Mathematica Numerica Sinica [illegible]
[illegible] in Chinese Journal of Numerical Mathematics
and Applications [illegible]

A Projection and Contraction Method for the Nonlinear Complementarity Problem and Its Extensions

Sun Defeng
Institute of Applied Mathematics, Academia Sinca
Beijing, 100080

Abstract

In this paper we propose a new global convergent iterative method for solving the nonlinear complementarity problem and it's related problems. The method behaves effectively not only for linear cases but also for general nonlinear cases. In special case, our method reduces to the same, which was also discussed in [9,11,22] for linear cases.

1 Introduction

The nonlinear complementarity problem, denoted by $NCP(R^n_+, F)$, is to find $x^* \in R^n_+$ such that

$$F(x^*)^T x^* = 0, \;\; F(x^*) \in R^n_+, \tag{1.1}$$

where F is a mapping from R^n_+ into R^n. In this paper, we consider the following general nonlinear complementarity problem: to find $x^* \in X = \{x \in R^n | l \le x \le u\}$ such that

$$F(x^*)^T(x - x^*) \ge 0 \;\; \forall x \in X, \tag{1.2}$$

where l and u are two vectors of $\{R \bigcup \infty\}^n$ and $l \le u$. For this problem, we will denote it by $NCP(X, F)$. When $X = R^n_+$, problem (1.2) reduces to problem (1.1). Let

$$X^* = \{x \in X | x \text{ solves } NCP(X, F)\}. \tag{1.3}$$

It is easy to see that $x \in X^*$ if and only if x satisfies the following projection equation (may be nonsmooth):

$$P_X[x - \beta F(x)] = x \;\; \text{for some or any } \beta > 0, \tag{1.4}$$

where for any $y \in R^n$, $P_X(y) = \text{argmin}\{x \in X | \|x - y\|\}$ (1.5)
and $\|\cdot\|$ denotes the l_2 norm of R^n or it's induced matrix norm. See [3] for a proof. For the simplicity of X, it is easy to implement $P_X(y)$. When X is any other nonempty convex subset of R^n, we can denote the variational inequality problem (abbreviated to $VIP(X, F)$) similar to that of (1.2). We can also define $X^* = \{x \in X | x \text{ solves } VIP(X, F)\}$. From [3], we know that x solves $VIP(X, F)$ if and only if x solves the projection equation (1.4) for some $\beta > 0$. Therefore, in this paper we put our main attention on solving the projection equation (1.4). The development of algorithms for the $NCP(X, F)$ and

$VIP(X,F)$ has a long history in mathematical programming, see [7] for a comprehensive review of this literature.

Defnition 1.1 The mapping $F: R^n \longrightarrow R^n$ is said to

(a) be monotone over a set X if

$$[F(x)-F(y)]^T(x-y) \geq 0 \quad \forall\, x,y \in X; \tag{1.6}$$

(b) be pseudomonotone over X if

$$F(y)^T(x-y) \geq 0 \text{ implies } F(x)^T(x-y) \geq 0 \quad \forall\, x,y \in X; \tag{1.7}$$

(c) be strongly monotone over X if there exists a positive constant $\beta > 0$ such that

$$[F(x)-F(y)]^T(x-y) \geq \beta\|x-y\|^2 \quad \forall\, x,y \in X; \tag{1.8}$$

(d) satisfy the solvability condition over X if X^* is nonempty and for any $x^* \in X$

$$F(x)^T(x-x^*) \geq 0 \quad \forall\, x \in X. \tag{1.9}$$

When F(x) is monotone, and Lipschitz continuous, i.e., there exists a constant $L > 0$ such that

$$\|F(x)-F(y)\| \leq L\|x-y\| \quad \forall\, x,y \in X, \tag{1.10}$$

Korpelevich [15] studied a certain modification of the gradient method that used the idea of extrapolation. His method, which is called Extragradient Method, is as follows:

$$\begin{cases} \bar{x}^k = P_X[x^k - \beta F(x^k)] \\ x^{k+1} = P_X[x^k - \beta F(\bar{x}^k)], \end{cases} \tag{1.11}$$

where $0 < \beta < 1/L$. The advantage of the above algorithm is that it only needs monotonicity, not as other algorithms need strong monotonicity, see [4,7,20]; the obvious drawback of the extragradient method is that it also needs the Lipschitz constant, in practice, however, it is not easy to implement. In [21], the author gave a modification of the extragradient method in which there introduced the inexact line searches into extragradient method and didn't need the Lipschitz constant. The modified extragradient method is as follows:

