, 27].

(A0). For any fixed x, the number of local minima of (5) is finite.

Note that this number depends upon x, and is unknown in general. Thus, if (A0) holds, we may solve the finite system (2), (4) and (6), find its solution x, and check if (3) holds for these finitely many minima of (5) at x. If (3) holds at these points, then x is a substationary point of the SIP problem. In some cases, (3) automatically holds. For example, if $g(x, \cdot)$ is concave, c is convex and $p \geqslant 1$, then a solution of the finite system (2), (4) and (6) is a substationary point of the SIP problem automatically.

Such omission of (3) under Assumption (A0) may omit some solutions x with $p = 0$. But those x are solutions of

$$\nabla f(x) = 0, \tag{7}$$

which is not a difficult problem in general. After finding some solutions of (7), we need to check if (3) holds or not for these solutions.

The system (2), (4) and (6), involves some equations and nonlinear complementarity conditions. In recent years, based on the superlinear convergence theory of generalized Newton methods for solving semismooth equations [18, 22, 15], the Fischer-Burmeister function and other semismooth NCP functions [7, 19], globally and locally superlinearly or quadratically convergent Newton methods have been developed for solving the nonlinear complementarity (NCP) problems and the KKT systems [5, 10, 20, 6, 32, 19]. We thus intend to reformulate the optimality condition system (2), (4) and (6), into a system of semismooth equations (SE), then develop globally and locally superlinearly (quadratically) convergent semismooth Newton methods for solving them.

In Section 2, we give a SE reformulation of the system (2), (4) and (6). This SE reformulation is a system of $n + (m + q + 1)p$ semismooth equations of $n + (m + q + 1)p$ variables. In Section 3, we discuss conditions for superlinear or quadratic convergence of the generalized Newton methods for solving this SE system. In Section 4, based upon line search, we construct globally and locally superlinearly (quadratically) convergent semismooth Newton methods for solving this SE system, and prove their convergence properties. Numerical results are reported in Section 5. Some conclusions are given in the last section.

Another approach for solving SE systems is by smoothing Newton methods [1, 21]. A smoothing Newton method for solving SE systems from the SIP problem is discussed in [11].

The following notation will be used. We denote the n-dimensional unit square matrix by I_n. For a continuously differentiable function $\Phi : \Re^m \to \Re^m$, we denote the Jacobian of Φ at $x \in \Re^m$ by $\Phi'(x)$, whereas the transposed Jacobian as $\nabla\Phi(x)$. $\|\cdot\|$ denotes the Euclidean norm. For a function $f : \Re^n \times \Re^m \to \Re$ we denote $\nabla_x f(x, y)$ the gradient of f at (x, y) with respect to x and $\nabla^2_{xx} f(x, y)$, $\nabla^2_{xy} f(x, y)$ and $\nabla^2_{yy} f(x, y)$ denote, respectively, the $n \times n$, $n \times m$ and $m \times m$ Hessian matrices of f at (x, y).

2. A SE Reformulation

We first briefly review NCP and semismooth functions.

A function $\phi : \Re^2 \to \Re$ is called an NCP function [19] if $\phi(a, b) = 0$ if and only if $a \geqslant 0, b \geqslant 0$ and $ab = 0$. Two well-known NCP functions are the minimum function

$$\phi_{min}(a, b) = \min\{a, b\}$$

and the Fischer-Burmeister function [7, 19]

$$\phi_{FB}(a, b) = \sqrt{a^2 + b^2} - a - b.$$

A locally Lipschitz function $F : \Re^n \to \Re^m$ is called semismooth [18, 22, 15] at $x \in \Re^n$ if F is directionally differentiable at x and for all $V \in \partial F(x + d)$ and $d \to 0$,

$$F'(x; d) = Vd + o(\|d\|),$$

where ∂F is the generalized Jacobian of F in the sense of Clarke [2]. F is called strongly semismooth [22, 20, 7, 19] at x if F is semismooth at x and for all $V \in \partial F(x + d)$ and $d \to 0$,

$$F(x + d) - F(x) = Vd + O(\|d\|^2).$$

Both the minimum function and the Fischer-Burmeister function are not smooth (continuously differentiable), but they are strongly semismooth. Suppose $n = m$. We say F is *CD-regular* at a point x if all $V \in \partial F(x)$ are nonsingular.

We now put (2), (4) and (6) together as:

$$
\begin{aligned}
&\nabla f(x) + \sum_{i=1}^{p} u_i \nabla_x g(x, v^i) = 0,\\
&u_i \geqslant 0, \quad g(x, v^i) = 0, \ \text{for } i = 1, \ldots, p,\\
&\nabla_v g(x, v^i) - \sum_{j=1}^{q} w_j^i \nabla c_j(v^i) = 0, \ \text{for } i = 1, \ldots, p,\\
&w_j^i \geqslant 0, \ c_j(v^i) \leqslant 0, \ \text{for } i = 1, \ldots, p, \ j = 1, \ldots, q,\\
&w_j^i c_j(v^i) = 0, \ \text{for } i = 1, \ldots, p, \ j = 1, \ldots, q.
\end{aligned}
\tag{8}
$$

Note that x, u, v^i and w^i are unknown vectors here. Since p depends upon x, we call it the finite variable-dimensional optimality condition of the SIP problem.

