JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 91, No. 1, pp. 123-140, OCTOBER 1996

A Class of Iterative Methods for Solving Nonlinear Projection Equations[1]

D. SUN[2]

Communicated by Z. Q. Luo

Abstract. A class of globally convergent iterative methods for solving nonlinear projection equations is provided under a continuity condition of the mapping F. When F is pseudomonotone, a necessary and sufficient condition on the nonemptiness of the solution set is obtained.

Key Words. Projection equations, variational inequalities, iterative methods, continuity.

1. Introduction

Assume that the mapping $F: X \subset R^n \rightarrow R^n$ is continuous and X is a closed convex subset of R^n. We consider the solution of the following projection equations:

$$x - \Pi_X[x - F(x)] = 0, \qquad (1)$$

where for any $y \in R^n$,

$$\Pi_X(y) = \operatorname{argmin}\{x \in X \mid \|x - y\|\}.$$

Here, $\|\cdot\|$ denotes the l_2-norm of R^n or its induced matrix norm. It is well known (see, e.g., Refs. 1 and 2) that the projection problem (1) is equivalent to a variational inequality problem, which is to find $x \in X$ such that

$$(y - x)^T F(x) \geq 0, \qquad \text{for all } y \in X. \qquad (2)$$

For any $\beta > 0$, define

$$E_X(x, \beta) = x - \Pi_X[x - \beta F(x)]. \qquad (3)$$

[1]The author would like to thank two referees for their useful comments on this paper and one of them, in particular, for bringing Ref. 15 to his attention. The author also thanks Professor He for sending him Ref. 23.

[2]PhD, Institute of Applied Mathematics, Academia Sinica, Beijing, PR China.

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) if and only if $E(x, \beta)=0$ for some or any $\beta>0$. Denote

$$X^*=\{x\in X \mid x \text{ is a solution of } (1)\}. \quad (4)$$

Definition 1.1. The mapping F: $R^n \to R^n$ is said to be:

(i) monotone over a set D if

$$[F(x)-F(y)]^T(x-y)\geq 0, \qquad \text{for all } x, y\in D; \quad (5)$$

(ii) pseudomonotone over a set D relative to a set $Y(\subset D)$ if

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

$$F(x)^T(x-y)\geq 0, \text{ for all } x\in D, y\in Y. \quad (6)$$

Remark 1.1. When $Y=D$, the pseudomonotonicity of F over a set D relative to Y is the usual pseudomonotonicity, and in this case we say directly that F is pseudomonotone over D.

For solving the projection equations (1) and related problems, there is a long history in the mathematical programming field; see Refs. 2–4 for details. Among the algorithms for solving (1), Newton's method is the basic method when the derivative of F exists and is easy to implement. In this paper, we will investigate a globally convergent method for solving (1) assuming only the continuity of the mapping F.

When F is monotone and Lipschitz continuous over X, i.e., there exists a positive number L such that

$$\|F(x)-F(y)\|\leq L\|x-y\|, \qquad \text{for all } x, y\in X,$$

Korpelevich (Ref. 5) proposed the following extragradient (EG) method:

$$\bar{x}^k=\Pi_X[x^k-\beta F(x^k)],$$

$$x^{k+1}=\Pi_X[x^k-\beta F(\bar{x}^k)],$$

where $\beta\in(0, 1/L)$ is a constant. Since the Lipschitz constant is difficult to estimate in practice, by introducing an inexact line search to relax β to β_k in the above formulas, Khobotov (Ref. 6), Marcotte (Ref. 7), and Sun (Ref. 8) proposed an improved extragradient (IEG) method. The IEG method does not need the value of the Lipschitz constant. EG and IEG methods have global convergence if F is monotone and Lipschitz continuous over X (Refs. 5–7). Moreover, in Sun (Ref. 8) a global convergence theorem was provided for the IEG method if F is pseudomonotone and continuous over X only. For algorithms with strong monotonicity and Lipschitz continuity assumptions, see Fukushima (Ref. 9) and Pang and Chan (Ref. 3).

