Journal of Computational Mathematics, Vol.13, No.4, 1995, 357–368.

A NEW STEP-SIZE SKILL FOR SOLVING A CLASS OF NONLINEAR PROJECTION EQUATIONS*

D. Sun
(*Institute of Applied Mathematics, Academia Sinica, Beijing, China*)

Abstract

In this paper, a new step-size skill for a projection and contraction method[10] for linear programming is generalized to an iterative method[22] for solving nonlinear projection equation. For linear programming, our scheme is the same as that of[10]. For complementarity problem and related problems, we give an improved algorithm by considering the new step-size skill and ALGORITHM B discussed in [22]. Numerical results are provided.

1. Introduction

In [11], an iterative projection and contraction (PC) method for linear complementarity problems was proposed. In practice, the algorithm behaves effectively, but in theory the step-size can not be proved to be bounded away from zero. So no statement can be made about the rate of convergence. Although a variant of the prime PC algorithm with constant step-size for linear programmming has a linear convergence[9], it converges much slower in parctice. In [10], He proposed a new step-size rule for the prime PC algorithm for the linear programming such that the resulting algorithm has a globally linear convergence property, and showed that the new resulting algorithm works better in practice than the prime PC algorithm. In this paper, we will introduce a new step-size skill to a projection and contraction method for nonlinear complementarity and its extensions[22]. In order to obtain this, we first make a slight modification of the prime algorithm in [22], and then give the new step-size rule. For linear programming, our ALGORITHM C discussed in this paper is the same as that of [10]. For the complementarity problem and related problems, we will give ALGORITHM D by considering ALGORITHM C in this paper and ALGORITHM B in [22]. Both in theoritical and in computational view point, ALGORITHM D is satisfactory.

Assume that the mapping $F : X \subset R^n \to R^n$ is continuous and X is a closed convex subset of R^n, we will consider the solution of the following projection equations:

$$x - P_X[x - F(x)] = 0, \tag{1.1}$$

* Received April 26, 1994.

where for any $y \in R^n$, $P_X(y) = \mathrm{argmin}\{x \in X \| \|x - y\|\}$. (1.2)
Here $\|\cdot\|$ denotes the l_2-norm of R^n or its induced matrix norm of $R^{n\times n}$. The linear programming , nonlinear complementarity problem and nonlinear variational inequality problem can all be casted as a special case of (1.1), see [3] for a proof. For any $\beta > 0$, define

$$e_X(x, \beta) = x - P_X[x - \beta F(x)]. \tag{1.3}$$

Without causing any confusion, we will use $e(x, \beta)$ to represent $e_X(x, \beta)$. It is easy to see that x is a solution of (1.1) if and only if $e(x, \beta) = 0$ for some or any $\beta > 0$. Denote

$$X^* = \{x \in X | x \text{ is a solution of } (1.1)\}. \tag{1.4}$$

Definition 1. The mapping $F : R^n \to R^n$ is said to
(a) be monotone over a set X if

$$[F(x) - F(y)]^T (x - y) \geq 0, \quad \text{for all } x, y \in X; \tag{1.5}$$

(b) be pseudomonotone over X if

$$F(y)^T (x - y) \geq 0 \ \text{ implies } F(x)^T (x - y) \geq 0, \quad \text{for all } x, y \in X. \tag{1.6}$$

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 1[18]. *If $F(x)$ is continuous over a nonempty compact convex set Y, then there exists $y^* \in Y$ suth that*

$$F(y^*)^T (y - y^*) \geq 0, \quad for\ all\ y \in Y.$$

Lemma 2[23]. *For the projection operator $P_X(\cdot)$, we have*
(i)*when $y \in X$, $[z - P_X(z)]^T [y - P_X(z)] \leq 0$, for all $z \in R^n$;* (2.1)
(ii) *$\|P_X(z) - P_X(y)\| \leq \|z - y\|$, for all $y, z \in R^n$.* (2.2)

Lemma 3[2,5]. *Given $x \in R^n$ and $d \in R^n$, then the 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 $\eta \in (0, 1)$ (e.g., $\eta = 1/2$). When $x \in X \backslash X^*$, define

$$\eta(x) = \begin{cases} \max\Big\{\eta, 1 - \dfrac{t(x)}{\|e(x,1)\|^2}\Big\}, & \text{if } t(x) > 0 \\ 1, & \text{otherwise} \end{cases}, \tag{2.4}$$

where $t(x) = [F(x) - F(P_X[x - F(x)])]^T e(x, 1)$.

