Appl Math Optim 52:73–92 (2005)
DOI: 10.1007/s00245-005-0823-4

Applied Mathematics & Optimization

A Nonsmooth L-M Method for Solving the Generalized Nonlinear Complementarity Problem over a Polyhedral Cone*

Yiju Wang,[1,2] Fengming Ma,[1] and Jianzhong Zhang[3]

[1]Institute of Operations Research, Qufu Normal University,
Rizhao Shandong 276800, People's Republic of China

[2]Department of Applied Mathematics, The Hong Kong Polytechnic University,
Kowloon, Hong Kong
yiju66@hotmail.com; qfmafengming@eyou.com

[3]Department of Mathematics, City University of Hong Kong,
Kowloon, Hong Kong
mazhang@cityu.edu.hk

Abstract. In this paper the generalized nonlinear complementarity problem (GNCP) defined on a polyhedral cone is reformulated as a system of nonsmooth equations. Based on this reformulation, the famous Levenberg–Marquardt (L-M) algorithm is employed to obtain its solution. Theoretical results that relate the stationary points of the merit function to the solution of the GNCP are presented. Under mild assumptions, we show that the L-M algorithm is both globally and superlinearly convergent. Moreover, a method to calculate a generalized Jacobian is given and numerical experimental results are presented.

Key Words. GNCP, Stationary point, Superlinear convergence.

AMS Classification. 65H10, 90C33, 90C30.

1. Introduction

The generalized nonlinear complementarity problem, denoted by GNCP(F, G, $\mathcal{K}$), is to find a vector $x^* \in R^n$ such that

$$F(x^*) \in \mathcal{K}, \qquad G(x^*) \in \mathcal{K}^\circ, \qquad F(x^*)^\top G(x^*) = 0, \tag{1}$$

* This subject was supported by CityU Strategic Research Grant (Grant No. 7001427), the NSF of Shandong, China (Grant No. Y2003A02), and the Postdoctoral Fellowship of Hong Kong Polytechnic University.

where F and G are continuous functions from R^n to R^m, $\mathcal{K}$ is a nonempty closed convex cone in R^m, and $\mathcal{K}^\circ$ denotes the polar cone of $\mathcal{K}$.

This problem has many interesting applications and its solution using special techniques has been considered extensively in the literature. See [11], [14], [22], and references therein. In particular, if $\mathcal{K} = R^n_+$ and $G(x) = x$, then the GNCP reduces to the classical nonlinear complementarity problem [7]. Furthermore, the GNCP is closely related to the variational inequality problem in the sense that x^* is a solution of GNCP(F, G, $\mathcal{K}$) if and only if $F(x^*)$ is a solution of VI($G \circ F^{-1}$, $\mathcal{K}$) if F is invertible (see Lemma 6 in [1]).

To solve the GNCP, one usually reformulates it as a minimization problem over a simple set or an unconstrained optimization problem, see [22] for the case that $\mathcal{K}$ is a general cone and see [11] and [14] for the case that $\mathcal{K} = R^n_+$. The conditions under which a stationary point of the reformulated optimization is a solution of the GNCP were provided in this literature.

Now, we consider the case that $m = n$, F and G are both continuously differentiable on R^n, and $\mathcal{K}$ is a polyhedral cone in R^n, that is, there exist $A \in R^{s\times n}$, $B \in R^{t\times n}$ such that

$$\mathcal{K} = \{v \in R^n \mid Av \geq 0,\ Bv = 0\},$$

where s and t are both positive integers. It is easy to verify that its polar cone $\mathcal{K}^\circ$ has the following representation:

$$\mathcal{K}^\circ = \{u \in R^n \mid u = A^\top\lambda_1 + B^\top\lambda_2, \lambda_1 \geq 0, \lambda_1 \in R^s, \lambda_2 \in R^t\}.$$

Obviously, if A is an identity matrix and $B = 0$, then this version of the GNCP reduces to the case considered in [14].

From now on, the GNCP is specialized over a polyhedral cone.

For the GNCP, Andreani et al. reformulated it as a smooth simple constrained optimization problem in which the objective function preserves all derivatives of the functions that define the GNCP [1]. They also gave some sufficient conditions under which a stationary point of the optimization problem is a solution of the GNCP. However, since the second-order derivative, i.e, Hessian matrix, of the objective function becomes complicated due to its structure, it may be difficult to establish superlinear convergence of the algorithm.

To propose a superlinearly convergent algorithm for the solution of the GNCP, we now formulate the GNCP as a system of equations via the Fischer function [9] $\phi : R^2 \to R^1$ defined by

$$\phi(a, b) = \sqrt{a^2 + b^2} - a - b, \qquad \text{for} \quad a, b \in R.$$

A basic property of this function is that

$$\phi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0.$$

For arbitrary vectors $a, b \in R^n$, we define a vector-valued function as follows:

$$\Phi(a, b) = \begin{pmatrix} \phi(a_1, b_1) \\ \phi(a_2, b_2) \\ \cdots \\ \phi(a_n, b_n) \end{pmatrix}.$$

Obviously,

$$\Phi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, a^\top b = 0.$$

Before giving our reformulation of the GNCP, we first give some notations needed in this paper. The inner product of vectors $x, y \in R^n$ is denoted by $x^\top y$. Let $\|\cdot\|$ denote the 2-norm of vectors in Euclidean space. The transposed Jacobian $F'(x)$ of a vector-valued function $F(x)$ is denoted by $\nabla F(x)$. The nonnegative orthant of R^n is denoted by R^n_+.

Now, we give some equivalent statements relative to the solution of the GNCP.

x^* is a solution of the GNCP($F, G, \mathcal{K}$)

$$\Leftrightarrow \begin{cases} F(x^*) \in \mathcal{K} = \{v \in R^n \mid Av \geq 0, Bv = 0\}, \\ G(x^*) \in \mathcal{K}^\circ = \{u \in R^n \mid u = A^\top \lambda_1 + B^\top \lambda_2,\ \lambda_1 \in R^s_+, \lambda_2 \in R^t\}, \\ F(x^*)^\top G(x^*) = 0 \end{cases}$$

$$\Leftrightarrow \text{ there exist } \lambda_1^* \in R^s, \lambda_2^* \in R^t, \text{ such that } \begin{cases} AF(x^*) \geq 0, \\ BF(x^*) = 0, \\ G(x^*) = A^\top \lambda_1^* + B^\top \lambda_2^*, \\ \lambda_1^* \geq 0, \\ F(x^*)^\top G(x^*) = 0 \end{cases}$$

$$\Leftrightarrow \text{ there exist } \lambda_1^* \in R^s, \lambda_2^* \in R^t, \text{ such that } \begin{cases} AF(x^*) \geq 0, \\ \lambda_1^* \geq 0, \\ (\lambda_1^*)^\top AF(x^*) = 0, \\ BF(x^*) = 0, \\ G(x^*) = A^\top \lambda_1^* + B^\top \lambda_2^* \end{cases}$$

$$\Leftrightarrow \text{ there exist } \lambda_1^* \in R^s, \lambda_2^* \in R^t, \text{ such that } \begin{cases} \Phi(AF(x^*), \lambda_1^*) = 0, \\ BF(x^*) = 0, \\ G(x^*) - A^\top \lambda_1^* - B^\top \lambda_2^* = 0. \end{cases}$$