Given constants $\eta \in (0,1), \alpha \in (0,1)$and $s \in (0,\infty)$, the iterative form is

$$\begin{cases} \bar{x}^k = P_X[x^k - \beta_k F(x^k)] \\ x^{k+1} = P_X[x^k - \beta_k F(\bar{x}^k)], \end{cases} \tag{1.12}$$

where $\beta_k = s\alpha^{m_k}$ and m_k is the smallest nonnegative integer such that

$$\|F(\bar{x}^k) - F(x^k)\| \leq \eta\|\bar{x}^k - x^k\| \tag{1.13}$$

holds. The modified algorithm is of good numerical results, but when reduced to linear cases , it can't vie with those developed by He [9,11] and He and Stoer [10] only for linear cases. In this paper our main aim is to design a new projection and contraction method for solving general nonlinear complementarity problem. Our algorithm behaves effectively not only for linear cases but also for nonlinear cases. In special case (in particular for linear programming) our algorithm reduces to the same, which was also discussed in [11,22] and the later is designed for linear complementarity problem.

2 Basic Preliminaries

Throughout this paper, we assume that X is a nonempty convex subset of R^n and $F(x)$ is continuous over X.

Lemma 2.1 [18]. If $F(x)$ is continuous over a nonempty compact convex set Y, then there exists $y^* \in Y$ such that

$$F(y^*)^T(y - y^*) \geq 0 \text{ for all } y \in Y.$$

Lemma 2.2 [23]. For the projection operator P_X, we have

(i) when $y \in X$, $[P_X(z) - z]^T[P_X(z) - y] \leq 0$ for all $z \in R^n$; (2.1)

(ii) $||P_X(z) - P_X(y)|| \leq ||x - y||$ for all $y, z \in R^n$. (2.2)

Lemma 2.3 [2,5]. Given $x \in R^n$ and $d \in R^n$, then function θ defined by

$$\theta(\beta) = \frac{||P_X(x + \beta d) - x||}{\beta}, \quad \beta > 0 \tag{2.3}$$

is antitone (nonincreasing).

Choose an arbitrary constant $\eta \in (0, 1]$, define

$$\varphi(x, \beta) = \eta F(x)^T\{x - P_X[x - \beta F(x)]\}, \tag{2.4}$$

$$\psi(x, \beta) = \eta||x - P_X[x - \beta F(x)]||^2/\beta, \tag{2.5}$$

where β is a positive constant.

From (i) of lemma 2.2, take $z = x - \beta F(x)$ and $y = x$, we have

$$\beta\{x - P_X[x - \beta F(x)]\}^T F(x) \geq ||x - P_X[x - \beta F(x)||^2. \tag{2.6}$$

Combining (2.4), (2.5) and (2.6), we have

Theorem 2.1. Let $\varphi(x, \beta)$ and $\psi(x, \beta)$ be defined as in (2.4) and (2.5), respectively, then

(i) $\varphi(x, \beta) \geq \psi(x, \beta)$ for all $x \in X$;

(ii) $x \in X$ and $\varphi(x, \beta) = 0$ iff $x \in X$ and $\psi(x, \beta) = 0$ iff $x \in X^*$.

Theorem 2.2. Suppose that $F(x)$ is continuous over X and $\eta \in (0, 1)$. If $S \subset X \backslash X^*$ is a compact set, then there exists a positive constant δ such that for all $x \in S$ and $\beta \in (0, \delta]$, we have

$$\begin{aligned} &\{F(x) - F(P_X[x - \beta F(x)])\}^T\{x - P_X[x - \beta F(x)]\} \\ &\leq (1 - \eta)F(x)^T\{x - P_X[x - \beta F(x)]\}. \end{aligned} \tag{2.7}$$

Proof. Since $S \subset X \backslash X^*$ is a compact set and $F(x)$ is continuous over X, there exists a positive number $\delta_0 > 0$ such that for $x \in S$

$$||P_X[x - F(x)] - x|| \geq \delta_0 > 0. \tag{2.8}$$

From lemma 2.3 and (2.8), for all $x \in S$ and $\beta \in (0, 1]$,

$$||x - P_X[x - \beta F(x)]||/\beta \geq ||x - P_X[x - F(x)]|| \geq \delta_0. \tag{2.9}$$