Assume that $u_i > 0$ for $i = 1, 2, \ldots, p$. Multiplying the third equation in (8) by u_i and substituting w_j^i by $u_i w_j^i$ for $i = 1, \ldots, p$, $j = 1, \ldots, q$, then the system (8) is equivalent to the following system. This substitution is necessary for convergence analysis of our algorithm.

$$
\begin{aligned}
&\nabla f(x) + \sum_{i=1}^{p} u_i \nabla_x g(x, v^i) = 0,\\
&u_i \geqslant 0, \quad g(x, v^i) = 0, \ \text{for } i = 1, \ldots, p,\\
&u_i \nabla_v g(x, v^i) - \sum_{j=1}^{q} w_j^i \nabla c_j(v^i) = 0, \ \text{for } i = 1, \ldots, p,\\
&w_j^i \geqslant 0, \ c_j(v^i) \leqslant 0, \ \text{for } i = 1, \ldots, p, \ j = 1, \ldots, q,\\
&w_j^i c_j(v^i) = 0, \ \text{for } i = 1, \ldots, p, \ j = 1, \ldots, q.
\end{aligned}
\tag{9}
$$

Let ϕ be a semismooth NCP function. Then we may reformulate (9) as a system of semismooth equations:

$$
\begin{aligned}
&\nabla f(x) + \sum_{i=1}^{p} u_i \nabla_x g(x, v^i) = 0,\\
&\phi(u_i, -g(x, v^i)) = 0, \ \text{for } i = 1, \ldots, p,\\
&u_i \nabla_v g(x, v^i) - \sum_{j=1}^{q} w_j^i \nabla c_j(v^i) = 0, \ \text{for } i = 1, \ldots, p,\\
&\phi(w_j^i, -c_j(v^i)) = 0, \ \text{for } i = 1, \ldots, p, \ j = 1, \ldots, q.
\end{aligned}
\tag{10}
$$

Define H by

$$H(z) = \begin{pmatrix} \nabla f(x) + \sum_{i=1}^{p} u_i \nabla_x g(x, v^i) \\ \phi(u_1, -g(x, v^1)) \\ \vdots \\ \phi(u_p, -g(x, v^p)) \\ \\ u_1 \nabla_v g(x, v^1) - \sum_{j=1}^{q} w_j^1 \nabla c_j(v^1) \\ \vdots \\ u_p \nabla_v g(x, v^p) - \sum_{j=1}^{q} w_j^p \nabla c_j(v^p) \\ \phi(w_1^1, -c_1(v^1)) \\ \vdots \\ \phi(w_q^1, -c_q(v^1)) \\ \vdots \\ \phi(w_1^p, -c_1(v^p)) \\ \vdots \\ \phi(w_q^p, -c_q(v^p)) \end{pmatrix}, \tag{11}$$

where $z = (x, u, v, w) \in \Re^{n+(m+q+1)p}$, $u \in \Re^p$, $v = (v^1, \ldots, v^p) \in \Re^{mp}$ and $w = (w^1, \ldots, w^p) \in \Re^{qp}$. Let $H = H_1$ if $\phi = \phi_{min}$ and $H = H_2$ if $\phi = \phi_{FB}$.

REMARK 1. In appearance, the system of equations (10) is not “totally” equivalent to (8). It allows the case that

$$u_i = 0, \quad g(x, v^i) < 0.$$

But if there is an $n + (m + q + 1)p$ dimensional vector satisfying (10), we may drop the part indexed by i where $u_i = 0$. Thus, we get a solution of (8). On the other hand, a solution of (8) obviously satisfies (10). In this sense, (8) and (10) are equivalent.

REMARK 2. The parameter p depends upon the problem. One possibility is to use $p_k = r(x^k)$ at the $(k + 1)$th iteration to find x^{k+1}. In the latter part of this paper, we study a simple case: p is known. This case happens in applications. For example, if (A0) holds and for any fixed x, $g(x, \cdot)$ is a concave function, then $p = 0$ or 1. Since the solution for $p = 0$ is checked by solving (7) as discussed before, under that additional assumption, a method for solving the case $p = 1$, combining a solution of (7) satisfying (3), will solve the problem. In fact, if p is unknown but small, say $p = 2$ or 3, we may try $p = 1$ first, if it fails, we may try $p = 2$, so on. In Section 5, we give three examples with $p = 1$ and one example with $p = 2$. The case that p is unknown and not small will be studied in future research.

3. Superlinear Convergence of Semismooth Newton Methods

A semismooth Newton method for solving $H(z) = 0$ may be defined as the following: having the vector z^k, compute z^{k+1} by

$$z^{k+1} = z^k - W_k^{-1} H(z^k), \quad W_k \in \partial H(z^k). \tag{12}$$

The following theorem was proved by Qi and Sun [22]:

THEOREM 1. *Suppose that z^* is a solution of $H(z) = 0$, H is locally Lipschitzian, semismoooth and CD-regular at z^*. Then the iteration method (12) is well defined and the sequence $\{z^k\}$ generated by (12) converges to z^* Q-superlinearly in a neighborhood of z^*. If in addition H is strongly semismooth at z^*, then the convergence is Q-quadratic.*

Let $z = (x, u, v, w) \in \Re^{n+(m+q+1)p}$ be a solution of (11), where $v = (v^1, \ldots, v^p) \in \Re^{mp}$ and $w = (w^1, \ldots, w^p) \in \Re^{qp}$. Let $Q = \{1, 2, \ldots, q\}$ and $P = \{1, 2, \ldots, p\}$. We make the following assumptions.