For any $x \in X$ and $\beta > 0$, define

$$\varphi(x,\beta) = F(x)^T e(x,\beta), \tag{2.5}$$

$$\psi(x,\beta) = \|e(x,\beta)\|^2/\beta. \tag{2.6}$$

From (i) of Lemma 2, taking $z = x - \beta F(x)$ and $y = x$, we have

$$\beta F(x)^T e(x,\beta) \geq \|e(x,\beta)\|^2. \tag{2.7}$$

By (2.5)-(2.7), and noticing that for any $x \in X \backslash X^*$, $\eta(x) \in [\eta, 1]$, we have

Theorem 1. *Let $\varphi(x,\beta)$ and $\psi(x,\beta)$ be defined as in (2.5) and (2.6), respectively, then for any $\beta > 0$*

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

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

For $x \in X \backslash X^*$, define

$$s(x) = \begin{cases} [1-\eta(x)]\dfrac{\|e(x,1)\|^2}{t(x)}, & \text{if } t(x) > 0 \\ 1, & \text{otherwise} \end{cases}, \tag{2.8}$$

where $t(x) = [F(x) - F(P_X[x - F(x)])]^T e(x,1)$. It is easy to see that $0 < s(x) \leq 1$.

Theorem 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 $\delta(\leq 1)$ such that for all $x \in S$ and $\beta \in (0,\delta]$, when $s(x) < 1$, we have*

$$[F(x) - F(P_X[x - \beta F(x)])]^T e(x,\beta) \leq [1-\eta(x)]\psi(x,\beta). \tag{2.9}$$

Proof. Note that for any $x \in X \backslash X^*$ with $s(x) < 1$, we have

$$[F(x) - F(P_X[x - F(x)])]^T e(x,1) > 0$$

and

$$\eta(x) > 1 - \frac{[F(x) - F(P_X[x - F(x)])]^T e(x,1)}{\|e(x,1)\|^2},$$

which, and the definition of $\eta(x)$, means that $\eta(x) = \eta$. The rest proof is similar to Theorem 2.2 in [22].

In [22], we proposed a projection and contraction method (ALGORITHM A) for solving nonlinear projection equations.

ALGORITHM A

Given $x^0 \in X$, positive constants $s \in (0,+\infty)$, η and $\alpha \in (0,1)$, and $0 < \Delta_1 \leq \Delta_2 < 2$ (In [22] we just take $\Delta_1 = \Delta_2$).

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

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

$$[F(x^k) - F(P_X[x^k - s\alpha^m F(x^k)])]^T e(x^k, s\alpha^m) \leq (1-\eta)\varphi(x^k, s\alpha^m)$$

holds.

2. Calculate $g(x^k, \beta_k) := F(P_X[x^k - \beta_k F(x^k)])$.
3. Calculate

$$\rho_k = \eta\varphi(x^k, \beta_k)/||g(x^k, \beta_k)||^2.$$

4. Take $\gamma_k \in [\Delta_1, \Delta_2]$ (In [22] we just take $\gamma_k = \gamma = \Delta_1 = \Delta_2$) and set

$$\begin{aligned} \bar{x}^k &= x^k - \gamma_k\rho_k g(x^k, \beta_k), \\ x^{k+1} &= P_X(\bar{x}^k). \end{aligned}$$

When X is of the form $X = \{x \in R^n \ | l \leq x \leq u\}$, where l and u are two vectors of $\{R \cup \{\infty\}\}^n$, we gave an improvement of ALGORITHM A, which is called ALGORITHM B, in [22]. As a comparision to ALGORITHM D, the iterative form of ALGORITHM B will be listed out in the last part of section 3.