From the continuity of $F(x)$ we know that $F(x)$ is uniformly continuous over X. Therefore, there must exist a constant $\delta > 0 (\delta \leq 1)$ such that for all $x \in S$ and $\beta \in (0, \delta]$,

$$||F(P_X[x - \beta F(x)]) - F(x)|| \leq (1 - \eta)\delta_0. \tag{2.10}$$

Combining (2.9) and (2.10), for all $x \in S$ and $\beta \in (0, \delta]$, we have

$$\begin{aligned}
&[F(x) - F(P_X[x - \beta F(x)])]^T \{x - P_X[x - \beta F(x)]\} \\
&\quad \le \|F(x) - F(P_X[x - \beta F(x)])\| \|x - P_X[x - \beta F(x)]\| \\
&\quad \le (1 - \eta)\|x - P_X[x - \beta F(x)]\|^2/\beta \\
&\quad \le (1 - \eta)F(x)^T \{x - P_X[x - \beta F(x)]\},
\end{aligned}$$

the last inequality follows from (2.6). So we complete the proof. □

Remark 2.1. Theorem 2.2 ensures that the algorithm given in the next section is reasonable.

Remark 2.2. When $F(x) = Dx + c$ and D is a skew-symmetric matrix (i.e., $D^T = -D$), then (2.7) holds for $\eta = 1$, $\beta \in (0, +\infty)$ and $x \in X$.

3 Algorithms and Convergence Properties

If we put,

$$g(x, \beta) = F(P_X[x - \beta F(x)]), \quad \beta > 0, \tag{3.1}$$

then we have

Theorem 3.1 Suppose that $F(x)$ is continuous over X and satisfies the solvability condition (1.9). If there exists a positive number β such that (2.7) holds for some $x \in X$, then

$$(x - x^*)^T g(x, \beta) \ge \varphi(x, \beta) \quad \forall x^* \in X^*. \tag{3.2}$$

Proof. Since $F(x)$ satisfies the solvability condition (1.9), for $x^* \in X^*$, we have

$$\{P_X[x - \beta F(x)] - x^*\}^T F(P_X[x - \beta F(x)]) \ge 0, \tag{3.3}$$

$$\begin{aligned}
(x - x^*)^T g(x, \beta) &= (x - x^*)^T F(P_X[x - \beta F(x)) \\
&= \{x - P_X[x - \beta F(x)]\}^T F(P_X[x - \beta F(x)]) \\
&\quad + \{P_X[x - \beta F(x)] - x^*\}^T F(P_X[x - \beta F(x)]) \\
&\ge \{x - P_X[x - \beta F(x)]\}^T F(P_X[x - \beta F(x)]) \quad \text{(using(3.3))} \\
&= \{x - P_X[x - \beta F(x)]\}^T \{F(P_X[x - \beta F(x)]) - F(x)\} + F(x)^T \{x - P_X[x - \beta F(x)]\} \\
&\ge (\eta - 1)F(x)^T \{x - P_X[x - \beta F(x)]\} + F(x)^T \{x - P_X[x - \beta F(x)]\},
\end{aligned} \tag{3.4}$$

the last inequality follows from (2.7). Therefore,

$$(x - x^*)^T g(x, \beta) \ge \eta F(x)^T \{x - P_X[x - \beta F(x)]\} = \varphi(x, \beta). \quad □$$

Choose positive constants $s \in (0, +\infty)$ and $\eta \in (0, 1)$, we can describe our algorithm:

ALGORITHM A

Given $x^0 \in X$, $\alpha \in (0, 1)$, $\gamma \in (0, 2)$

For $k = 0, 1, \ldots$, if $x^k \notin X^*$, then do

1. Determine $\beta_k = s\alpha^{m_k}$, where m_k is the smallest nonnegative integer such that

$$\begin{aligned}
&\{x^k - P_X[x^k - \beta_k F(x^k)]\}^T \{F(x^k) - F(P_X[x^k - \beta_k F(x^k)])\} \\
&\quad \le (1 - \eta)F(x^k)^T \{x^k - P_x[x^k - \beta_k F(x^k)]\};
\end{aligned} \tag{3.5}$$