(A1). For all $i \in P$, $u_i > 0$.

(A2). The vectors $\nabla_x g(x, v^i)$, $i \in P$ are linearly independent.

Define

$$F(x, u, v) = \nabla f(x) + \sum_{i=1}^{p} u_i \nabla_x g(x, v^i). \tag{13}$$

For $i \in P$, define

$$I(v^i) = \{j : j \in Q, c_j(v^i) = 0\},$$

$$J(v^i) = Q \backslash I(v^i).$$

and

$$L(x, u_i, v^i, w^i) = u_i g(x, v^i) - \sum_{j=1}^{q} w_j^i c_j(v^i). \tag{14}$$

Let $G(x, v)$ be the set of all $(d, \xi_1, \ldots, \xi_p) \in \Re^n \times \Re^{mp}$ satisfying

$$d^T \nabla_x g(x, v^i) + \xi_i^T \nabla_v g(x, v^i) = 0 \text{ for } i \in P,$$

and

$$\xi_i^T \nabla c_j(v^i) = 0 \text{ for } i \in P,\ \ j \in I(v^i).$$

We further suppose that the following assumptions hold.

(A3). For each $i \in P$, the vectors $\nabla c_j(v^i)$, $j \in I(v^i)$ are linearly independent.
(A4). $w_j^i - c_j(v^i) \neq 0$, for $i \in P$ and $j \in Q$.
(A5). For all $(d, \xi_1, \dots, \xi_p) \in G(x, v)\backslash\{0\}$,

$$d^T \nabla_x F(x, u, v)d + 2\sum_{i=1}^{p} u_i d^T \nabla_{xv}^2 g(x, v^i)\xi_i + \sum_{i=1}^{p} \xi_i^T \nabla_{vv}^2 L(x, u_i, v^i, w^i)\xi_i > 0.$$

Note that **(A5)** is similar to (not the same as) the second order optimality conditions for semi-infinite programming problems, given in [25].

THEOREM 2. *Suppose that $z = (x, u, v, w)$ is a solution of (11) and satisfies (A1)-(A5). Then both H_1 and H_2 are CD-regular at z.*

Proof. For any z, H_2 is differentiable at z if and only if

$$u_i^2 + (g(x, v^i))^2 > 0 \text{ for all } i \in P$$

and

$$(w_j^i)^2 + (c_j(v^i))^2 > 0 \text{ for all } i \in P,\ j \in Q.$$

For these points, we have

$$H_2'(z) = \begin{pmatrix} F_x'(x, u, v) & \nabla_x g(x, v)^T & (\nabla_{xv}^2 L) & 0 \\ \Lambda \nabla_x g(x, v) & \Gamma & \Psi & 0 \\ \nabla_{vx}^2 L & \Psi^T & \nabla_{vv}^2 L & -\nabla c(v)^T \\ 0 & 0 & A & B \end{pmatrix}, \tag{15}$$

where

$$\nabla_x g(x, v) = \begin{pmatrix} \nabla_x g(x, v^1)^T \\ \vdots \\ \nabla_x g(x, v^p)^T \end{pmatrix}, \tag{16}$$

$$\nabla_{xv}^2 L = [u_1 \nabla_{xv}^2 g(x, v^1), \dots, u_p \nabla_{xv}^2 g(x, v^p)], \tag{17}$$

$$\nabla_{vx}^2 L = (\nabla_{xv}^2 L)^T, \tag{18}$$

$$\Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_p\}, \Gamma = \text{diag}\{\gamma_1, \ldots, \gamma_p\},$$

$$\lambda_i = \frac{g(x, v^i)}{\sqrt{u_i^2 + (g(x, v^i))^2}} + 1, \text{ for } i \in P,$$

$$\gamma_i = \frac{u_i}{\sqrt{u_i^2 + (g(x, v^i))^2}} - 1, \text{ for } i \in P,$$

$$\Psi = \begin{pmatrix} \lambda_1 \nabla_v g(x, v^1)^T & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p \nabla_v g(x, v^p)^T \end{pmatrix}, \tag{19}$$

$$\nabla_{vv}^2 L = \text{diag}\{\nabla_{vv}^2 L(x, u_1, v^1, w^1), \ldots, \nabla_{vv}^2 L(x, u_p, v^p, w^p)\}, \tag{20}$$

$$\nabla c(v)^T = \text{diag}\{\nabla c(v^1)^T, \ldots, \nabla c(v^p)^T\},$$

$$\nabla c(v^i) = \begin{pmatrix} \nabla c_1(v^i)^T \\ \vdots \\ \nabla c_q(v^i)^T \end{pmatrix}, \text{ for } i \in P,$$

$$A = \text{diag}\{A_1 \nabla c(v^1), \ldots, A_p \nabla c(v^p)\}, B = \text{diag}\{B_1, \ldots, B_p\},$$

$$A_i = \text{diag}\{a_{i1}, \ldots, a_{iq}\}, B_i = \text{diag}\{b_{i1}, \ldots, b_{iq}\}, \text{ for } i \in P,$$

$$a_{ij} = \frac{c_j(v^i)}{\sqrt{(w_j^i)^2 + (c_j(v^i))^2}} + 1, \text{ for } i \in P, j \in Q$$

and

$$b_{ij} = \frac{w_j^i}{\sqrt{(w_j^i)^2 + (c_j(v^i))^2}} - 1, \text{ for } i \in P, j \in Q.$$