It is known that $G(x)$ is the gradient of a function $f : R^n \to R$ if $G'(x)$ is symmetric. So in this case, if $F(x) \equiv x$ in addition, from the last equivalent statement above, we can see that x^* is a solution of the GNCP if and only if it is a KKT point of the following linearly constrained optimization:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & Ax \geq 0, \\ & Bx = 0. \end{array}$$

Define a vector-valued function $\Psi : R^{n+s+t} \to R^{n+s+t}$ and a real-valued function $f : R^{n+s+t} \to R$ as follows:

$$\Psi(x, \lambda_1, \lambda_2) := \begin{pmatrix} \Phi(AF(x), \lambda_1) \\ BF(x) \\ G(x) - A^\top \lambda_1 - B^\top \lambda_2 \end{pmatrix}, \tag{2}$$

$$f(x, \lambda_1, \lambda_2) := \tfrac{1}{2}\Psi(x, \lambda_1, \lambda_2)^\top \Psi(x, \lambda_1, \lambda_2) = \tfrac{1}{2}\|\Psi(x, \lambda_1, \lambda_2)\|^2, \tag{3}$$

then the following result is straightforward.

Theorem 1.1. *x^* is a solution of the GNCP if and only if there exist $\lambda_1^* \in R^s$, $\lambda_2^* \in R^t$ such that $\Psi(x^*, \lambda_1^*, \lambda_2^*) = 0$.*

2. Preliminaries

In this section we review some definitions and basic results which will be used in what follows.

Definition 2.1.

(a) An $n \times n$ matrix M is a P_0-matrix if every principal minor of M is nonnegative.
(b) An $n \times n$ matrix M is a P-matrix if every principal minor of M is positive.

The properties and applications of a $P(P_0)$-matrix are given in [4].

For vector $a \in R^n$, $D_a = \operatorname{diag}(a)$ denotes the diagonal matrix in which the ith diagonal element is a_i. The following result is useful in the analysis of our algorithm, and its proof can be found in [15] and [17].

Lemma 2.1.

(a) *Let M be a P_0-matrix and let D_a and D_b be negative definite diagonal matrices in $R^{n\times n}$, then the matrix $D_a + D_b M$ is nonsingular.*
(b) *Let M be a P-matrix and let D_a and D_b be negative semidefinite diagonal matrices in $R^{n\times n}$ such that $D_a + D_b$ is negative definite, then the matrix $D_a + D_b M$ is nonsingular.*

The function $\Phi(AF(x), \lambda_1)$ is not differentiable everywhere with respect to $(x, \lambda_1) \in R^n \times R^s$ even though $F(x)$ is. However, it is locally Lipschitzian due to the continuity of F, and therefore has a nonempty generalized Jacobian in the sense of Clarke [3]. In the following, for a locally Lipschitzian mapping $\Theta : R^n \to R^m$, we let $\partial\Theta(x)$ denote Clarke's generalized Jacobian of $\Theta(x)$ at $x \in R^n$ which can be expressed as the convex hull of the set $\partial_B\Theta(x)$ [20], where

$$\partial_B\Theta(x) = \left\{ V \in R^{n\times n} \mid V = \lim_{x^k \to x} \Theta'(x^k),\ \Theta(x) \text{ is differentiable at } x^k \text{ for all } k \right\}.$$

Now, we recall some basic definitions about semismoothness and strong semismoothness. A locally Lipschitz continuous vector-valued function $\Theta : R^n \to R^m$ is said to be semismooth at $x \in R^n$, if the limit

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

exists for any $h \in R^n$. It is well known that the directional derivative, denoted by $\Theta'(x; h)$, of Θ at x in the direction h exists for any $h \in R^n$ if Θ is semismooth at x. The following properties about the semismooth function are due to Qi and Sun in [21].

Lemma 2.2. *Suppose that $\Theta : R^n \to R^m$ is a locally Lipschitz function and semismooth, then*

(a) *for any* $V \in \partial\Theta(x+h),\ h \to 0,$

$$Vh - \Theta'(x; h) = o(\|h\|);$$

(b) *for any* $h \to 0,$

$$\Theta(x+h) - \Theta(x) - \Theta'(x; h) = o(\|h\|).$$

Semismooth functions lie between Lipschitz functions and continuously differentiable functions, and both continuously differentiable functions and convex functions are semismooth. A stronger notion than semismoothness is strong semismoothness. The function $\Theta : R^n \to R^m$ is said to be strongly semismooth at x if Θ is semismooth at x and for any $V \in \partial\Theta(x+h),\ h \to 0$, it holds that

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

Now, we discuss the differential properties of the functions defined by (2) and (3). In particular, we present an overestimate of Clarke's generalized Jacobian of $\Phi(AF(x), \lambda_1)$. For simplicity, we denote Clarke's generalized Jacobian of $\Phi(AF(x), \lambda_1)$ with respect to $(x, \lambda_1) \in R^n \times R^s$ by $\partial\Phi(AF(x), \lambda_1)$. Similar to the discussion of Proposition 3.1 in [8], we have the following result.

Lemma 2.3. *For any $x \in R^n$ and $\lambda_1 \in R^s$, we have*

$$\partial\Phi(AF(x), \lambda_1) \subseteq \big(D_a AF'(x) + D_b\big),$$

where

$$a_i = \frac{[AF(x)]_i}{\sqrt{[\lambda_1]_i^2 + [AF(x)]_i^2}} - 1, \qquad b_i = \frac{[\lambda_1]_i}{\sqrt{[\lambda_1]_i^2 + [AF(x)]_i^2}} - 1$$

if

$$[\lambda_1]_i^2 + [AF(x)]_i^2 > 0$$

and

$$a_i = \xi_i - 1, \quad b_i = \eta_i - 1 \qquad \textit{for every} \quad (\xi_i, \eta_i) \in R^2 \quad \textit{such that} \quad \|(\xi_i, \eta_i)\| \le 1$$

if

$$[\lambda_1]_i^2 + [AF(x)]_i^2 = 0.$$

A favorable property of the function $f(x, \lambda_1, \lambda_2)$ is that it is continuously differentiable on the whole space R^{n+s+t} although $\Psi(x, \lambda_1, \lambda_2)$ is not in general. We summarize the differential properties of Ψ and f defined by (2) and (3) in the following lemma [8], [23].