2. Calculate $\varphi(x^k, \beta_k)$ and $g(x^k, \beta_k)$;
3. Calculate $\rho_k = \varphi(x^k, \beta_k)/||g(x^k, \beta_k)||^2$. (3.6)
4. Set $x^{k+1} = P_X[x^k - \gamma \rho_k g(x^k, \beta_k)]$. (3.7)

when $X = \{x \in R^n | l \leq x \leq u\}$, we can improve ALGORITHM A.
For $x \in X$, let

$$N = \{i | (x_i = l_i \text{ and } (g(x,\beta))_i \geq 0) \text{ or } (x_i = u_i \text{ and } (g(x,\beta))_i \leq 0)\},$$

$$B = \{1, 2, ..., n\} \backslash N. \tag{3.8}$$

Let

$$(g_N(x,\beta))_i = \begin{cases} 0 & i \in B \\ (g(x,\beta))_i & i \in N \end{cases},$$

$$(g_B(x,\beta))_i = (g(x,\beta))_i - (g_N(x,\beta))_i, \tag{3.9}$$

$i = 1, 2, ..., n$. Then for any $x^* \in X^*$ and $x \in X$, $(x - x^*)^T g_N(x) \leq 0$ and

$$(x - x^*)^T g_B(x,\beta) \geq (x - x^*)^T g(x,\beta). \tag{3.10}$$

Theorem 3.2 Assume that the conditions of theorem 3.1 holds, then

$$(x - x^*)^T g_B(x,\beta) \geq \varphi(x,\beta) \quad \forall x^* \in X^*. \tag{3.11}$$

Proof. The result follows from theorem 3.1 and (3.10). □

<u>**ALGORITHM B**</u> (An improvement of ALGORITHM A)
Given $x^0 \in X$, $s \in (0, +\infty)$, $\eta, \alpha \in (0,1)$ and $\gamma \in (0,2)$
For $k = 0, 1, ...$, if $x^* \notin X^*$, then do
1. Determine $\beta_k = s\alpha^{m_k}$, where m_k is the smallest nonnegative integer such that

$$\begin{aligned} &\{x^k - P_X[x^k - \beta_k F(x^k)]\}^T \{F(x^k) - F(P_X[x^k - \beta_k F(x^k)])\} \\ &\leq (1-\eta) F(x^k)^T \{x^k - P_X[x^k - \beta_k F(x^k)]\} \end{aligned} \tag{3.12}$$

holds;
2. Calculate $\varphi(x^k, \beta_k)$ and $g(x^k, \beta_k)$ by (2.4) and (3.1), respectively;
3. Determine $g_B(x^k, \beta_k)$ by (3.8) and (3.9) and calculate

$$\rho_k = \varphi(x^k, \beta_k)/||g_B(x^k, \beta_k)||^2; \tag{3.13}$$

4. Set $x^{k+1} = P_X[x^k - \gamma \rho_k g_B(x^k, \beta_k)]$. (3.14)

Remark 3.1. When $F(x) = Dx + c$ and D is skew-symmetric, (2.7) holds if we take $\eta = 1$ and $\beta = 1$, and in this case ALGORITHM B is also discussed by He [11]. We note that in [11] the search direction $g(x)$ is given by

$$g(x) = D^T\{x - P_X[x - (Dx + c)]\} + Dx + c.$$

Since D is skew-symmetric, we have

$$\begin{aligned} g(x) &= -D\{x - P_X[x - (Dx + c)]\} + Dx + c \\ &= DP_X[x - (Dx + c)] + c = g(x, 1). \end{aligned} \tag{3.15}$$

So the search direction $g_B(x)$ given in [11] is the same to $g_B(x, 1)$ and the step-sizes are also the same. In particular, for linear programming, our algorithm can generate the same sequence to that of [11]. When $D^T \neq -D$, the algorithm of [11] is different with ours. When $F(x)$ is a nonlinear mapping, there is no correspondent algorithm in [11], but here our algorithm is generalized to nonlinear case.

For the convergent properties, ALGORITHM A and ALGORITHM B are similar, so we only consider ALGORITHM B.