Let z be a solution of $H_2(z) = 0$ and satisfy (A1)-(A5). Then we have $u_i > 0$ and $g(x, v^i) = 0$ for all $i \in P$. By the definition of the generalized Jacobian of H_2, if $W \in \partial H_2(z)$, we have

$$W = \begin{pmatrix} F_x'(x, u, v) & \nabla_x g(x, v)^T & \nabla_{xv}^2 L & 0 \\ \nabla_x g(x, v) & 0 & \Psi & 0 \\ \nabla_{vx}^2 L & \Psi^T & \nabla_{vv}^2 L & -\nabla c(v)^T \\ 0 & 0 & A & B \end{pmatrix}, \tag{21}$$

where

$$\Psi = \begin{pmatrix} \nabla_v g(x, v^1)^T & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \nabla_v g(x, v^p)^T \end{pmatrix},$$

$$A = \mathrm{diag}\{A_1 \nabla c(v^1), \ldots, A_p \nabla c(v^p)\}, B = \mathrm{diag}\{B_1, \ldots, B_p\},$$

$$A_i = \mathrm{diag}\{a_{i1}, \ldots, a_{iq}\}, B_i = \mathrm{diag}\{b_{i1}, \ldots, b_{iq}\}, \text{ for } i \in P,$$

$$a_{ij} = 1, \; b_{ij} = 0, \text{ for } i \in P, \; j \in I(v^i), \tag{22}$$

and

$$a_{ij} = 0, \; b_{ij} = -1, \text{ for } i \in P, \; j \in Q \backslash I(v^i). \tag{23}$$

Suppose that

$$W \begin{pmatrix} d_1 \\ d_2 \\ \xi \\ \zeta \end{pmatrix} = 0, \tag{24}$$

where $d_1 \in \Re^n$, $d_2 \in \Re^p$, $\xi = (\xi_1^T, \xi_2^T, \ldots, \xi_p^T)^T \in \Re^{mp}$ and $\zeta = (\zeta_1^T, \zeta_2^T, \ldots, \zeta_p^T)^T \in \Re^{qp}$. Then (24) implies

$$F'_x(x, u, v)d_1 + \nabla_x g(x, v)^T d_2 + \sum_{i=1}^{p} u_i \nabla^2_{xv} g(x, v^i)\xi_i = 0, \tag{25}$$

$$\nabla_x g(x, v)^T d_1 + \Psi\xi = 0, \tag{26}$$

$$u_i \nabla^2_{vx} g(x, v^i) d_1 + d_{2i} \nabla_v g(x, v^i) + \nabla^2_{vv} L(x, u_i, v^i, w^i)\xi_i$$

$$-\nabla c(v^i)^T \zeta_i = 0, \text{ for } i \in P, \tag{27}$$

where d_{2i} is the ith component of d_2,

$$a_{ij} \nabla c_j(v^i)^T \xi_i + b_{ij}\zeta_{ij} = 0, \text{ for } i \in P, \; j \in Q. \tag{28}$$

Let ζ_{ij} denote the jth element of ζ_i. By (28), (22) and (23), for $i \in P$,

$$\zeta_{ij} = 0 \tag{29}$$

if $j \in Q \backslash I(v^i)$; and

$$\nabla c_j(v^i)^T \xi_i = 0, \tag{30}$$

if $j \in I(v^i)$. For $i \in P$, multiplying (27) by ξ_i^T, by (29) and (30),

$$u_i \xi_i^T \nabla^2_{vx} g(x, v^i) d_1 + d_{2i} \xi_i^T \nabla_v g(x, v^i) + \xi_i^T \nabla^2_{vv} L(x, u_i, v^i, w^i) \xi_i = 0. \tag{31}$$

Multiplying (25) by d_1^T, we have

$$d_1^T F'_x(x, u, v) d_1 + d_1^T \nabla_x g(x, v)^T d_2 + \sum_{i=1}^{p} u_i d_1^T \nabla^2_{xv} g(x, v^i) \xi_i = 0. \tag{32}$$

From (26),(31) and (32), we have

$$d_1^T F'_x(x, u, v) d_1 + 2 \sum_{i=1}^{p} u_i d^T \nabla^2_{xv} g(x, v^i) \xi_i$$

$$+ \sum_{i=1}^{p} u_i \xi_i^T \nabla^2_{vv} L(x, u^i, v^i, w^i) \xi_i = 0. \tag{33}$$

By (26), (30), (33) and (A5), $d_1 = 0$ and $\xi_i = 0$ for $i \in P$. From (25), $\nabla_x g(x, v)^T d_2 = 0$. By (A2), $d_2 = 0$. Now (27) yields

$$\sum_{j \in I(v^i)} \nabla c_j(v^i) \zeta_{ij} = 0, \text{ for } i \in P.$$

By (A3), $\zeta_{ij} = 0$ for $i \in P$ and $j \in I(v^i)$. Hence $[d_1^T, d_2^T, \xi^T, \zeta^T]^T = 0$. This shows that W is nonsingular. Therefore, H_2 is CD-regular at z.