Lemma 2.4. *For the vector-valued function* Ψ *and the real-valued function* f *defined by* (2) *and* (3), *the following statements hold*:

(a) *If F and G are both continuously differentiable*, *then* Ψ *is semismooth*, *and if* F' *and* G' *are both locally Lipschitzian in addition*, *then* Ψ *is strongly semismooth*.
(b) *If F and G are both continuously differentiable*, *then* f *is continuously differentiable*, *and its gradient at a point* $(x, \lambda_1, \lambda_2) \in R^n \times R^s \times R^t$ *is given by* $\nabla f(x, \lambda_1, \lambda_2) = V^\top \Psi(x, \lambda_1, \lambda_2)$, *where* V *is an arbitrary element belonging to* $\partial\Psi(x, \lambda_1, \lambda_2)$.

Finally, we give the definition of BD-regularity which plays a crucial role in the proof of the convergence rate of our algorithm in Section 4.

Definition 2.2. A function $\Theta : R^n \to R^n$ is said to be BD-regular at x if any $V \in \partial\Theta(x)$ is nonsingular.

The following result is an immediate consequence of z^* being a BD-regular solution to the semismooth equation $\Theta(z) = 0$ [12], [16], [19], [21], [23].

Lemma 2.5. *Suppose that* $\Theta : R^n \to R^n$ *is semismooth and* $z^* \in R^n$ *is a solution of* $\Theta(z) = 0$. *Then for sufficiently small* $\varepsilon > 0$, *there exists a constant* $c_1 > 0$ *such that*

$$\|\Theta(z)\| \le c_1\|z - z^*\|, \qquad \text{for } z \text{ with} \quad \|z - z^*\| \le \varepsilon.$$

Moreover, *if* Θ *is BD-regular at* z^*, *then there exists a constant* $c_2 > 0$ *such that the matrices* $V \in \partial\Theta(z)$ *are nonsingular and*

$$\|V^{-1}\| \le c_2, \qquad \text{for } z \text{ with} \quad \|z - z^*\| \le \varepsilon.$$

3. Existence and Uniqueness of the Solutions

In this section we discuss the existence and uniqueness of the solutions to GNCP $(F, G, \mathcal{K})$. First, we prove that GNCP$(F, G, \mathcal{K})$ has at most one solution if the mappings F and G satisfy the following condition: there exists a constant $\mu > 0$ such that

$$(F(x) - F(y))^\top (G(x) - G(y)) \ge \mu\|x - y\|^2, \qquad \forall\, x, y \in R^n. \tag{4}$$

Lemma 3.1. *Assume that the mappings* $F, G : R^n \to R^n$ *satisfy condition* (4) *for some* $\mu > 0$, *then* GNCP$(F, G, \mathcal{K})$ *has at most one solution*.

Proof. Suppose that x^* and y^* are two different solutions of GNCP(F, G, $\mathcal{K}$), then there exist nonnegative vectors $\lambda_1^{x^*}, \lambda_1^{y^*} \in R^s$ and vectors $\lambda_2^{x^*}, \lambda_2^{y^*} \in R^t$ such that

$$G(x^*) = A^\top \lambda_1^{x^*} + B^\top \lambda_2^{x^*}, \qquad G(y^*) = A^\top \lambda_1^{y^*} + B^\top \lambda_2^{y^*}.$$

By (4) and the definition of GNCP(F, G, $\mathcal{K}$), we have

$$\begin{aligned}
0 &< \mu \|x^* - y^*\|^2 \\
&\leq (F(x^*) - F(y^*))^\top (G(x^*) - G(y^*)) \\
&= -[F(x^*)]^\top G(y^*) - [F(y^*)]^\top G(x^*) \\
&= -[F(x^*)]^\top \big(A^\top \lambda_1^{y^*} + B^\top \lambda_2^{y^*}\big) - [F(y^*)]^\top \big(A^\top \lambda_1^{x^*} + B^\top \lambda_2^{x^*}\big) \\
&= -[AF(x^*)]^\top \lambda_1^{y^*} - [AF(y^*)]^\top \lambda_1^{x^*} \\
&\leq 0.
\end{aligned}$$

This contradiction implies that GNCP(F, G, $\mathcal{K}$) has at most one solution. $\square$

As noted in Section 1, GNCP(F, G, $\mathcal{K}$) is equivalent to the variational inequality problem VI($G \circ F^{-1}$, $\mathcal{K}$) if the mapping F is invertible. The following result is concerned with the existence and uniqueness of the solutions to VI(H, $\mathcal{K}$) (see Corollary 3.2 in [7]).

Theorem 3.1. *Suppose the continuous mapping $H : R^n \to R^n$ is strongly monotone, i.e., there exists a constant $\rho > 0$ such that*

$$(H(x) - H(y))^\top (x - y) \geq \rho \|x - y\|^2, \qquad \forall x, y \in R^n.$$

Then VI(H, $\mathcal{K}$) *has a unique solution*.

Using a similar way to the proof of Lemma 5.2 in [14], we can prove the following result.

Lemma 3.2. *Assume that the mappings F, G satisfy condition* (4) *for some $\mu > 0$ and F is invertible and Lipschitz continuous on R^n. Then the mapping $G \circ F^{-1}$ is strongly monotone*.

Using Lemma 3.2 and Theorem 3.1, we can easily obtain the existence and uniqueness of solutions to GNCP(F, G, $\mathcal{K}$).

Theorem 3.2. *Assume that the conditions in Lemma* 3.2 *hold and the mapping F^{-1} is continuous in addition. Then* GNCP(F, G, $\mathcal{K}$) *has a unique solution*.

4. Stationary Point and Nonsingularity Conditions

From Theorem 1.1, we know that a point x^* is a solution of the GNCP if and only if there exist $\lambda_1^* \in R^s$, $\lambda_2^* \in R^t$ such that $z^* = (x^*, \lambda_1^*, \lambda_2^*)$ solves the following system of equations:

$$\Psi(z) = 0,$$

or, equivalently, z^* is a global minimizer with a zero objective function value of the unconstrained optimization problem

$$\min_{z \in R^{n+s+t}} f(z). \tag{5}$$

Since most unconstrained minimization methods always generate a sequence converging to a local minimizer or a stationary point rather than a global minimizer, it is therefore crucial to study the conditions under which a stationary point of (5) is its global minimizer with the objective value zero. The following theorem gives a suitable condition which guarantees that every stationary point of (5) solves GNCP$(F, G, \mathcal{K})$.

Theorem 4.1. *Suppose that $z^* = (x^*, \lambda_1^*, \lambda_2^*)$ is a stationary point of* (5), *$F'(x^*)$ is nonsingular, and $G'(x^*)[F'(x^*)]^{-1}$ is positive definite in the null space of B, then x^* is a solution of* GNCP$(F, G, \mathcal{K})$.