3. Algorithms and Convergence Properties

If we set

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

then we have

Theorem 3. *Suppose that $F(x)$ is continuous and pseudomonotone over X. If $X^* \neq \emptyset$ and there exists a positive number β such that (2.9) holds for some $x \in X \backslash X^*$, then*

$$(x - x^*)^T g(x, \beta) \geq \varphi(x, \beta) - [1 - \eta(x)]\psi(x, \beta), \quad \textit{for all } x^* \in X^*. \tag{3.2}$$

Proof. Since $X^* \neq \emptyset$, from [3] we know that for any $x^* \in X^*$, $y \in X$ we have

$$F(x^*)^T(y - x^*) \geq 0,$$

which, and the pseudomonotonicity of $F(x)$, means

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

Therefore,

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

the last inequality follows from (2.9). Therefore,

$$(x - x^*)^T g(x, \beta) \geq \varphi(x, \beta) - [1 - \eta(x)]\psi(x, \beta).$$

Now , we state our algorithm with new step-size rule.

ALGORITHM C

Given $x^0 \in X$, positive constants $\eta, \alpha \in (0,1)$ and $0 < \Delta_1 \leq \Delta_2 < 2$.

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

1. Calculate $\eta(x^k)$ and $s(x^k)$. If $s(x^k) = 1$, let $\beta_k = 1$; otherwise determine $\beta_k = s(x^k)\alpha^{m_k}$, where m_k is the smallest nonnegative integer m such that

$$[F(x^k) - F(P_X[x^k - s(x^k)\alpha^m F(x^k)])]^T e(x^k, s(x^k)\alpha^m) \leq [1 - \eta(x)]\psi(x^k, s(x^k)\alpha^m) \tag{3.5}$$

holds.

2. Calculate $g(x^k, \beta_k)$ by (3.1).
3. Calculate

$$\rho_k = \frac{\eta(x^k)\beta_k\|e(x^k, \beta_k)\|^2}{\|e(x^k, \beta_k) - \beta_k[F(x^k) - F(P_X[x^k - \beta_k F(x^k)])]\|^2}. \tag{3.6}$$

4. Take $\gamma_k \in [\Delta_1, \Delta_2]$ and set

$$\bar{x}^k = x^k - \gamma_k \rho_k g(x^k, \beta_k), \tag{3.7}$$

$$x^{k+1} = P_X(\bar{x}^k). \tag{3.8}$$

Remark 1. Theorem 2 ensures that β_k can be obtained in finine number of trials if $s(x^k) < 1$. When $s(x^k) = 1$, (3.5) holds for $m = 0$.

Remark 2. When $F(x) = Dx + c$ and D is a skew-symmetric matrix (i.e., $D^T = -D$), then we have $\eta(x^k) = s(x^k) = 1$, which results that $\beta_k = 1$ for each step, and

$$\begin{aligned}
\rho_k &= \frac{\eta(x^k)\beta_k\|e(x^k, \beta_k)\|^2}{\|e(x^k, \beta_k) - \beta_k[F(x^k) - F(P_X[x^k - \beta_k F(x^k)])]\|^2} \\
&= \frac{\|e(x^k, 1)\|^2}{\|e(x^k, 1) - De(x^k, 1)\|^2} \\
&= \frac{\|e(x^k, 1)\|^2}{\|e(x^k, 1)\|^2 + \| - De(x^k, 1)\|^2 - 2e(x^k, 1)^T De(x^k, 1)} \\
&= \frac{\|e(x^k, 1)\|^2}{\|e(x^k, 1)\|^2 + \|D^T e(x^k, 1)\|^2}.
\end{aligned} \tag{3.9}$$

So for linear programming (when translated into an equivalent linear complementarity problem), our ALGORITHM C is the same as that of [10].

Theorem 4. *Suppose that $F(x)$ is continuous and pseudomonotone over X. If $X^* \neq \emptyset$, then for any $x^* \in X^*$, the sequence $\{x^k\}$ generated by* ALGORITHM C *satisfies*

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_k(2 - \gamma_k)\rho_k\psi(x^k, \beta_k). \tag{3.10}$$