When F is an affine map, i.e.,

$$F(x) = Mx + c,$$

where $M \in R^{n \times n}$ and $c \in R^n$, He (Refs. 10–13) and He and Stoer (Ref. 14) proposed a class of projection and contraction (PC) methods for solving (1). Some of the results of He (Refs. 10–13) and He and Stoer (Ref. 14) were also obtained recently by Solodov and Tseng (Ref. 15). The numerical results show that PC methods behave much better than the EG method or IEG method in linear cases, i.e., $F(x) = Mx + c$. This stimulates us to investigate such algorithms that not only can compete with the PC methods in linear cases, but also behave much better than the EG method or IEG method in nonlinear cases. By introducing some parameters, Sun (Ref. 16) made a first step toward this. The method introduced by the author of Ref. 16 behaves promisingly in practice, but the stepsize cannot be proved to be bounded away from zero if F is Lipschitz continuous over X. So, a new stepsize procedure was introduced in Sun (Ref. 17). Unfortunately, although the new stepsize is bounded away from zero under the Lipschitz continuity assumption, the corresponding algorithm does not perform as promisingly as the algorithm in Sun (Ref. 16) does. In this paper, we propose a class of iterative methods, which include the method in Sun (Ref. 16) and a new method that not only performs promisingly in practice, but also has the property that the stepsize is bounded away from zero under the Lipschitz continuity assumption, for solving the projection problem (1). In addition, we introduce a practically useful strategy to choose the initial value of β_k at each step. This strategy can greatly reduce the total number of inner iterations. When $F(x) = Mx + c$ and M is a skew-symmetric matrix (i.e., $M^T = -M$), some of our algorithms were also discussed by He (Refs. 10 and 12). For linear convergence of descent methods for solving convex smooth minimization problem, see Luo and Tseng (Ref. 18). Here, we focus our main attention on projection-type methods for the projection problem (1) or the variational inequality problem (2).

In Section 2, we give some preliminaries. In Section 3, we give a class of abstract search directions and the corresponding algorithms. In Section 4, we discuss two forms of search directions which satisfy the requirements. In Section 5, we establish a necessary and sufficient condition on the nonemptiness of the solution set when F is pseudomonotone. Numerical results are presented in Section 6. In Section 7, we give a discussion.

2. Basic Preliminaries

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

Lemma 2.1. See Ref. 19. If F is continuous over a nonempty compact convex set Y, then there exists $y^* \in Y$ such that

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

Lemma 2.2. See Ref. 20. For the projection operator $\Pi_X(\cdot)$, we have:

(i) $[z-\Pi_X(z)]^T[y-\Pi_X(z)] \leq 0$, when $y \in X$, for all $z \in R^n$;
(ii) $\|\Pi_X(z)-\Pi_X(y)\| \leq \|z-y\|$, for all y, $z \in R^n$.

Lemma 2.3. See Refs. 21 and 22. Given $x \in R^n$ and $d \in R^n$, the function θ defined by

$$\theta(\beta) = \|\Pi_X(x+\beta d)-x\|/\beta, \qquad \beta > 0$$

is antitone (nonincreasing).

Choose an arbitrary constant $\eta \in (0, 1)$, e.g., $\eta = 1/2$. When $x \in X$, define

$$\eta(x) = \begin{cases} \max\{\eta, 1-t(x)/\|E(x, 1)\|^2\}, & \text{if } t(x) > 0, \\ 1, & \text{otherwise,} \end{cases} \tag{7}$$

$$s(x) = \begin{cases} [1-\eta(x)]\|E(x, 1)\|^2/t(x), & \text{if } t(x) > 0, \\ 1, & \text{otherwise,} \end{cases} \tag{8}$$

where

$$t(x) = \{F(x) - F(\Pi_X[x-F(x)])\}^T E(x, 1).$$

It is easy to see that $0 < s(x) \leq 1$.