Proof. Let

$$\begin{aligned}
U^* &= \Phi(AF(x^*), \lambda_1^*),\\
V^* &= BF(x^*),\\
W^* &= G(x^*) - A^\top\lambda_1^* - B^\top\lambda_2^*,\\
M^* &= G'(x^*)[F'(x^*)]^{-1}.
\end{aligned}$$

Since z^* is a stationary point of (5),

$$\nabla f(z^*) = 0,$$

which, according to Lemmas 2.3 and 2.4, implies that

$$V^\top \Psi(z^*) = 0,$$

where

$$V^\top = \begin{pmatrix} \nabla F(x^*)A^\top D_a & \nabla F(x^*)B^\top & \nabla G(x^*)\\ D_b & 0 & -A\\ 0 & 0 & -B \end{pmatrix},$$

and both D_a and D_b are diagonal matrices given in Lemma 2.3. Thus

$$\nabla F(x^*)A^\top D_a U^* + \nabla F(x^*)B^\top V^* + \nabla G(x^*)W^* = 0, \tag{6}$$

$$D_b U^* - A W^* = 0, \tag{7}$$

$$-B W^* = 0. \tag{8}$$

Since $F'(x^*)$ is nonsingular, from (6) and (7), it holds that

$$A^\top D_a U^* + B^\top V^* + (M^*)^\top W^* = 0, \tag{9}$$

$$A W^* = D_b U^*. \tag{10}$$

Combining (8)–(10) yields

$$\begin{aligned} 0 &= \langle W^*, A^\top D_a U^* \rangle + \langle W^*, B^\top V^* \rangle + \langle W^*, (M^*)^\top W^* \rangle \\ &= \langle A W^*, D_a U^* \rangle + \langle B W^*, V^* \rangle + \langle W^*, (M^*)^\top W^* \rangle \\ &= \langle D_b U^*, D_a U^* \rangle + \langle W^*, (M^*)^\top W^* \rangle \\ &= (U^*)^\top D_b D_a U^* + \langle W^*, (M^*)^\top W^* \rangle. \end{aligned}$$

Since W^* belongs to the null space of B, $\mathcal{N}(B)$, using the positive semidefiniteness of $D_a D_b$ from Lemma 2.3 and the positive definiteness of M^* in $\mathcal{N}(B)$, we have

$$W^* = 0. \tag{11}$$

Substituting this result into (9) and (10), we have

$$A^\top D_a U^* + B^\top V^* = 0, \tag{12}$$

$$D_b U^* = 0. \tag{13}$$

Premultiplying (12) by $F(x^*)^\top$ yields

$$(A F(x^*))^\top D_a U^* + \| B F(x^*) \|^2 = 0. \tag{14}$$

Now, we assert that

$$U_i^* = 0, \qquad i = 1, 2, \ldots, s.$$

Otherwise, there exists an index $1 \le i_0 \le s$ such that

$$U_{i_0}^* \neq 0. \tag{15}$$

From (13), we can deduce that $(D_b)_{i_0} = 0$, and hence from Lemma 2.3, $(D_a)_{i_0} = -1$ and

$$[\lambda_1^*]_{i_0} \ge 0, \qquad [A F(x^*)]_{i_0} = 0.$$

Thus $U_{i_0}^* = 0$, which contradicts (15). Therefore,

$$U^* = 0. \tag{16}$$

By (14), we have

$$V^* = B F(x^*) = 0. \tag{17}$$

Finally, (11), (16), and (17) imply that

$$\Psi(x^*, \lambda_1^*, \lambda_2^*) = 0. \qquad \square$$

From this theorem, the condition under which a stationary point of (5) is a solution of the GNCP is the same as that of the reformulation given by Andreani et al. in [1].

To establish a superlinear (quadratic) convergence rate of our algorithm, it is necessary to study the conditions under which every element of the generalized Jacobian $\partial\Psi(z)$ is nonsingular at a solution point z^* of the equation $\Psi(z) = 0$.

Theorem 4.2. *For $z^* = (x^*, \lambda_1^*, \lambda_2^*) \in R^n \times R^s \times R^t$, if $F'(x^*)$ and $G'(x^*)$ are nonsingular, matrix B has full row rank, and $AF'(x^*)[G'(x^*)]^{-1}A^\top$ is a P-matrix, then for any $V \in \partial\Psi(z^*)$, V is nonsingular.*

Proof. From Lemma 2.3 we know that for any $V \in \partial\Psi(z^*)$, it has the following form:

$$V^\top = \begin{pmatrix} \nabla F(x^*)A^\top D_a & \nabla F(x^*)B^\top & \nabla G(x^*) \\ D_b & 0 & -A \\ 0 & 0 & -B \end{pmatrix},$$

where D_a and D_b are defined in Lemma 2.3.

We denote $D = V^\top$ and $M = [G'(x^*)[F'(x^*)]^{-1}]^\top$.

In what follows, we prove the nonsingularity of D through a system of transformations without changing its rank.

Since $F'(x^*)$ is nonsingular, we make a transformation to D as follows:

$$\begin{pmatrix} [\nabla F(x^*)]^{-1} & 0 & 0 \\ 0 & I_s & 0 \\ 0 & 0 & I_t \end{pmatrix} D = \begin{pmatrix} A^\top D_a & B^\top & M \\ D_b & 0 & -A \\ 0 & 0 & -B \end{pmatrix} := D_1.$$

We partition B as $B = \begin{pmatrix} B_1 & B_2 \end{pmatrix}$ such that B_2 is a square matrix. Since B has full row rank, without loss of generality, we can assume that B_2 is nonsingular. Now A has the corresponding partition $A = \begin{pmatrix} A_1 & A_2 \end{pmatrix}$ and D_1 has the following corresponding partition:

$$D_1 = \begin{pmatrix} A_1^\top D_a & B_1^\top & M_{11} & M_{12} \\ A_2^\top D_a & B_2^\top & M_{21} & M_{22} \\ D_b & 0 & -A_1 & -A_2 \\ 0 & 0 & -B_1 & -B_2 \end{pmatrix},$$

where

$$A_1 \in R^{s\times(n-t)}, \quad A_2 \in R^{s\times t}, \quad B_1 \in R^{t\times(n-t)}, \quad B_2 \in R^{t\times t}, \quad M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix},$$
$$M_{11} \in R^{(n-s)\times(n-s)}, \quad M_{12} \in R^{(n-s)\times s}, \quad M_{21} \in R^{t\times(n-t)}, \quad M_{22} \in R^{t\times t}.$$

Since B_2 is nonsingular, we can make a column transformation to D_1 and obtain

$$\begin{pmatrix} A_1^\top D_a & B_1^\top & M_{11} - M_{12}B_2^{-1}B_1 & M_{12} \\ A_2^\top D_a & B_2^\top & M_{21} - M_{22}B_2^{-1}B_1 & M_{22} \\ D_b & 0 & -A_1 + A_2B_2^{-1}B_1 & -A_2 \\ 0 & 0 & 0 & -B_2 \end{pmatrix}.$$