Proof. From (i) of Lemma 2 , we have

$$\begin{aligned}
&\|x^{k+1} - x^*\|^2 = \|P_X(\bar{x}^k) - x^*\|^2 \\
&= \|\bar{x}^k - x^*\|^2 - \|\bar{x}^k - P_X(\bar{x}^k)\|^2 + 2[\bar{x}^k - P_X(\bar{x}^k)]^T[x^* - P_X(\bar{x}^k)] \\
&\leq \|\bar{x}^k - x^*\|^2 - \|\bar{x}^k - x^{k+1}\|^2 \\
&= \|x^k - x^*\|^2 - 2\gamma_k\rho_k(x^k - x^*)^T g(x^k, \beta_k) + \gamma_k^2\rho_k^2\|g(x^k, \beta_k)\|^2 \\
&\quad -[\gamma_k^2\rho_k^2\|g(x^k, \beta_k)\|^2 + \|x^k - x^{k+1}\|^2 - 2\gamma_k\rho_k g(x^k, \beta_k)^T(x^k - x^{k+1})] \\
&= \|x^k - x^*\|^2 - 2\gamma_k\rho_k(x^k - x^*)^T g(x^k, \beta_k) - \|x^k - x^{k+1}\|^2 \\
&\quad +2\gamma_k\rho_k F(P_X[x^k - \beta_k F(x^k)])^T(x^k - x^{k+1}).
\end{aligned}$$

Therefore,

$$\begin{aligned}\|x^{k+1}-x^*\|^2 \le{}& \|x^k-x^*\|^2 - 2\gamma_k\rho_k(x^k-x^*)^Tg(x^k,\beta_k) - \|x^k-x^{k+1}\|^2\\ &-\frac{\gamma_k^2\rho_k^2}{\beta_k^2}\|e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\|^2\\ &+\frac{\gamma_k^2\rho_k^2}{\beta_k^2}\|e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\|^2\\ &+\frac{2\gamma_k\rho_k}{\beta_k}\{e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\}^T(x^k-x^{k+1})\\ &-\frac{2\gamma_k\rho_k}{\beta_k}[e(x^k,\beta_k)-\beta_kF(x^k)]^T(x^k-x^{k+1}),\end{aligned}$$

which means

$$\begin{aligned}\|x^{k+1}-x^*\|^2 \le{}& \|x^k-x^*\|^2 - 2\gamma_k\rho_k(x^k-x^*)^Tg(x^k,\beta_k)\\ &-\|\frac{\gamma_k\rho_k}{\beta_k}\{e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta F(x^k)])]\}-(x^k-x^{k+1})\|^2\\ &+\frac{\gamma_k^2\rho_k^2}{\beta_k^2}\|e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\|^2\\ &-\frac{2\gamma_k\rho_k}{\beta_k}[e(x^k,\beta_k)-\beta_kF(x^k)]^T(x^k-x^{k+1}).\end{aligned}$$

Hence from Theorem 3 and the above formula, we have

$$\begin{aligned}\|x^{k+1}-x^*\|^2 \le{}& \|x^k-x^*\|^2 - 2\gamma_k\rho_k\{\varphi(x^k,\beta_k)-[1-\eta(x^k)]\psi(x^k,\beta_k)\}\\ &+\frac{\gamma_k^2\rho_k^2}{\beta_k^2}\|e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\|^2\\ &-\frac{2\gamma_k\rho_k}{\beta_k}[e(x^k,\beta_k)-\beta_kF(x^k)]^T(x^k-x^{k+1})\\ ={}& \|x^k-x^*\|^2 - 2\gamma_k\rho_k\{\psi(x^k,\beta_k)-[1-\eta(x^k)]\psi(x^k,\beta_k)\}\\ &+\frac{2\gamma_k\rho_k}{\beta_k}\{-\beta_k\varphi(x^k,\beta_k)+\beta_k\psi(x^k,\beta_k)-[e(x^k,\beta_k)-\beta_kF(x^k)]^T(x^k-x^{k+1})\}\\ &+\frac{\gamma_k^2\rho_k^2}{\beta_k^2}\|e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\|^2.\end{aligned}$$