Theorem 2.1. Suppose that F is continuous over X and $\eta \in (0, 1)$ is a constant. If $S \subset X \backslash X^*$ is a compact set, then there exists a positive constant $\delta \leq 1$ such that, for all $x \in S$ with $s(x) < 1$ and $\beta \in (0, \delta]$, we have

$$\begin{aligned} &\{F(x) - F(\Pi_X[x-\beta F(x)])\}^T E(x, \beta) \\ &\leq [1-\eta(x)]\|E(x, \beta)\|^2/\beta. \end{aligned} \tag{9}$$

In addition, for all $x \in X \backslash X^*$ with $s(x) = 1$, (9) holds for $\beta = 1$.

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

$$t(x) > 0 \quad \text{and} \quad \eta(x) > 1 - t(x)/\|E(x, 1)\|^2,$$

which means that $\eta(x)=\eta$. Since $S\subset X\backslash X^*$ is a compact set and F is continuous over X, there exists a constant $\delta_0>0$ such that, for all $x\in S$, we have

$$\|\Pi_X[x-F(x)]-x\|\geq\delta_0>0. \tag{10}$$

From Lemma 2.3 and (10), for all $\beta\in(0, 1]$ and $x\in S$, we have

$$\|x-\Pi_X[x-\beta F(x)]\|/\beta\geq\|x-\Pi_X[x-F(x)]\|\geq\delta_0. \tag{11}$$

From the continuity of F, we know that F is uniformly continuous over compact sets. So, from (ii) of Lemma 2.2, we know that there exists a positive constant $\delta\leq 1$, for all $x\in S$ with $s(x)<1$ and $\beta\in(0, \delta]$, such that

$$\|F(\Pi_X[x-\beta F(x)])-F(x)\|\leq(1-\eta)\delta_0. \tag{12}$$

Combining (11) and (12), for all $x\in S$ and $\beta\in(0, \delta]$, we have

$$\begin{aligned}&\{F(x)-F(\Pi_X[x-\beta F(x)])\}^T E(x, \beta)\\ &\leq\|F(x)-F(\Pi_X[x-\beta F(x)])\|\,\|E(x, \beta)\|\\ &\leq(1-\eta)\|E(x, \beta)\|^2/\beta\\ &=[1-\eta(x)]\|E(x, \beta)\|^2/\beta,\end{aligned}$$

which completes the proof of (9). If in addition $x\in X\backslash X^*$ with $s(x)=1$, then from the definition of $s(x)$ we know that

$$t(x)\leq 0 \quad\text{or}\quad t(x)=[1-\eta(x)]\|E(x, 1)\|^2,$$

which means that (9) holds for $\beta=1$. □

3. Algorithms and Convergence

Suppose that $g: R^n\times R^1_{++}\to R^n$ is a continuous mapping. We will use $g(x, \beta)$ as a search direction in this section. The various forms of $g(x, \beta)$ will be given in Section 4. First, we describe our algorithm in the abstract form of $g(x, \beta)$.

Projection and Contraction Method. Given $x^0\in X$, positive constants η, $\alpha\in(0, 1)$, and $0<\Delta_1\leq\Delta_2<2$. For $k=0, 1, \ldots$, if $x^k\notin X^*$, do the following steps.

Step 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

$$\begin{aligned}&\{F(x^k)-F(\Pi_X[x^k-s(x^k)\alpha^m F(x^k)])\}^T E(x^k, s(x^k)\alpha^m)\\ &\leq[1-\eta(x^k)]\|E(x^k, s(x^k)\alpha^m)\|^2/(s(x^k)\alpha^m).\end{aligned} \tag{13}$$

Step 2. Calculate $g(x^k, \beta_k)$.