Thus, to prove that D is nonsingular, it suffices to show the nonsingularity of the principal submatrix of the matrix above by deleting the last t rows and columns, which we denote by D_2, i.e.,

$$D_2 = \begin{pmatrix} A_1^\top D_a & B_1^\top & M_{11} - M_{12}B_2^{-1}B_1 \\ A_2^\top D_a & B_2^\top & M_{21} - M_{22}B_2^{-1}B_1 \\ D_b & 0 & -A_1 + A_2B_2^{-1}B_1 \end{pmatrix}.$$

Making a simple row transformation to D_2 leads to the following matrix:

$$\begin{pmatrix} A_1^\top D_a - (B_2^{-1}B_1)^\top A_2^\top D_a & 0 & M_{11} - M_{12}B_2^{-1}B_1 - (B_2^{-1}B_1)^\top (M_{21} - M_{22}B_2^{-1}B_1) \\ A_2^\top D_a & B_2^\top & M_{21} - M_{22}B_2^{-1}B_1 \\ D_b & 0 & -A_1 + A_2B_2^{-1}B_1 \end{pmatrix}.$$

Hence, to prove the nonsingularity of D, it suffices to show the nonsingularity of the following matrix:

$$D_3 := \begin{pmatrix} A_1^\top D_a - (B_2^{-1}B_1)^\top A_2^\top D_a & M_{11} - M_{12}B_2^{-1}B_1 - (B_2^{-1}B_1)^\top (M_{21} - M_{22}B_2^{-1}B_1) \\ D_b & -A_1 + A_2B_2^{-1}B_1 \end{pmatrix}.$$

Now, we make a decomposition of matrix D_3.

$$D_3 = \begin{pmatrix} \begin{pmatrix} I_{n-t} & -(B_2^{-1}B_1)^\top \end{pmatrix} \begin{pmatrix} A_1^\top D_a \\ A_2^\top D_a \end{pmatrix} & \begin{pmatrix} I_{n-t} & -(B_2^{-1}B_1)^\top \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} I_{n-t} \\ -(B_2^{-1}B_1) \end{pmatrix} \\ D_b & -\begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} I_{n-t} \\ -(B_2^{-1}B_1) \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} I_{n-t} & -(B_2^{-1}B_1)^\top \end{pmatrix} & 0 \\ 0 & I_s \end{pmatrix} \begin{pmatrix} A^\top D_a & M \\ D_b & -A \end{pmatrix} \begin{pmatrix} I_s & 0 \\ 0 & \begin{pmatrix} I_{n-t} \\ -(B_2^{-1}B_1) \end{pmatrix} \end{pmatrix}.$$

Since

$$\begin{pmatrix} \begin{pmatrix} I_{n-t} & -(B_2^{-1}B_1)^\top \end{pmatrix} & 0 \\ 0 & I_s \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} I_s & 0 \\ 0 & \begin{pmatrix} I_{n-t} \\ -(B_2^{-1}B_1) \end{pmatrix} \end{pmatrix}$$

have full row rank and full column rank, respectively, we only need to show that

$$\begin{pmatrix} A^\top D_a & M \\ D_b & -A \end{pmatrix}$$

is nonsingular. Since M is nonsingular, a row transformation to the matrix above yields that

$$\begin{pmatrix} A^\top D_a & M \\ D_b & -A \end{pmatrix} \longrightarrow \begin{pmatrix} A^\top D_a & M \\ D_b + AM^{-1}A^\top D_a & 0 \end{pmatrix}.$$

Since $AM^{-1}A^\top$ is a P-matrix, by Lemma 2.2,

$$D_b + AM^{-1}A^\top D_a$$

is nonsingular, and thus the matrix $\begin{pmatrix} A^\top D_a & M \\ D_b & -A \end{pmatrix}$ is nonsingular. Hence, V is nonsingular and the desired result follows. □

If $z^* = (x^*, \lambda_1^*, \lambda_2^*)$ is a stationary point of (5), i.e.,

$$0 = \nabla f(z^*) = V^\top \Psi(z^*),$$

where $V \in \partial\Psi(z^*)$, then x^* is a solution of GNCP$(F, G, \mathcal{K})$ under the condition of Theorem 4.2. Thus, the following conclusion holds.

Corollary 4.1. *If $z^* = (x^*, \lambda_1^*, \lambda_2^*)$ is a stationary point of* (5), *if either of the conditions given in Theorems* 4.1 *or* 4.2 *holds*, *then x^* is a solution of* GNCP$(F, G, \mathcal{K})$.

If $B = 0$, then the optimization problem (5) becomes

$$\min_{x,\lambda} \ \frac{1}{2}\Big(\|\Phi(AF(x), \lambda)\|^2 + \|G(x) - A^\top\lambda\|^2\Big). \tag{18}$$

In this case, from Theorem 4.2, we have the following result.

Corollary 4.2. *If (x^*, λ^*) is a stationary point of* (18), *$G'(x^*)$ is nonsingular*, *and the matrix $AF'(x^*)[G'(x^*)]^{-1}A^\top$ is a P-matrix*, *then for any $V \in \partial\bar{\Psi}(x^*, \lambda^*)$, V is nonsingular*, *where*

$$\bar{\Psi}(x, \lambda) = \begin{pmatrix} \Phi(AF(x), \lambda) \\ G(x) - A^\top\lambda \end{pmatrix}.$$

To guarantee the superlinear (quadratic) convergence of the L-M method for solving a system of equations $H(x) = 0$, Yamashita and Fukushima provided a weaker condition than nonsingularity of the Jacobian recently [24], which requires that the local error bound of $H(x)$ holds near a solution point x^*, i.e., there exists a positive scalar c such that

$$\|H(x)\| \geq c \cdot \mathrm{d}(x, S), \qquad \forall x \in N(x^*, \varepsilon),$$

where S is the solution set of $H(x) = 0$, $N(x^*, \varepsilon)$ is a neighborhood of the solution x^*, and $\mathrm{d}(x, S)$ is the distance from x to S.

For the linear complementarity problem, one reformulated system of equations can provide a local error bound near its solution set [10]. It is uncertain whether the function $\Psi(z)$ can provide a local error bound near the solution set of the affine GNCP$(F, G, \mathcal{K})$ such that F and G are both affine, this is a topic for further research.

5. Algorithm and Convergence

In this section an L-M method for solving the GNCP is outlined. It is similar to that in [6]. For convenience, let $z^k = (x^k, \lambda_1^k, \lambda_2^k)$ in what follows.

Algorithm 5.1

Step 1: Choose any point $z^0 \in R^{n+s+t}$, parameters $\sigma, \beta \in (0, 1)$ and $\varepsilon \geq 0$. Let $k = 0$.

Step 2: If $\|\nabla f(z^k)\| \leq \varepsilon$, stop; otherwise, go to Step 3.

Step 3: Choose an element $V^k \in \partial\Psi(z^k)$. Let $d^k \in R^{n+s+t}$ be the solution of the linear system

$$((V^k)^\top V^k + \mu^k I)d = -(V^k)^\top \Psi(z^k),$$

where $\mu^k = \|\Psi(z^k)\|$.