Theorem 3.3. Suppose that X^* is nonempty and $F(x)$ is continuous over $X = \{x \in R^n | l \leq x \leq u\}$. If $F(x)$ satisfies the solvability condition (1.9), then for any $x^* \in X^*$, the sequence $\{x^k\}$ generated by ALGORITHM B satisfies

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma(2-\gamma)\varphi(x^k, \beta_k)^2 / \|g_B(x^k)\|^2. \tag{3.16}$$

Proof. From (ii) of lemma 2.2 and theorem 3.2, we have

$$\begin{aligned}
&\|x^{k+1} - x^*\|^2 = \|P_X[x^k - \gamma\rho_k g_B(x^k, \beta_k)] - x^*\|^2 \\
&\leq \|x^k - \gamma\rho_k g_B(x^k, \beta_k) - x^*\|\|^2 \\
&= \|x^k - x^*\|^2 - 2\gamma\rho_k g_B(x^k, \beta_k)^T(x^k - x^*) + \gamma^2\rho_k^2\|g_B(x^k, \beta_k)\|^2 \\
&\leq \|x^k - x^*\|^2 - 2\gamma\rho_k\varphi(x^k, \beta_k) + \gamma^2\rho_k^2\|g_B(x^k, \beta_k)\|^2 \\
&= \|x^k - x^*\|^2 - 2\gamma\varphi(x^k, \beta_k)/\|g_B(x^k, \beta_k)\|^2 + \gamma^2\varphi(x^k, \beta_k)^2/\|g_B(x^k, \beta_k)\|^2 \\
&= \|x^k - x^*\|^2 - \gamma(2-\gamma)\varphi(x^k, \beta_k)^2/\|g_B(x^k, \beta_k)\|^2,
\end{aligned}$$

which proves (3.16). □

Define

$$\text{dist}(x, X^*) = \inf\{\|x - x^*\| | x^* \in X^*\}. \tag{3.17}$$

Since (3.16) holds for any $x \in X^*$, then form theorem 3.3,

$$[\text{dist}(x^{k+1}, X^*)]^2 \leq [\text{dist}(x^k, X^*)]^2 - \gamma(2-\gamma)\varphi(x^k, \beta_k)^2/\|g_B(x^k, \beta_k)\|^2, \tag{3.18}$$

i.e., the sequence $\{x^k\}$ is Féjer-monotone relatively to X^*.

Theorem 3.4. If the conditions of theorem 3.3 hold, then there exists $\bar{x}^* \in X^*$ such that $x^k \longrightarrow \bar{x}^*$ as $k \longrightarrow \infty$.

Proof. Let $x^* \in X^*$. It is easy to check that each Féjer-monotone sequence is bounded. Suppose

$$\lim_{k \to \infty} \text{dist}(x^k, X^*) = \delta_0 > 0, \tag{3.19}$$

then $\{x^k\} \subset S = \{x \in X | \delta_0 \leq \text{dist}(x, X^*),\ \|x - x^*\| \leq \|x^0 - x^*\|\}$ and S is a compact set. Since $S \subset X \backslash X^*$ is a compact set, then from theorem 2.2 there exists a positive constant δ such that for all $x \in S$ and $\beta \in (0, \delta]$ (2.7) holds. Therefore,

$$\beta_k \geq \min\{\alpha\delta, s\} > 0, \quad \forall k. \tag{3.20}$$

From theorem 2.1 and (3.20),

$$\inf \varphi(x^k, \beta_k) > 0. \tag{3.21}$$

From the definition of $g_B(x, \beta)$ and the continuity of $F(x)$,

$$\sup \|g_B(x^k, \beta_k)\| < +\infty. \tag{3.22}$$

Combining (3.21) and (3.22) we have

$$\inf \varphi(x^k, \beta_k)^2/\|g_B(x^k, \beta_k)\|^2 = \varepsilon_0 > 0. \tag{3.23}$$

From (3.19) there exists an integer $k_0 > 0$ such that for all $k \geq k_0$

$$[\text{dist}(x^k, X^*)]^2 \leq \delta_0^2 + \varepsilon_0(2-\gamma)\gamma/2. \tag{3.24}$$

On the other hand , (3.18), (3.23) and (3.24) gives

$$\begin{aligned}[\text{dist}(x^{k+1}, X^*)]^2 &\leq [\text{dist}(x^k, X^*)]^2 - \varepsilon_0(2-\gamma)\gamma \\ &\leq \delta_0^2 - \varepsilon_0(2-\gamma)\gamma/2 \qquad \forall\, k \geq k_0,\end{aligned}$$

which contradicts (3.19). Therefore,

$$\lim_{k\to\infty} \text{dist}(x^k, X^*) = 0. \tag{3.25}$$

From (3.25) and (3.18) there exists $x^* \in X^*$ such that $x^k \longrightarrow \bar{x}^*$ as $k \longrightarrow \infty$. □

Remark 3.2. When X^* is nonempty and $F(x)$ is pseudomonotone over X, the conclutions of thereoms hold, for in this case solvability condition holds.