Let z be a solution of $H_1(z) = 0$ and satisfy (A1)-(A5). Then we have $u_i > 0$ and $g(x, v^i) = 0$ for all $i \in P$. By the definition of the generalized Jacobian of H_1, if $W \in \partial H_1(z)$, we have

$$W = \begin{pmatrix} F'_x(x, u, v) & \nabla_x g(x, v)^T & \nabla^2_{xv} L & 0 \\ -\nabla_x g(x, v) & 0 & -\Psi & 0 \\ \nabla^2_{vx} L & \Psi^T & \nabla^2_{vv} L & -\nabla c(v)^T \\ 0 & 0 & -A & B \end{pmatrix}, \tag{34}$$

where

$$\Psi = \begin{pmatrix} \nabla_v g(x, v^1)^T & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \nabla_v g(x, v^p)^T \end{pmatrix},$$

$$A = \text{diag}\{A_1 \nabla c(v^1), \ldots, A_p \nabla c(v^p)\}, \; B = \text{diag}\{B_1, \ldots, B_p\},$$

$$A_i = \text{diag}\{a_{i1}, \ldots, a_{iq}\}, \; B_i = \text{diag}\{b_{i1}, \ldots, b_{iq}\}, \text{ for } i \in P,$$

$$a_{ij} = 1, \; b_{ij} = 0, \text{ for } i \in P, \; j \in I(v^i), \tag{35}$$

and

$$a_{ij} = 0, \; b_{ij} = 1, \text{ for } i \in P, \; j \in Q \backslash I(v^i). \tag{36}$$

Suppose that

$$W \begin{pmatrix} d_1 \\ d_2 \\ \xi \\ \zeta \end{pmatrix} = 0, \tag{37}$$

where $d_1 \in \Re^n$, $d_2 \in \Re^p$, $\xi = (\xi_1^T, \xi_2^T, \ldots, \xi_p^T)^T \in \Re^{mp}$ and $\zeta = (\zeta_1^T, \zeta_2^T, \ldots, \zeta_p^T)^T \in \Re^{qp}$. Then (37) implies

$$F'_x(x, u, v)d_1 + \nabla_x g(x, v)^T d_2 + \sum_{i=1}^{p} u_i \nabla^2_{xv} g(x, v^i)\xi_i = 0, \tag{38}$$

$$\nabla_x g(x, v)d_1 + \Psi\xi = 0, \tag{39}$$

$$u_i \nabla^2_{vx} g(x, v^i)d_1 + d_{2i} \nabla_v g(x, v^i) + \nabla^2_{vv} L(x, u_i, v^i, w^i)\xi_i$$

$$-\nabla c(v^i)^T \zeta_i = 0, \text{ for } i \in P, \tag{40}$$

$$-a_{ij} \nabla c_j(v^i)^T \xi_i + b_{ij}\zeta_{ij} = 0, \text{ for } i \in P, \; j \in Q. \tag{41}$$

By (41), for $i \in P$,

$$\zeta_{ij} = 0 \tag{42}$$

if $j \in Q \backslash I(v^i)$, here ζ_{ij} denotes the jth element of ζ_i,

$$\nabla c_j(v^i)^T \xi_i = 0, \tag{43}$$

if $j \in I(v^i)$. For $i \in P$, multiplying (40) by ξ_i^T, by (42) and (43),

$$u_i \xi_i^T \nabla^2_{vx} g(x, v^i)d_1 + d_{2i}\xi_i^T \nabla_v g(x, v^i) + \xi_i^T \nabla^2_{vv} L(x, u_i, v^i, w^i)\xi_i = 0. \tag{44}$$

Multiplying (38) by d_1^T, we have

$$d_1^T F_x'(x,u,v)d_1 + d_1^T \nabla_x g(x,v)d_2 + \sum_{i=1}^{p} u_i d_1^T \nabla_{xv}^2 g(x,v^i)\xi_i = 0. \tag{45}$$

From (39), (44) and (45), we have

$$d_1^T F_x'(x,u,v)d_1 + 2\sum_{i=1}^{p} u_i d^T \nabla_{xv}^2 g(x,v^i)\xi_i$$

$$+ \sum_{i=1}^{p} u_i \xi_i^T \nabla_{vv}^2 L(x,u_i,v^i,w^i)\xi_i = 0. \tag{46}$$

By (39), (43), (46) and (A5), $d_1 = 0$ and $\xi_i = 0$ for $i \in P$. From (38), $\nabla_x g(x,v)^T d_2 = 0$. By (A2), $d_2 = 0$. Now (40) yields

$$\sum_{j \in I(v^i)} \nabla c_j(v^i)\zeta_{ij} = 0, \text{ for } i \in P.$$

By (A3), $\zeta_{ij} = 0$ for $i \in P$ and $j \in I(v^i)$. Hence $[d_1^T, d_2^T, \xi^T, \zeta^T]^T = 0$. This shows that W is nonsingular. Therefore, H_1 is CD-regular at z. This completes the proof.

By Theorems 1 and 2, we have the following theorem.