Step 4: Let m_k be the smallest nonnegative integer m such that

$$f(z^k + \sigma^m d^k) \leq f(z^k) + \beta\sigma^m \nabla f(z^k)^\top d^k.$$

Let $z^{k+1} := z^k + \sigma^{m_k} d^k$, $k := k + 1$, go to Step 2.

It is easy to verify that d_k is a descent direction of $f(z)$ at z_k and the algorithm is well defined. Obviously, if $\nabla f(z^k) = 0$, then z^k is a stationary point of problem (5), and thus x^k is a solution of GNCP($F, G, \mathcal{K}$) under suitable conditions. In the following convergence analysis, we assume that $\varepsilon = 0$ and Algorithm 5.1 generates an infinite sequence. Following the proof in [13], we can obtain the convergence and superlinear convergence of Algorithm 5.1.

Theorem 5.1. *Any accumulation point of the sequence* $\{z^k\}$ *generated by Algorithm* 5.1 *is a stationary point of* (5).

Theorem 5.2. *Let* $\{z^k\}$ *be the sequence generated by Algorithm* 5.1. *Assume that* z^* *is an accumulation point of* $\{z^k\}$ *and a* BD*-regular solution of* $\Psi(z) = 0$, *then*

(a) *the entire sequence* $\{z^k\}$ *superlinearly converges to* z^*, *and*
(b) $\{z^k\}$ *converges to* z^* *Q-quadratically if* F' *and* G' *are both Lipschitzian in addition*.

From Theorem 2.1 in [24], we know that for a system of equations $H(x) = 0$, where H is a twice continuously differentiable mapping from R^n to R^n, if $H(x)$ provides a local error bound near a solution of $H(x) = 0$, then the sequence generated by the L-M method converges locally and quadratically [24].

Now, consider the system of equations $\Psi(z) = 0$ defined in (2). If one of its solutions z^* is nondegenerate, i.e., $(\lambda_1^*)_i + (AF(x^*))_i > 0$ for $i = 1, 2, \ldots, s$, and F and G are both twice continuously differentiable, then $\Psi(z)$ is twice continuously differentiable near z^*. Thus, from Theorem 2.1 in [24], we have the following convergence result of Algorithm 5.1.

Theorem 5.3. *Suppose $F(x)$ and $G(x)$ are both twice continuously differentiable, if an accumulation point z^* of the sequence $\{z_k\}$ generated by Algorithm 5.1 is nondegenerate, the local error bound of $\Psi(z)$ holds near z^*, and the condition in Theorem 4.1 holds at z^*, then the sequence converges to z^*, a solution of $\Psi(z) = 0$, quadratically.*

6. Computational Experiments

Before making our computational experiments, we should find a way to calculate an element of $\partial\Phi(AF(x), \lambda_1)$. The following theorem gives an approach to calculate an element of $\partial\Phi(AF(x), \lambda_1)$, and its proof can be referred to Theorem 27 of [17].

Lemma 6.1. *For $x \in R^n$ and $\lambda_1 \in R^s$, choose $v \in R^s$ such that $v_i \neq 0$ for any index i with $[\lambda_1]_i = 0$ and $[AF(x)]_i = 0$.*

Let

$$W = D_a AF'(x) + D_b,$$

where

$$a_i = \frac{[AF(x)]_i}{\sqrt{[\lambda_1]_i^2 + [AF(x)]_i^2}} - 1, \qquad b_i = \frac{[\lambda_1]_i}{\sqrt{[\lambda_1]_i^2 + [AF(x)]_i^2}} - 1$$

if

$$[\lambda_1]_i^2 + [AF(x)]_i^2 > 0,$$

and

$$a_i = \frac{[AF'(x)x]_i}{\sqrt{v_i^2 + [AF'(x)x]_i^2}} - 1, \qquad b_i = \frac{v_i}{\sqrt{v_i^2 + [AF'(x)x]_i^2}} - 1$$

if

$$[\lambda_1]_i^2 + [AF(x)]_i^2 = 0.$$

Then

$$W \in \partial\Phi(AF(x), \lambda_1),$$

or more precisely,

$$W \in \partial_B\Phi(AF(x), \lambda_1).$$

Changing v, we will obtain a different element of $\partial_B\Phi(AF(x), \lambda_1)$. In our code, we choose to set $v_i = 0$ if $[\lambda_1]_i^2 + [AF(x)]_i^2 > 0$ and $v_i = 1$ otherwise. Thus, an element

$V \in \partial\Psi(z)$ can be calculated as

$$V^\top = \begin{pmatrix} \nabla F(x) A^\top D_a & \nabla F(x) B^\top & \nabla G(x) \\ D_b & 0 & -A \\ 0 & 0 & -B \end{pmatrix},$$

where D_a and D_b are defined in Lemma 6.1.

In the following we implement Algorithm 5.1 in Matlab and run it on a Pentium IV computer. Here, we do not give a comparison with the method given by Andreani in [1], since the efficiency of their method mainly depends on the efficiency of the method for solving smooth box-constrained minimization problem.

Throughout our computation, we take parameters $\sigma = 0.6$, $\beta = 0.4$, and terminate our computation whenever $\|\nabla f(z)\| \le 10^{-14}$.

First, we consider the example given in [18] which was also considered by Jiang et al. in [11]. For completeness, we give the example in detail.

Example 6.1. Find $x^* \in R^n$ such that

$$F(x^*) \ge 0, \qquad x^* - m(x^*) \ge 0, \quad \text{and} \quad \langle F(x^*), x^* - m(x^*)\rangle = 0,$$

where

$$F(x) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

and $m(x) = \varphi(F(x)) : R^n \to R^n$ is twice continuously differentiable.

Table 1 lists the numerical results of this example with different starting points for the following two choices of function $\varphi(\cdot)$:

(1) $\varphi_i(t) = -0.5 - t_i,\ i = 1, 2, 3, 4;$
(2) $\varphi_i(t) = -0.5 - 1.5t_i + 0.25t_i^2,\ i = 1, 2, 3, 4.$

In Table 1, **Iter** denotes the number of iterations, which is equal to the number of Jacobian evaluations for the functions F and G, and is also equal to the number of linear

Table 1. Numerical results for Example 6.1.

φ	Starting point	**Iter**	f^*	**NF**	**CT**
(1)	$(0, \dots, 0)^\top$	10	3.2×10^{-30}	19	0.06
(2)	$(0, \dots, 0)^\top$	11	6.2×10^{-32}	21	0.05
(1)	$(-0.5, \dots, -0.5)^\top$	8	1.8×10^{-32}	15	0.05
(2)	$(-0.5, \dots, -0.5)^\top$	8	1.8×10^{-31}	15	0.05
(1)	$(-1, \dots, -1)^\top$	9	2.6×10^{-32}	17	0.06
(2)	$(-1, \dots, -1)^\top$	9	4.6×10^{-32}	17	0.11
(1)	$(0.5, \dots, 0.5)^\top$	15	2.6×10^{-32}	29	0.11
(2)	$(0.5, \dots, 0.5)^\top$	16	6.6×10^{-32}	31	0.11

systems solved. **NF** represents the number of evaluations for the function f, f^* is the final value of f when the algorithm terminates, and **CT** denotes the computing time.