After rearrangement, we have

$$\begin{aligned}\|x^{k+1}-x^*\|^2 \le{}& \|x^k-x^*\|^2 - 2\gamma_k\rho_k\eta(x^k)\psi(x^k,\beta_k)\\ &+\frac{2\gamma_k\rho_k}{\beta_k}[e(x^k,\beta_k)-\beta_kF(x^k)]^T[e(x^k,\beta_k)-(x^k-x^{k+1})]\\ &+\frac{\gamma_k^2\rho_k^2}{\beta_k^2}\|e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\|^2\\ ={}& \|x^k-x^*\|^2 - 2\gamma_k\rho_k\eta(x^k)\psi(x^k,\beta_k)\\ &+\frac{2\gamma_k\rho_k}{\beta_k}\{x^k-\beta_kF(x^k)-P_X[x^k-\beta_kF(x^k)]\}^T\{x^{k+1}-P_X[x^k-\beta_kF(x^k)]\}\\ &+\frac{\gamma_k^2\rho_k^2}{\beta_k^2}\|e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\|^2,\end{aligned}$$

which, and (i) of Lemma 2, means

$$\begin{aligned}\|x^{k+1}-x^*\|^2 &\le \|x^k-x^*\|^2 - 2\gamma_k\rho_k\eta(x^k)\psi(x^k,\beta_k)\\&\quad+\frac{\gamma_k^2\rho_k^2}{\beta_k^2}\|e(x^k,\beta_k)-\beta_k[F(x^k)-F(P_X[x^k-\beta_kF(x^k)])]\|^2\\&= \|x^k-x^*\|^2-\gamma_k(2-\gamma_k)\eta(x^k)\rho_k\psi(x^k,\beta_k),\end{aligned}$$

which proves (3.10).

Define

$$\mathrm{dist}(x,X^*)=\inf\{\|x-x^*\| \mid x^*\in X^*\}. \tag{3.11}$$

Since (3.10) holds for any $x^*\in X^*$, then from Theorem 4,

$$[\mathrm{dist}(x^{k+1},X^*)]^2 \le [\mathrm{dist}(x^k,X^*)]^2-\gamma_k(2-\gamma_k)\eta(x^k)\rho_k\psi(x^k,\beta_k), \tag{3.12}$$

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

Theorem 5. *If the conditions of Theorem 4 hold, then there exists $\bar{x}^*\in X^*$ such that*

$$x^k\to\bar{x}^* \quad \textit{as } k\to\infty.$$

Proof. For the sake of simplicity, we take $\gamma_k=1$. Let $x^*\in X^*$. It is easy to see that each Féjer-monotone sequence is bounded. Suppose that

$$\lim_{k\to\infty}\mathrm{dist}(x^k,X^*)=\underline{\delta}>0, \tag{3.13}$$

then $\{x^k\}\subset S=\{x\in X \mid \underline{\delta}\le \mathrm{dist}(x,X^*),\ \|x-x^*\|\le\|x^0-x^*\|\}$ and S is a compact set. Since $S\subset X\backslash X^*$ is a compact set, then from Theorem 2 there exists a positive number $\delta(\le 1)$ such that for all $x\in S$ with $s(x)<1$ and $\beta\in(0,\delta]$, (2.9) holds. Hence for each k with $s(x^k)<1$, we have

$$\beta_k\ge\min\{\alpha\delta,\ s(x^k)\}. \tag{3.14}$$

From the definition of $s(x^k)$, we know that if $s(x^k)<1$, then

$$[F(x^k)-F(P_X[x^k-F(x^k)])]^Te(x^k,1)>0 \text{ and } \eta(x^k)=\eta,$$

and

$$\begin{aligned}s(x^k)&=(1-\eta)\frac{\|e(x^k,1)\|^2}{[F(x^k)-F(P_X[x^k-F(x^k)])]^Te(x^k,1)}\\&\ge(1-\eta)\frac{\|e(x^k,1)\|}{\|F(x^k)\|+\|F(P_X[x^k-F(x^k)])\|}.\end{aligned} \tag{3.15}$$