Step 3. Calculate

$$\rho_k = E(x^k, \beta_k)^T g(x^k, \beta_k)/\|g(x^k, \beta_k)\|^2. \tag{14}$$

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

$$\bar{x}^k = x^k - \gamma_k \rho_k g(x^k, \beta_k), \tag{15}$$

$$x^{k+1} = \Pi_X(\bar{x}^k). \tag{16}$$

Remark 3.1. Theorem 2.1 ensures that β_k can be obtained in a finite number of trials if $s(x^k) < 1$. When $s(x^k) = 1$, (13) holds for $m = 0$.

For $\beta > 0$, define

$$\psi(x, \beta) = \eta(x)\|E(x, \beta)\|^2/\beta. \tag{17}$$

Theorem 3.1. Suppose that F, g are continuous over $X, X \times R^1_{++}$ respectively. If $X^* \neq \varnothing$, and if there exists $x^* \in X^*$ such that the infinite sequence $\{x^k\}$ generated by PC methods satisfies

$$(x^k - x^*)^T g(x^k, \beta_k) \geq E(x^k, \beta_k)^T g(x^k, \beta_k) \geq \psi(x^k, \beta_k), \tag{18}$$

then

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_k(2 - \gamma_k)\psi^2(x^k, \beta_k)/\|g(x^k, \beta_k)\|^2. \tag{19}$$

Proof. From (ii) of Lemma 2.2 and (18), we have

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &= \|\Pi_X[x^k - \gamma_k \rho_k g(x^k, \beta_k)] - x^*\|^2 \\
&\leq \|x^k - \gamma_k \rho_k g(x^k, \beta_k) - x^*\|^2 \\
&= \|x^k - x^*\|^2 - 2\gamma_k \rho_k (x^k - x^*)^T g(x^k, \beta_k) \\
&\quad + \gamma_k^2 \rho_k^2 \|g(x^k, \beta_k)\|^2 \\
&\leq \|x^k - x^*\|^2 - 2\gamma_k \rho_k E(x^k, \beta_k)^T g(x^k, \beta_k) \\
&\quad + \gamma_k^2 \rho_k^2 \|g(x^k, \beta_k)\|^2 \\
&= \|x^k - x^*\|^2 - \gamma_k(2 - \gamma_k) \\
&\qquad \times [E(x^k, \beta_k)^T g(x^k, \beta_k)]^2/\|g(x^k, \beta_k)\|^2 \\
&\leq \|x^k - x^*\|^2 - \gamma_k(2 - \gamma_k)\psi^2(x^k, \beta_k)/\|g(x^k, \beta_k)\|^2,
\end{aligned}$$

which verifies (19). □

Define

$$\mathrm{dist}(x, X^*) = \inf\{\|x - x^*\| \,|\, x^* \in X^*\}. \tag{20}$$

Theorem 3.2. Suppose that the conditions of Theorem 3.1 hold. Then, the infinite sequence $\{x^k\}$ generated by the PC method is bounded and $\liminf_{k\to\infty} \mathrm{dist}(x^k, X^*) = 0$. Furthermore, if (18) holds for each $x^* \in X^*$, then there exists $\bar{x} \in X^*$ such that $x^k \to \bar{x}$ as $k \to \infty$.

Proof. For the sake of simplicity, we take $\gamma_k = 1$.

From (19) we know that $\{\|x^k - x^*\|\}$ is a decreasing sequence. So, the sequence $\{x^k\}$ generated by the PC method is bounded and the sequence $\{\mathrm{dist}(x^k, X^*)\}$ is also bounded. Suppose that there exists a positive constant ϵ such that

$$\mathrm{dist}(x^k, X^*) \geq \epsilon > 0, \qquad \text{for all } k.$$

Define

$$S = \{x \in X \,|\, \mathrm{dist}(x, X^*) \geq \epsilon,\ \|x - x^*\| \leq \|x^0 - x^*\|\}.$$