For this problem, Jiang et al. also gave an encouraging numerical experiment using the trust region method in [11].

Our next numerical experiment is about the following two sets of problems constructed by Andreani et al. in [1] and [2]. For simplicity of description, we make a slight modification.

Example 6.2. Consider the problem of finding $x^* \in R^n$ such that

$$\begin{cases} x \in \mathcal{K} = \{v \in R^n \mid Av \geq 0\}, \\ Nx + d \in \mathcal{K}^\circ = \{v \in R^n \mid v = A^\top \lambda,\ \lambda \in R^s_+\}, \\ x^\top (Nx + d) = 0, \end{cases}$$

where the polyhedral cone $\mathcal{K}$ is generated by s faces whose edges are the following lines:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\left(\dfrac{2\pi}{s} i\right) \\ r \sin\left(\dfrac{2\pi}{s} i\right) \\ 1 \end{pmatrix} \tau, \qquad \tau \in R, \quad i = 1, 2, \ldots, s.$$

Thus, the ith row of matrix $A \in R^{s\times 3}$ can be computed as

$$\begin{pmatrix} \sin\left(\dfrac{2\pi}{s} i\right)\left(\cos\dfrac{2\pi}{s} - 1\right) - \cos\left(\dfrac{2\pi}{s} i\right)\sin\dfrac{2\pi}{s} \\ \cos\left(\dfrac{2\pi}{s} i\right)\left(1 - \cos\dfrac{2\pi}{s}\right) - \sin\left(\dfrac{2\pi}{s} i\right)\sin\dfrac{2\pi}{s} \\ r \sin\dfrac{2\pi}{s} \end{pmatrix}^\top.$$

For each family, we choose $r \in \{0.1, 1, 10\}$ and $s \in \{3, 5, 9, 12\}$. The vector d is generated randomly from the interval $(-10, 10)$. Matrix N is generated as follows. Denote the orthogonal Householder matrix $Q_{(\cdot)} = I - 2(u_{(\cdot)} u_{(\cdot)}^\top / \|u_{(\cdot)}\|^2)$, where the components of vector $u_{(\cdot)}$ are generated randomly from $(-1, 1)$. Let D_N be the diagonal matrices whose diagonal elements are generated randomly from $(1, 10)$. We define matrix $N = Q_{NL} D_N Q_{NR}$.

For this problem, we divide the set of test problems into three families:

(1) N is nonsymmetric and indefinite;
(2) N is symmetric and positive definite;
(3) N is symmetric and positive semidefinite.

Obviously, the problems in Families (1) and (3) do not satisfy the hypothesis of Theorem 4.1. For each family with different r and s, 20 problems are tested with $z = (0, \ldots, 0)^\top$ being the starting point. The numerical results are reported in Table 2.

To take into account the possibility of convergence to a stationary point of $f(z)$ which is not a solution of the GNCP, we call a case successful if the value of f is

Table 2. Average numerical results for Example 6.2.

s	Family	r	**Iter**	**SP**	**NF**
3	(1)	0.1	187.67	0.3	374.30
		1	50.92	0.55	100.82
		10	78.8	0.5	157.4
	(2)	0.1	399.92	0.6	799.25
		1	68.35	1	135.70
		10	28.47	0.95	57
	(3)	0.1	319.30	0.5	637.60
		1	77.63	0.8	154.25
		10	98.8	0.5	197
5	(1)	0.1	442	0.55	883
		1	122.33	0.45	243.67
		10	76.67	0.30	152.33
	(2)	0.1	344.83	0.30	690.33
		1	29.5	1	58
		10	63.4	1	129.05
	(3)	0.1	497.36	0.55	993.82
		1	76.59	0.85	152.18
		10	40.75	0.40	82.50
9	(1)	0.1	273.43	0.35	545.86
		1	208.78	0.45	417.22
		10	134.67	0.45	269.22
	(2)	0.1	397.25	0.20	793.5
		1	221.33	0.9	441.72
		10	28.4	1	56.6
	(3)	0.1	382.83	0.3	764.67
		1	234.87	0.75	469.07
		10	215.46	0.65	431.23
12	(1)	0.1	417.89	0.45	834.78
		1	262.5	0.3	554.31
		10	86.80	0.5	173.8
	(2)	0.1	523.75	0.2	1046
		1	106.50	1	213.55
		10	67.1	1	134.7
	(3)	0.1	291.33	0.15	581.67
		1	221.70	0.50	442.80
		10	71.23	0.65	143.85

less than 10^{-10} within 1000 iterations and we denote by **SP** the successful rate. For all successful cases, **Iter** denotes the average number of iterations, and **NF** denotes the average number of evaluations for the function f. The numerical results are reported in Table 2, from which we can see that except for the case that $r = 0.1$, Algorithm 5.1 performs well for this set of problems.

Example 6.3. Consider the GNCP of finding $x^* \in R^n$ such that

$$Mx + c \in \mathcal{K}, \qquad Nx + d \in \mathcal{K}^\circ, \quad \text{and} \quad (Mx + c)^\top (Nx + d) = 0,$$

where

$$\mathcal{K} = \{v \in R^n \mid Av \geq 0,\ Bv = 0\},$$
$$\mathcal{K}^\circ = \{u \in R^n \mid u = A^\top \lambda_1 + B^\top \lambda_2, \lambda_1 \in R^s_+,\ \lambda_2 \in R^t\},$$

with $M, N \in R^{n\times n}$, $A \in R^{s\times n}$, $B \in R^{t\times n}$, and $c, d \in R^n$.

The matrices M, N, A and vectors c, d will be generated following the pattern in Example 6.2. Let the diagonal elements of D_M and D_N be generated randomly from $(-5, 5)$, and let $D_A \in R^{s\times s}$ be the diagonal matrix whose diagonal elements are generated randomly from $(-10, 10)$. We define matrices M, N, and A as follows:

$$M = Q_{ML} D_M Q_{MR}, \qquad N = Q_{NL} D_N Q_{NR}, \qquad A = Q_{AL} (D_A \quad \mathbf{0}_{s\times(n-s)}) Q_{AR}.$$

Let the components of vectors $c, d \in R^n$ be generated randomly from $(-10, 10)$. To complete the problem data, we generate matrix B as follows: take $y = Md + c$ and make a QR factorization to the matrix

$$\left(y \begin{pmatrix} I_t \\ \mathbf{0}_{(n-t)\times t} \end{pmatrix} \right),$$

and form matrix $B^\top$ by the last t columns of the Q factor.

For this problem, we divide the set of test problems into three families:

(1) $M = N$, indefinite and nonsymmetric;
(2) $M = N$, indefinite and symmetric;
(3) $M \neq N$, indefinite, nonsymmetric.