Remark 3.3. From lemma 2.1, when X is a nonempty compact convex subset of R^n, $X^* \neq \emptyset$. When X is unbounded, the nonempty conditions of X^* can be found in the comprehensive paper [7].

4 Numerical Experiments

In this section, the results of applying ALGORITHM B to a number of examples which have appeared in literature will be reported, and will be compared against the use of other algorithms. The term"NCP" is used to denote ALGORITHM B, "EGM" and "$MEGM$" refer to the extragradient method [15] and the modified extragradient method [21], respectively, while "LCP" refers to the method of [11], which was designed for the linear complementarity problem. In all the examples, we will take $\alpha = 0.5$ and $\eta = 0.95$ (in numerical experiments, a large value η will cause faster convergence). All the algorithms will terminate when $\varphi(x,1) \leq \eta\varepsilon^2$, where ε is a small tolerance. (note $\varphi(x,1) \geq \eta\|x - P_X[x - F(x)]\|^2$).

Example 1. This example is a 4-variable nonlinear complementarity problem reported by Kojima and Shindo [14]. We take the fisrt steplenth $s = \sqrt{\eta}/4$ and $\varepsilon^2 = 10^{-16}$. Table 1 lists the result for this example. Our algorithm is much faster than the modified extragradient method for two different values of γ. Since this example is nonlinear and nonlipschitz continuous, we didn't consider the methods of "LCP" and "EGM".

Table 1.
Results for example 1

Algorithm	Starting point	Number of iterations	Number of inner iterations
MEGM	(0,0,0,0)	380	758
NCP($\gamma = 1.95$)	(0,0,0,0)	22	22
NCP($\gamma = 1.0$)	(0,0,0,0)	52	52
MEGM	(1,1,1,1)	395	785
NCP($\gamma = 1.95$)	(1,1,1,1)	28	27
NCP($\gamma = 1.0$)	(1,1,1,1)	73	63

Example 2. In this example, we consider a 4-variable equilibrium problem [17]. The function $F : R \times R_+^2 \times R_+ \longrightarrow R^4$ is of the form

$$F(y, p_1, p_2, p_3) = \begin{pmatrix} -p_1 + p_2 + p_3 \\ y - \alpha(b_2p_2 + b_3p_3)/p_1 \\ b_2 - y - (1-\alpha)(b_2p_2 + b_3p_3)/p_2 \\ b_3 - y \end{pmatrix},$$

with the data being the constants $b_2, b_3 > 0$ and $\alpha \in (0, 1)$. For this problem, we will take two sets of constants: $(\alpha, b_2, b_3) = (0.75, 1, 0.5)$ and $(0.75, 1, 2)$. In this example, we take $s = \sqrt{\eta}/2$ and $\varepsilon^2 = 10^{-16}$.

Table 2.
Results for example 2.

Parameter values	Algorithm	Starting point	Number of iterations	Number of inner iterations
$\alpha = 0.75$	MEGM	(1,1,1,1)	103	0
$b_2 = 1.0$	NCP($\gamma = 1.95$)	(1,1,1,1)	42	0
$b_3 = 0.5$	NCP ($\gamma = 1.0$)	(1,1,1,1)	56	0
$\alpha = 0.75$	MEGM	(1,1,1,1)	41	0
$b_2 = 1.0$	NCP($\gamma = 1.95$)	(1,1,1,1)	36	0
$b_3 = 2.0$	NCP($\gamma = 1.0$)	(1,1,1,1)	43	0

Example 3. In this example, we investigate the problem which was also discussed by Ahn [1]. $F(x) = Dx + c$, where

$$D = \begin{pmatrix} 4 & -2 & & & & \\ 1 & 4 & -2 & & & \\ & 1 & 4 & -2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 4 & -2 \\ & & & & 1 & 4 \end{pmatrix}$$

and c in a n-vector. $X = [l, u]$, where $l = (0, ..., 0)^T$ and $u = (1, ..., 1)^T$. For "*EGM*" and "*MEGM*", we take the first steplength $s = \sqrt{\eta}/7$ while for "*NCP*" we take $s = \sqrt{\eta}/4$. In this example, we take $\varepsilon^2 = n10^{-14}$, where n is the dimension of the problem. "*NCP*($\gamma = 1.95$)" can view with "*LCP*($\gamma = 1.0$)", and both of them behave better than "*EGM*" and "*MEGM*" do. Different starting points are of similar conclusions.