THEOREM 3. *Let $H = H_i$ for $i = 1, 2$. Suppose that $z^* = (x^*, u^*, v^*, w^*)$ is a solution of $H(z) = 0$ and satisfies (A1)-(A5). Then the iteration method (12) is well defined, and the sequence $\{z^k\}$ generated by (12) converges to z^* Q-superlinearly in a neighborhood of z^*. If in addition f, g and c are three times continuously differentiable functions, then the convergence is Q-quadratic.*

4. Damped Newton and Gauss-Newton Methods

In this section we present damped Newton and Gauss-Newton methods for solving $H_2(z) = 0$. Let

$$\theta(z) = \frac{1}{2} H_2(z)^T H_2(z).$$

θ is continuously differentiable with the gradient given by

$$\nabla\theta(z) = W^T H_2(z),$$

where $W \in \partial H_2(z)$.

For any $z = (x, u, v, w) \in \Re^{n+(m+q+1)p}$, where $v = (v^1, \dots, v^p) \in \Re^{mp}$ and $w = (w^1, \dots, w^p) \in \Re^{qp}$, let

$$J_0(x) = \{i \in P : u_i = g(x, v^i) = 0\}$$

and

$$I_0(v^i) = \{j \in Q : w^i_j = c_j(v^i) = 0\}.$$

Define

$$W = \begin{pmatrix} F'_x(x,u,v) & \nabla_x g(x,v)^T & \nabla^2_{xv} L & 0 \\ \Lambda \nabla_x g(x,v) & \Gamma & \Psi & 0 \\ \nabla^2_{vx} L & \Psi^T & \nabla^2_{vv} L & -\nabla c(v)^T \\ 0 & 0 & A & B \end{pmatrix}, \tag{47}$$

where

$$\Lambda = \mathrm{diag}\{\lambda_1, \dots, \lambda_p\}, \Gamma = \mathrm{diag}\{\gamma_1, \dots, \gamma_p\},$$

$$\lambda_i = \frac{g(x, v^i)}{\sqrt{u_i^2 + (g(x, v^i))^2}} + 1, \ \text{for } i \in P \backslash J_0(x),$$

$$\gamma_i = \frac{u_i}{\sqrt{u_i^2 + (g(x, v^i))^2}} - 1, \ \text{for } i \in P \backslash J_0(x),$$

and $\lambda_i = 1$, $\gamma_i = 0$ for $i \in J_0(x)$,

$$A = \mathrm{diag}\{A_1 \nabla c(v^1), \dots, A_p \nabla c(v^p)\}, B = \mathrm{diag}\{B_1, \dots, B_p\},$$

$$A_i = \mathrm{diag}\{a_{i1}, \dots, a_{iq}\}, B_i = \mathrm{diag}\{b_{i1}, \dots, b_{iq}\}, \ \text{for } i \in P,$$

$$a_{ij} = \frac{c_j(v^i)}{\sqrt{(w^i_j)^2 + (c_j(v^i))^2}} + 1, \ \text{for } i \in P, \ j \in Q \backslash I_0(v^i),$$

and

$$b_{ij} = \frac{w^i_j}{\sqrt{(w^i_j)^2 + (c_j(v^i))^2}} - 1, \ \text{for } i \in P, \ j \in Q \backslash I_0(v^i),$$

and $a_{ij} = 1$, $b_{ij} = 0$ for $i \in P$ and $j \in I_0(v^i)$. Then $W \in \partial_B H_2(z) \subseteq \partial H_2(z)$.

ALGORITHM 1. (Generalized Damped Newton Method).

Step 1. Let $z^0 \in \Re^{n+(m+q+1)p}$, $\sigma, \rho \in (0, 1)$, $\eta > 0$, $a > 2$ and $k = 0$.
Step 2. If $H_2(z^k) = 0$, stop. Otherwise, let d^k be a solution of

$$H_2(z^k) + W^k d = 0, \ W^k \in \partial H_2(z^k). \tag{48}$$

If (48) is not solvable, or if $\nabla\theta(z^k)^T d^k > -\eta\|d^k\|^a$, set $d^k = -\nabla\theta(z^k)$.
Step 3. Let $\alpha_k = \rho^{j_k}$, where j_k is the smallest nonnegative integer j such that

$$\theta(z^k + \rho^j d^k) - \theta(z^k) \leqslant \sigma\rho^j \nabla\theta(z^k)^T d^k,$$

where ρ^j means the jth power of ρ.
Step 4. Let $z^{k+1} := z^k + \alpha_k d^k$ and $k := k + 1$. Go to Step 2.

This algorithm is a generalization of the corresponding algorithm for the Fischer-Burmeister equation in [5]. It is also similar to the corresponding algorithm in [10], with an additional steepest descent direction consideration at Step 2, which treats the case that (48) is not solvable. Similar to Theorem 11 of [5] or Theorem 4.1 of [10], we have the following global and superlinear convergence theorem for this algorithm. We omit its proof since it is similar to the proof of Theorem 4.1 of [10], by using Theorem 3 of this paper.

THEOREM 4. *Assume that $z^* = (x^*, u^*, v^*, w^*)$ is an accumulation point of $\{z^k\}$ generated by Algorithm 1. If (A1)-(A5) hold at z^*, then z^* is a solution of $H_2(z) = 0$, and $\{z^k\}$ converges to z^* Q-superlinearly. If in addition f, g and c are three times continuously differentiable functions, then the convergence is Q-quadratic.*

We may also use the Gauss-Newton technique. The following is a generalized damped Gauss-Newton method, which is a generalization of the corresponding algorithm in [9].