Since $\{x^k\}$ is bounded and $F(x)$ is continuous over X, there exists a positive constant M such that

$$\|F(x^k)\|,\ \|F(P_X[x^k-F(x^k)])\|\le M. \tag{3.16}$$

From the continuity of F and $\{x^k\}\subset S\subset X\backslash X^*$, we know that there exists a positive constant δ_0 such that

$$\|e(x^k,1)\|\ge\delta_0. \tag{3.17}$$

From (3.15)—(3.17), for each k with $s(x^k) < 1$ we have

$$s(x^k) \geq (1-\eta)\frac{\delta_0}{2M}. \tag{3.18}$$

Substituting (3.18) into (3.14), gives

$$\beta_k \geq \min\Big\{\alpha\delta,\ (1-\eta)\frac{\delta_0}{2M}\Big\}, \tag{3.19}$$

for all k such that $s(x^k) < 1$. From ALGORITHM C we know that if $s(x^k) = 1$, we have

$$\beta_k = 1. \tag{3.20}$$

Hence there exists a positive constant $\bar{\delta}(\leq 1)$ such that

$$1 \geq \beta_k \geq \bar{\delta} > 0, \quad \text{for all } k. \tag{3.21}$$

From the continuity of $F(x)$ and the boundedness of S,

$$\sup\|e(x^k, 1) - \beta_k[F(x^k) - F(P_X[x^k - \beta_k F(x^k)])]\| < \infty. \tag{3.22}$$

Combining (3.17), (3.21)—(3.22) and Lemma 3, we have

$$\begin{aligned}\inf\{\eta(x^k)\rho_k\psi(x^k,\beta_k)\} &= \inf\Big\{\frac{\eta(x^k)^2\|e(x^k,\beta_k)\|^4}{\|e(x^k,\beta_k) - \beta_k[F(x^k) - F(P_X[x^k - \beta_k F(x^k)])]\|^2}\Big\}\\ &\geq \inf\Big\{\frac{\eta^2\beta_k^4\|e(x^k,1)\|^4}{\|e(x^k,\beta_k) - \beta_k[F(x^k) - \beta_k F(P_X[x^k - \beta_k F(x^k)])]\|^2}\Big\} = \varepsilon_0 > 0.\end{aligned} \tag{3.23}$$

From (3.13) there exists an integer $k_0 > 0$ such that for all $k \geq k_0$,

$$[\mathrm{dist}(x^k, X^*)]^2 \leq \underline{\delta}^2 + \varepsilon_0/2. \tag{3.24}$$

On the other hand, (3.12), (3.23) and (3.24) gives

$$\begin{aligned}[\mathrm{dist}(x^{k+1}, X^*)]^2 &\leq [\mathrm{dist}(x^k, X^*)]^2 - \varepsilon_0\\ &\leq \underline{\delta}^2 - \varepsilon_0/2, \quad \text{for all } k \geq k_0,\end{aligned}$$

which contradicts (3.13). Therefore,

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

From (3.25) and (3.12) there exists $\bar{x}^* \in X^*$ such that

$$x^k \to \bar{x}^* \quad \text{as } k \to \infty.$$

Now the proof is completed.

When X is of the following form

$$X = \{x \in R^n \mid l \leq x \leq u\}, \tag{3.26}$$

where l and u are two vectors of $\{R \cup \{\infty\}\}^n$, we can give a modification of ALGORITHM C. For any $x \in X$, $\beta > 0$, denote

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

$$B(x,\beta) = \{1, ..., n\} \backslash N(x,\beta). \tag{3.27}$$

Denote $g_N(x,\beta)$ and $g_B(x,\beta)$ as follows:

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

$$(g_B(x,\beta))_i = (g(x,\beta))_i - (g_N(x,\beta))_i, \quad i = 1, ..., n. \tag{3.28}$$

As a comparision we will first list out ALGORITHM B[22].

ALGORITHM B (An improvement of ALGORITHM A)

Given $x^0 \in X$, positive constants $s \in (0, +\infty)$, η and $\alpha \in (0,1)$, and $0 < \Delta_1 \leq \Delta_2 < 2$ (In [22] we just take $\Delta_1 = \Delta_2$).

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

1. This step is the same as 1 of ALGORITHM A.
2. Calculate $g(x^k, \beta_k)$ and $g_B(x^k, \beta_k)$ by (3.1) and (3.28), respectively.
3. Calculate

$$\rho_k = \eta\varphi(x^k, \beta_k)/\|g_B(x^k, \beta_k)\|^2.$$

4. Take $\gamma_k \in [\Delta_1, \Delta_2]$ (In [22] we just take $\gamma_k = \gamma = \Delta_1 = \Delta_2$) and set

$$\bar{x}^k = x^k - \gamma_k \rho_k g_B(x^k, \beta_k),$$
$$x^{k+1} = P_X(\bar{x}^k).$$

Now we describe ALGORITHM D.