Table 3.

Results for example 3 with starting point $(0,...,0)$.

Algorithm	Number of iterations(left) and number of inner iterations(right)									
	n=10		n=100		n=200		n=500		n=1000	
EGM	59	0	59	0	59	0	59	0	58	0
MEGM	59	0	59	0	59	0	59	0	58	0
NCP($\gamma = 1.95$)	11	9	14	11	14	10	17	10	16	10
NCP($\gamma = 1.0$)	31	27	31	26	31	25	31	25	31	24
LCP($\gamma = 1.95$)	39	0	39	0	39	0	39	0	38	0
LCP($\gamma = 1.0$)	18	0	19	0	19	0	19	0	19	0

Example 4. In this example, we investigate a linear complementarity problem for Lemke's algorithm is known to run in exponential time (see [19, Chapter 6]). This Problem has a special structure with $c = (-1, -1, ..., -1)^T$ and

$$D = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 2 & \cdots & 2 \\ 0 & 0 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

$F(x) = Dx + c$. This example was also discussed by Harker and Pang [6], Harker and Xiao [8]. For "EGM", we take $s = \sqrt{\eta}/(\sqrt{2}n)$; for "$MEGM$" take $s = \sqrt{2\eta}/(4\sqrt{\eta})$; for "$NCP$" take $s = \sqrt{\eta}/2$. Here, n is the dimension of the problem. In this example, we take $\varepsilon^2 = n10^{-14}$. From the results, we see that for the starting point $x^0 = (0, ..., 0)^T$, "$NCP(\gamma = 1.95)$" converges slower than "$LCP(\gamma = 1.0$ or $1.95)$" does. However, even in this case, our algorithm behaves better than "EGM" and "$MEGM$" do. When we take the starting point $x^0 = (0.6, ..., 0.6)$ or $(1, ..., 1)$, "$NCP(\gamma = 1.95)$" behaves better than "$LCP(\gamma = 1.95$ or $1.0)$" does. Here we only list out the computational results for $x^0 = (0, ..., 0)$.

Table 4.
Results for example 4 with starting point (1,...,1)

Algorithm	Number of iterations(left) and number of inner iterations(right)											
	n=10		n=20		n=50		n=100		n=200		n=500	
EGM	227	0	434	0	/		/		/		/	
MEGM	150	5	202	5	305	13	372	16	456	21	593	43
NCP($\gamma = 1.95$)	12	8	15	17	20	42	26	73	44	172	64	317
NCP($\gamma = 1.0$)	32	16	36	30	56	100	63	158	71	221	85	359
LCP($\gamma = 1.95$	10	0	11	0	5	0	11	0	7	0	12	0
LCP($\gamma = 1.0$)	26	0	26	0	25	0	26	0	26	0	27	0

"/" indicates that the number of iterations exceeds 1000.

Example 5. In this example, we consider $F(x) = F_1(x) + F_2(x), x = (x_1, ..., x_n)^T, x_0 = x_{n+1} = 0, F_1(x) = (f_1(x), ..., f_n(x))^T, F_2(x) = Dx + c$, where $f_i(x) = x_{i-1}^2 + x_i^2 + x_{i-1}x_i + x_i x_{i+1}$, $i = 1, ..., n$ and D, c is the same to those of example 3. We take $X = [l, u]$, where $l = (0, ..., 0)^T$ and $u = (1, ..., 1)^T$. For "$MEGM$" and "NCP", we take $s = \sqrt{\eta}/4$ and $\varepsilon^2 = n10^{-14}$, where n is the dimension of the problem. Also note that "NCP" behaves much better than "$MEGM$" does.

Table 5 Results for example 5 with starting point $x^0 = l$.