To generate a singular matrix, we force 20% of the diagonal elements of the diagonal matrix $D_{(\cdot)}$ to be zero. To generate an indefinite matrix, each diagonal element of $D_{(\cdot)}$ is multiplied by the sign of a random number.

For each family, six sets for the dimensions (n, s, t) are considered:

$$(5, 5, 1),\ (10, 10, 1),\ (10, 15, 1),\ (10, 5, 5),\ (10, 10, 5),\ (10, 15, 5).$$

For each set of dimensions, 20 problems are generated with different seeds. The detailed numerical experiments are reported in Table 3. From Table 3, we see that the algorithm performs well except for Family (3).

7. Discussion

In this paper we formulate the GNCP defined on a polyhedral cone as an unconstrained optimization problem and a globally and superlinearly convergent L-M method was proposed to solve the transformed optimization problem. Some encouraging numerical results are also reported in this paper. Certainly, there exists an accumulation point of the

Table 3. Average numerical results for Example 6.3.

(n, s, t)	Family	**Iter**	**SP**	**NF**
(5, 5, 1)	(1)	157.47	0.85	319.76
	(2)	194.67	0.75	390.40
	(3)	402.00	0.35	866.71
(10, 10, 1)	(1)	387.08	0.75	840.87
	(2)	228.21	0.70	506.50
	(3)	113.00	0.05	225.00
(10, 15, 1)	(1)	294.92	0.60	588.83
	(2)	400.45	0.55	817.00
	(3)	854.00	0.10	1759
(10, 5, 5)	(1)	228.64	0.55	463.45
	(2)	354.62	0.65	783.87
	(3)	721.00	0.15	1442.3
(10, 10, 5)	(1)	498.56	0.45	1185.8
	(2)	513.5	0.60	1015.3
	(3)	538.33	0.15	1323.00
(10, 15, 5)	(1)	248.08	0.60	495.83
	(2)	486.22	0.45	988.11
	(3)	493.50	0.20	1092.9

generated sequence if the level set $L(z_0) = \{z \mid f(z) \leq f(z^0)\}$ for some $z^0 \in R^{n+s+t}$ is bounded. For this issue, we consider a special case such that $B = 0$ and A is a symmetric and nonsingular matrix. Then the function defined by (2) becomes

$$\Psi(x, \lambda_1) = \begin{pmatrix} \Phi(AF(x), \lambda_1) \\ G(x) - A\lambda_1 \end{pmatrix}.$$

Thus, the function defined in (3) can be written as

$$f(x) = \tfrac{1}{2}\|\Phi(AF(x), A^{-1}G(x))\|^2.$$

For this function, we have a result which is similar to Theorem 6.2 in [14].

Theorem 7.1. *Assume that mappings $F, G : R^n \to R^n$ are Lipschitz continuous and satisfy* (4) *for some $\mu > 0$. Then the level set $L(x_0) = \{x \in R^n \mid f(x) \leq f(x_0)\}$ is bounded for all $x_0 \in R^n$.*

Obviously, this stronger condition confines the wide applications of Algorithm 5.1, and thus under what conditions the level set $L(z_0) = \{z \mid f(z) \leq f(z^0)\}$ is bounded is an interesting topic for further research.

Acknowledgments

The first author gives grateful thanks to Prof. Liqun Qi and Dr. Guanglu Zhou for their useful suggestions and comments.

References

1. R. Andreani, A. Friedlander, and S.A. Santos, On the resolution of the generalized nonlinear complementarity problem, SIAM J. Optim., 2001, 12:303–321.
2. R. Andreani, A. Friedlander, and S.A. Santos, Solving Generalized Nonlinear Complementarity Problems: Numerical Experiments on Polyhedral Cones, Technical Report, IMECC, State University of Campinas, Campinas, Brazil, 2001.
3. F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
4. R.W. Cottle, J.S. Pang, and R.E. Stone, The Linear Complementarity Problem, Academic Press, New York, 1992.
5. F. Facchinei, Minimization of SC^1 functions and the Maratos effect, Oper. Res. Lett., 1995, 17:131–137.
6. F. Facchinei and C. Kanzow, A nonsmooth inexact Newton method for the solution of large scale nonlinear complementarity problems, Math. Program., 1997, 76:493–512.
7. F. Facchinei and J.S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Springer, New York, 2003.
8. F. Facchinei and J. Soares, A new merit function for nonliner complementarity problems and a related algorithm, SIAM J. Optim., 1997, 7:225–247.
9. A. Fischer, A special Newton-type optimization method, Optimization, 1992, 24:269–284.
10. A. Fischer, On the local superlinear convergence of a Newton-type method for LCP under weak conditions, Optim. Methods Soft., 1995, 6:83–107.
11. H. Jiang, M. Fukushima, L. Qi, and D. Sun, A trust region method for solving generalized complementarity problems, SIAM J. Optim., 1998, 8:140–157.
12. H. Jiang and L. Qi, A new nonsmooth equations approach to nonlinear complementarity problems, SIAM J. Control. Optim., 1997, 35:178–193.
13. H. Jiang and D. Ralph, Global and local superlinear convergence analysis of Newton-type methods for semismooth equations with smooth least squares, in Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, M. Fukushima and L. Qi (eds.), Kluwer, Dordrecht, 1998, pp. 181–209.
14. C. Kanzow and M. Fukushima, Equivalence of the generalized complementarity problem to differentiable unconstrained minimization, J. Optim. Theory Appl., 1996, 90:581–603.
15. C. Kanzow and H. Kleinmichel, A new class of semismooth Newton-type methods for nonlinear complementarity problems, Comput. Optim. Appl., 1998, 11:227–251.
16. C. Kanzow and H. Qi, A QP-free constrained Newton-type method for variational inequality problems, Math. Program., 1999, 85:81–106.
17. T.D. Luca, F. Facchinei, and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, Math. Program., 1996, 75:407–439.
18. J.V. Outrata and J. Zowe, A Newton method for a class of quasi-variational inequalities, Comput. Optim. Appl., 1995, 4:5–21.
19. J.S. Pang and L. Qi, Nonsmooth equations: motivation and algorithms, SIAM J. Optim., 1993, 3:443–465.
20. L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Math. Oper. Res., 1993, 18:227–244.
21. L. Qi and J. Sun, A nonsmooth version of Newton's method, Math. Program., 1993, 58:353–367.
22. P. Tseng, N. Yamashita and M. Fukushima, Equivalence of complementarity problems to differentiable minimization: a unified approach, SIAM J. Optim., 1996, 6:446–460.
23. N. Yamashita and M. Fukushima, Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems, Math. Program., 1997, 76:469–491.
24. N. Yamashita and M. Fukushima, On the rate of convergence of the Levenberg–Marquardt method, Computing (Suppl), 2001, 15:239–249.

Accepted 8 *December* 2004. *Online publication* 10 *May* 2005.