ALGORITHM D (An improvement of ALGORITHM C)

Given $x^0 \in X$, positive constants $\eta, \alpha \in (0,1)$ and $0 < \Delta_1 \leq \Delta_2 < 2$.

1. This step is the same as 1 of ALGORITHM C.
2. Calculate $g(x^k, \beta_k)$ and $g_B(x^k, \beta_k)$ by (3.1) and (3.28), respectively.
3. Calculate

$$\rho_k = \max\Big\{\frac{\eta(x^k)\beta_k\|e(x^k,\beta_k)\|^2}{\|e(x^k,\beta_k) - \beta_k[F(x^k) - F(P_X[x^k - \beta_k F(x^k)])]\|^2}, \frac{\eta(x^k)\varphi(x^k,\beta_k)}{\|g_B(x^k,\beta_k)\|^2}\Big\}.$$

4. Take $\gamma_k \in [\Delta_1, \Delta_2]$ and set (3.29)

$$\bar{x}^k = x^k - \gamma_k \rho_k g_B(x^k, \beta_k), \tag{3.30}$$

$$x^{k+1} = P_X(\bar{x}^k). \tag{3.31}$$

4. Numerical Experiments

In the following examples, we take $\eta = \alpha = 0.5$, and $\Delta_1 = \Delta_2 = 1.95$ (Numerical results show that when γ_k approaches 2, the resulting algorithms behave better. This phenomeneon is also observed by He[10] for solving linear programming. The reason may exist in that some uncertain terms are lost by enlarging the inequalities.) and use $\varphi(x,1) \leq \varepsilon^2$ as a stop criteria, where ε is a small nonegative number. In practice, we will use ALGORITHM D instead of ALGORITHM C, although it is reported in [10] that for linear programming ALGORITHM C behaves better than the prime algorithm. For using ALGORITHM B in [22], we use $s(x^k)$ and $\eta(x^k)$ to substitute s and η, respectively. "ALGORITHM B" and "ALGORITHM D" will be abbreviated as Alg. B and Alg. D, respectively. For Alg. B and Alg. D, the computing cost of each (out) iteration is nearly the same and the inner iteration takes about half of the computing cost of the (out) iteration. So the effeciency of Alg. B and Alg. D can be measured by the sum of the number of iterations and the the half of the number of inner iterations. Numerical results show that both Alg. B and Alg. D behave effectively, and Alg. D behaves slightly better than Alg. B does. But one point should be stressed is that the step-size in Alg. D can be proved to be bounded away from zero under the local Lipschitzian condition of the mapping F while Alg. B can not have such a conclusion. So both in practice and in theory, Alg. D is an appropriate choice. Just as a referee pointed out that it was easy to construct a small example to make the present algorithms converge very slowly. Nevertheless, we did not find a more effective method under the conditions given in this paper.

Example 1. This example is discussed in [1, 22], the numerical results are given by Table 1.

Table 1

Results for example 1 with starting point (0,...,0)

Algorithm	Number of iterations (left) and number of inner iterations (right)									
	n=10		n=50		n=100		n=200		n=500	
Alg. B	18	8	20	11	20	11	18	9	20	12
Alg. D	18	9	19	10	15	4	16	6	13	2

where $F(x) = Dx + c$, c is a vector and D is a non-symmetric matrix of the form

$$D = \begin{pmatrix} 4 & -2 & & & & & \\ 1 & 4 & -2 & & & & \\ & 1 & 4 & -2 & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & \cdot \\ & & & & & \cdot \quad \cdot & -2 \\ & & & & & 1 & 4 \end{pmatrix}.$$

$X = [l, u]$, where $l = (0, ..., 0)$ and $u = (1, ..., 1)$. We take $\varepsilon^2 = n10^{-14}$, where n is the dimension of the problem.

Example 2. This example is a linear complementarity problem dicussed in [6, 22].