Algorithm	Number of iterations(left) and number of inner iterations(right)							
	n=10		n=20		n=50		n=100	
MEGM	58	57	60	59	61	60	62	60
NCP($\gamma = 1.95$)	14	13	14	13	13	12	13	11
NCP($\gamma = 1.0$)	20	19	19	18	19	18	19	17

Acknowlegement. The author is grateful to Dr. B. He for his helpful suggestions to shorten the proof of theorem 3.4.

References

[1] Ahn, Byong-Hun, "Iterative Methods for Linear Complementarity Problem With Upperbounds and Lowerbounds", Math. Prog. 26 (1983), P265-315.

[2] Calamai, P. H. and Moré, J. J., "Projected Gradient Method for Linearly Constrained Problems", Math. Prog. 39 (1987), P93-116.

[3] Eaves, B. C., "On the Basic Theorem of Complementarity", Math. Prog. 1 (1971), P68-75.

[4] Fukushima, M., "Equivalent Differentiable Optimization Problems and Descent Methods for Asymmetric Variational Inequality Problems", Math. Prog. 53 (1992), P99-110.

[5] Gafni, E. H. and Bertsekas, D. P., "Two-metric Projection Methods for Constrained Optimization", SIAM J. on Contr. Opti. 22 (1984), P936-964.

[6] Harker, P. T. and Pang, J. S., "A damped-Newton Method for Linear Complementarity Problem", in G. Allgower and K. Georg, eds., *Computational Solutions of Nonlinear Systems of Equations*, Lectures in Applied Mathematics, Vol. 26 (American Mathematical Society, Province, RI, 1990) P265-284.

[7] Harker, P. T. and Pang, J. S., "Finite-dimentional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms and Applications", Math. Prog. 48 (1990), P161-220.

[8] Harker, P. T. and Xiao, B., "Newton's Method for the Nonlinear Complementarity Problem: A B-differentiable Equation Approach", Math. Prog. 48 (1990), P339-357.

[9] He, B., "A Saddle Point Algorithm for Linear Programming", Nanjing Daxue Xuebao, Shu Xue Banian Kan 6 (1989) 41-48.

[10] He, B. and Stoer, J., "Solution of Projection Problems Over Polytopes", Numer. Math. 61 (1992) 73-90.

[11] He, B. and Stoer, J., "A Projection and Contraction Method for a Class of Linear Complementarity problems and It's Application in Convex Quadratic Programming", Appl. Math. and Opti. 25 (1992) 247-262.

[12] Karamardian, S., "Generalized Complementarity Problems", JOTA 8 (1971), P747-756.

[13] Karamardian, S and Schaible, S., "Seven Kinds of Monotone Maps", JOTA 66 (1990), P37-46.

[14] Kojima, M, and Shindo, S., "Extensions Of Newton and Quasi-Newton Methods to Systems of PC' Equations", J. of Oper. Res. Society of Japan 29 (1986), P352-374.

[15] Korpelevich, G. M., "Ekstragradientnyi Method Dlia Otyskaniia Sedlovykh Tochek I Drugikh Zadach", Ekonomika I Matematicheski Metody 12 (1976) 747-756.

[16] Lemke, C. E., "On Complementarity Pivot Theory", in Mathematics of the Decision Sciences, G. B. Dantzig and A. F. Veionott (eds) 1968.

[17] Mathiesen, L., "An algorithm Based on a Sequence of Linear Complementarity Problems Applied to a Walrasian Equillibrium Model: An Example", Math. Prog. 37 (1987), P1-18.

[18] Moré, J.J., "Coercivity Conditions in Nonlinear Complementarity Problems", SIAM Review 16 (1974) 1-16.

[19] Murty, K. G., *Linear Complementarity, Linear and Nonlinear Programming* (Helderman, Berlin, 1988).

[20] Pang, J. S., and Chan, D., "Iterative Methods for Variational and complementarity Problems", Math. Prog. 24 (1982), P284-313

[21] Sun, D. F., "I. An Extragradient Method for Solving Variational Inequality Problems", Master Thesis, Dept. of Math., Nanjing University, 210008, P. R. China (July, 1992).

[22] He, B. and He, X., "A Class of Saddle Point Method for Solving Convex Programming", Computational Mathematics 1 (1991), P51-57.

[23] Zarantonello, E. H., "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).