ALGORITHM 2. (Generalized Damped Gauss-Newton Method).
Step 1. Let $z^0 \in \Re^{n+(m+q+1)p}$, $\sigma \in (0, \frac{1}{2})$, $\rho \in (0, 1)$, $\beta_0 > 0$, $k = 0$.
Step 2. If $(W^k)^T H_2(z^k) = 0$, where $W^k \in \partial H_2(z^k)$, stop. Otherwise, let d^k be a solution of

$$(W^k)^T H_2(z^k) + [(W^k)^T (W^k) + \beta_k I]d = 0.$$

Step 3. Let $\alpha_k = \rho^{j_k}$, where j_k is the smallest nonnegative integer j such that

$$\theta(z^k + \rho^j d^k) - \theta(z^k) \leqslant \sigma\rho^j \nabla\theta(z^k)^T d^k,$$

where ρ^j means the jth power of ρ.
Step 4. Choose $\beta_{k+1} > 0$. Let $z^{k+1} := z^k + \alpha_k d^k$ and $k := k + 1$. Go to Step 2.

Similar to Theorem 5.1 of [9], we have the following global and superlinear convergence theorem for this algorithm. Again we omit its proof since it is similar to the proof of Theorem 5.1 of [9], by using Theorem 3 of this paper.

Table 1. Results for Algorithm 1

Example	p	Iter	NG	x^k	v^k	$f(x^k)$
1	1	3	0	(-9.53e-02,9.53e-02)	1	2.20e+00
2	1	4	0	(-2.13e-01,-1.36e+00,1.85e+00)	1	5.33e+00
3	1	6	0	(7.20e-01,-1.45e+00)	0	9.72e+01
4	2	7	0	(-7.50e-01,-6.18e-01)	(1,0)	1.94e-01

THEOREM 5. *Let* $\beta_k = \min\{\theta(z^k), \|\nabla\theta(z^k)\|\}$ *in Algorithm 2. Then Algorithm 2 is well-defined. Assume that* $z^* = (x^*, u^*, v^*, w^*)$ *is an accumulation point of* $\{z^k\}$ *generated by Algorithm 2. If (A1)-(A5) hold at* z^*, *then* z^* *is a solution of* $H_2(z) = 0$, *and* $\{z^k\}$ *converges to* z^* *Q-superlinearly. If in addition* f, g *and* c *are three times continuously differentiable functions, then the convergence is Q-quadratic.*

5. Numerical Results

To illustrate the computational behavior of the proposed algorithms in Section 4, Algorithm 1 was implemented in MATLAB and run on a DEC Alpha Server 8200 for the following examples from [4]. Throughout the computational experiments, the parameters used in the algorithm were $\rho = 0.5, a = 2.1, \eta = 10^{-8}$, and $\sigma = 10^{-4}$. We terminated our iteration when $\|H_2(z^k)\| < 10^{-6}$. The numerical results are summarized in Table 1, where *Iter* denotes the number of iterations, *NG* the number of gradient steps, x^k and v^k the final iterate and $f(x^k)$ the function value of f at the final iterate x^k.

EXAMPLE 1. $f(x) = 1.21\exp(x_1) + \exp(x_2)$, $g(x, v) = v - \exp(x_1 + x_2)$, $V = [0, 1]$. $p = 1$, $x^0 = (0, 0)^T$, $v^0 = 0$.

EXAMPLE 2. $f(x) = x_1^2 + x_2^2 + x_3^2$, $g(x, v) = x_1 + x_2\exp(x_3 v) + \exp(2v) - 2\sin(4x)$, $V = [0, 1]$. $p = 1$, $x^0 = (1, 1, 1)^T$, $v^0 = 1$.

EXAMPLE 3. $f(x) = (x_1 - 2x_2 + 5x_2^2 - x_2^3 - 13)^2 + (x_1 - 14x_2 + x_2^2 + x_2^3 - 29)^2$, $g(x, v) = x_1^2 + 2x_2v^2 + \exp(x_1 + x_2) - \exp(v)$, $V = [0, 1]$. $p = 1$, $x^0 = (1, -1)^T$, $v^0 = 0.5$.

EXAMPLE 4. $f(x) = \frac{1}{3}x_1^2 + \frac{1}{2}x_1 + x_2^2$, $g(x, v) = (1 - x_1^2v^2)^2 - x_1v^2 - x_2^2 + x_2$, $V = [0, 1]$. $p = 2$, $x^0 = (-1, -1)^T$, $v_1^0 = 0.5$, $v_2^0 = 1$.

6. Conclusion

In this paper we proposed some semismooth Newton methods for solving the semi-infinite programming problem. Compared with the methods proposed in [3, 4, 16, 17, 27, 28, 29, 30, 31], the advantage of the methods proposed in this paper is that only a system of linear equations needs to be solved at each iteration. The numerical tests reported in this paper are preliminary. Further experience with testing and with actual applications will be necessary. Furthermore, it will be interesting to determine the value p numerically. Thus, the next research topic is to combine the approach in this paper with some other approaches, where p is determined numerically [13, 14, 17, 27].

Acknowledgements

The first author's work was supported by the Australian Research Council and the Research Grant Council of Hong Kong. This work was initiated when he visited National Cheng-Kung University. That trip was supported by the National Sciences Council of Taiwan. The third author's work was supported by the Australian Research Council. The authors would like to thank Masao Fukushima, Lucien Polak, Alexander Shapiro, Donghui Li, Defeng Sun and the anonymous referees for their helpful comments.