$F(x) = Dx + c$, where $c = (-1, ..., -1)$ and

$$D = \begin{pmatrix} 1 & 2 & 2 & \cdot & \cdot & \cdot & 2 \\ 0 & 1 & 2 & \cdot & \cdot & \cdot & 2 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & 2 \\ \cdot & \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 1 \end{pmatrix}.$$

We take $\varepsilon^2 = n10^{-16}$, where n is the dimension of the problem.

Table 2

Results for example 2 with starting point (0,..,0)

Algorithm	Number of iterations (left) and number of inner iterations (right)									
	n=10		n=50		n=100		n=200		n=500	
Alg. B	14	3	22	8	20	2	25	5	31	10
Alg. D	12	2	18	3	23	6	24	4	29	5

Example 3. This example is a 4-dimensional nonlinear complementarity problem [14, 22]. We take $\varepsilon^2 = 10^{-16}$. For starting point (0,0,0,0), the number of iterations and the number of inner iterations for Alg. B are 63 and 5 respectively and for Alg. D are 64 and 5 respectively.

Acknowledgement Special thanks are due to Prof. Jiye Han and Dr. Bingsheng He for their guidance and helpness. I also thank two anonymous referees for very helpful comments and suggestions.

References

[1] Byong-hun Ahn, Iterative methods for linear complementarity problem with upperbounds and lowerbounds, *Math. Prog.*, 26 (1983), 295–315.

[2] P. H. Calamai and J. J. Moré, Projected gradient method for linearly constrained problems, *Math. Prog.*, 39 (1987), 93–116.

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

[4] M. Fukushima, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, *Math. Prog.*, 52 (1992), 99–110.

[5] E. H. Gafni and D. P. Bertsekas, Two-metric projection methods for constrained optimization, *SIAM J. Contr. Opti.*, 22 (1984), 936–964.

[6] P.T. Harker and J.S. Pang, 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, 265–284.

[7] P. T. Harker and J.S. Pang, Finite-dimentional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications, *Math. Prog.*, 48 (1990), 161–220

[8] P. T. Harker and B. Xiao, Newton's method for the nonlinear complementarity problem: A B-differentiable equation approach, *Math. Prog.*, 48 (1990), 339–357.

[9] B. S. He, On a class of iterative projection and contraction methods for linear programming, *JOTA*, 78 (1993).

[10] B. S. He, Further developments in an iterative projection and contraction method for linear programming, *J. Comp. Math.*, 11 (1993), 350–364.

[11] B. S. He, A projection and contraction method for a class of linear complementarity problems and its applications in convex quadratic programming, *Appl. Math. Opti.*, 25 (1992), 247–262.

[12] S. Karamardian, Generalized complementarity problems, *JOTA*, 8 (1971), 747–756.

[13] S. Karamardian and S. Schaible, Seven kinds of monotone maps, *JOTA*, 66 (1990), 37–46.

[14] M. Kojima and S. Shindo, Extentions of Newton and quasi-Newton methods to systems of PC′ equations, *J. Oper. Res. Society of Japan*, 29 (1986), 352–374.

[15] G. M. Korpelevich, The extragradient method for finding saddle points and other problems, *Ekonomika i matematicheskie metody*, 12 (1976), 747–756.

[16] C. E. Lemke, On complementarity pivot theory, in Mathematics of the decision sciences, G. B. Dantzig and A. F. Veinott (eds.), 1968.

[17] L. Mathiesen, An algorithm based on a sequence of linear complementarity problems applied to a Walrasian equilibrium model: An example, *Math. Prog.*, 37 (1987), 1–18.

[18] J. J. Moré, Coercivity conditions in nonlinear complementarity problems, *SIAM Review* , 16 (1974), 1–16.

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

[20] J. S. Pang and D. Chan, Iterative methods for variational and complementarity problems, *Math. Prog.*, 24 (1982), 284–313.

[21] D. Sun, An iterative method for solving variational inequality problems and complementarity problems, *Numerical Mathematics, A Journal of Chinese Universities*, 16 (1994), 145–153.

[22] D. Sun, A projection and contraction method for the nonlinear complementarity problem and its extensions, *Mathematica Numerica Sinica*, 16 (1994), 183–194.

[23] E. H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory, in E. H. Zarantonello, ed., Contributions to nonlinear Functional Analysis (Academic Press, New York, 1971).