References

1. Chen, X., Qi, L. and Sun, D. 1998, Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities, *Math. Comput.* 67, 519–540.
2. Clarke, F.H. 1983, *Optimization and Nonsmooth Analysis*, Wiley, New York.
3. Conn, A.R. and Gould, N.I.M. 1987, An exact penalty function for semi-infinite programming, *Math. Programming* 37, 19–40.
4. Coope, I.D. and Watson, G.A. 1985, A projected Lagrangian algorithm for semi-infinite programming, *Math. Programming* 32, 337–356.
5. De Luca, T., Facchinei, F. and Kanzow, C. 1996, A semismooth equation approach to the solution of nonlinear complementarity problems, *Math. Programming* 75, 407–439.
6. Facchinei, F. and Kanzow, C. 1997, A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems, *Math. Programming* 76, 493–512.
7. Fischer, A., 1997, Solution of monotone complementarity problems with locally Lipschitzian functions, *Math. Programming* 76, 513–532.
8. Janin, R., 1984, Direction derivative of the marginal function in nonlinear programming, *Math. Programming Study* 21, 127–138.
9. Jiang, H., 1999, Global convergence analysis of the generalized Newton and Gauss-Newton methods of the Fischer-Burmeister equation for the complementarity problem, *Math. Oper. Res.* 24, 529–543.
10. Jiang, H. and Qi, L. 1997, A new nonsmooth equations approach to nonlinear complementarity problems, *SIAM J. Control and Optimization* 35, 178–193.

11. Li, D., Qi, L., Tam, J. and Wu, S.Y. Smoothing Newton methods for solving semi-infinite programming problems, to appear in: *J. Global Optimization*.
12. Mangasarian, O.L., 1969 *Nonlinear Programming*, McGraw-Hill, New York, 1969.
13. Mayne, D.Q., Michalska, H. and Polak, E. 1994, An efficient algorithm for solving semi-infinite inequality problems with box constraints, *J. Appl. Math. and Optim.* 30, 135–157.
14. Mayne, D.Q. and Polak, E. 1982, A quadratically convergent algorithm for solving infinite dimensional inequalities, *J. Appl. Math. and Optim.* 9, 25–40.
15. Pang, J.-S. and Qi, L. 1993, Nonsmooth equations: Motivation and algorithms, *SIAM J. Optimization* 3, 443–465.
16. Polak, E., 1997, *Optimization: Algorithms and Consistent Approximation*, Springer-Verlag, New York.
17. Polak, E., and Tits, A. 1982, A recursive quadratic programming algorithm for semi-infinite optimization problems, *Appl. Math. Optim.* 8, 325–349.
18. Qi, L. 1993, Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.* 18 227–244.
19. Qi, L., 1999, Regular pseudo-smooth NCP and BVIP functions and globally and quadratically convergent generalized Newton methods for complementarity and variational inequality problems, *Math. Oper. Res.* 24, 440–471.
20. Qi, L. and Jiang, H. 1997, Semismooth Karush-Kuhn-Tucker equations and convergence analysis of Newton methods and quasi-Newton methods for solving these equations, *Math. Oper. Res.* 22, 301–325.
21. Qi, L., Sun, D. and Zhou, G. 2000, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, *Math. Programming* 87, 1–35.
22. Qi, L. and Sun, J. 1993, A nonsmooth version of Newton's method, *Math. Programming* 58, 353–367.
23. Qi, L., and Wei, Z. 2000, On the constant positive linear dependence condition and its applications to SQP methods, *SIAM J. Optimization* 10, 963–981.
24. Reemsten, R. and Rückmann, J. 1998, Semi-Infinite Programming, Kluwer Academic Publishers, Boston.
25. Rückmann, J.J. and Shapiro, A. 2001, Second Order Optimality Conditions in Generalized Semi-Infinite Programming, *Set-Valued Analysis* 9, 169–186.
26. Shapiro, A. 1998, First and second order optimality conditions and perturbation analysis of semi-infinite programming problems, in: R. Reemsten and J. Rückmann, eds., pp. 103-133 Semi-Infinite Programming, Kluwer Academic Publishers, Boston.
27. Tanaka, Y., Fukushima, M. and Ibaraki, T. 1988, A globally convergent SQP method for semi-infinite nonlinear optimization, *J. Comput. Appl. Math.* 23, 141–153.
28. Tanaka, Y., Fukushima, M. and Ibaraki, T. 1988, A comparative study of several semi-infinite nonlinear programming algorithms, *European J. Oper. Res.* 36, 92–100.
29. Teo, K.L., Yang, X.Q. and Jennings, L.S. 2001, Computational discretization algorithms for functional inequality constrained optimization, *Ann. Oper. Res.* 98, 215–234.
30. Watson, G.A., 1981, Globally convergent methods for semi-infinite programming, *BIT*, 21 362–373.
31. Watson, G.A., 1983, Numerical experiments with globally convergent methods for semi-infinite programming problems, in: Semi-infinite programming and applications, A.V. Fiacco and K.O. Kortanek, eds., (Austin, Tex., 1981), 193–205, Lecture Notes in Econom. and Math. Systems, 215, Springer, Berlin-New York.
32. Yamashita, N. and Fukushima, M. 1997, Modified Newton methods for solving semismooth reformulations of monotone complementarity problems, *Math. Programming* 76, 469–491.