Then, $S \subset X \backslash X^*$ is a compact set and $\{x^k\} \subset S$. From Theorem 2.1, we know that there exists a positive constant $\delta \leq 1$ such that, for all $x \in S$ with $s(x) < 1$ and $\beta \in (0, \delta]$, (9) holds. Hence, for each k with $s(x^k) < 1$, we have

$$\beta_k \geq \min\{\alpha\delta, s(x^k)\}. \tag{21}$$

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

$$\{F(x^k) - F(\Pi_X[x^k - F(x^k)])\}^T E(x^k, 1) > 0, \qquad \eta(x^k) = \eta,$$

and

$$\begin{aligned} s(x^k) &= (1-\eta)\|E(x^k, 1)\|^2 / \{F(x^k) - F(\Pi_X[x^k - F(x^k)])\}^T E(x^k, 1) \\ &\geq (1-\eta)\|E(x^k, 1)\| / [\|F(x^k)\| + \|F(\Pi_X[x^k - F(x^k)])\|]. \end{aligned} \tag{22}$$

From the continuity of F and $\{x^k\} \subset S \subset X \backslash X^*$, we know that

$$\inf_k \|E(x^k, 1)\| > 0. \tag{23}$$

From (21)–(23), there exists a positive constant $\bar{\delta} \leq 1$ such that

$$\beta_k \geq \bar{\delta} > 0, \qquad \text{for all } k \text{ with } s(x^k) < 1.$$

If $s(x^k) = 1$, then $\beta_k = 1$. Hence,

$$1 \geq \beta_k \geq \bar{\delta} > 0, \qquad \text{for all } k. \tag{24}$$

Therefore,

$$\inf_k \psi(x^k, \beta_k)/\|g(x^k, \beta_k)\| = \epsilon_0 > 0,$$

which in light of (19) (note that we just take $\gamma_k = 1$) means that

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \epsilon_0^2.$$

Taking limits in both sides of the above inequality, we can derive a contradiction, since $\{\|x^k - x^*\|\}$ is a convergent sequence. So, we have

$$\liminf_{k \to \infty} \operatorname{dist}(x^k, X^*) = 0. \tag{25}$$

Furthermore, if (18) holds for each $x^* \in X^*$, we can conclude that there exists $\bar{x} \in X^*$ such that $x^k \to \bar{x}$, as $k \to \infty$, by the following argument. Since X^* is closed, (25) and the boundedness of $\{x^k\}$ mean that there exist $\bar{x} \in X^*$ and a subsequence $\{x^{k_j}\}$ such that $x^{k_j} \to \bar{x}$ as $j \to \infty$. Since $\{\|x^k - \bar{x}\|\}$ is a decreasing sequence and $x^{k_j} \to \bar{x}$ as $j \to \infty$, the whole sequence $\{x^k\}$ also converges to $\bar{x}$. □

When X is of the following form:

$$X = \{x \in R^n | l \leq x \leq u\}, \tag{26}$$

where l and u are two vectors of $\{R \cup \{-\infty, \infty\}\}^n$, we can give an improved form of the PC method. For any $x \in X$ and $\beta > 0$, denote

$$N(x, \beta) = \{i | (x_i = l_i \text{ and } (g(x, \beta))_i \geq 0) \text{ or } (x_i = u_i \text{ and } (g(x, \beta))_i \leq 0)\}. \tag{27}$$

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

$$(g_N(x, \beta))_i = \begin{cases} (g(x, \beta))_i, & \text{if } i \in N(x, \beta), \\ 0, & \text{otherwise,} \end{cases} \tag{28a}$$

$$(g_B(x, \beta))_i = (g(x, \beta))_i - (g_N(x, \beta))_i, \qquad i = 1, \ldots, n. \tag{28b}$$

Then, for any $x^* \in X^*$ and $x \in X$,

$$(x - x^*)^T g_N(x, \beta) \leq 0,$$

which means that

$$(x - x^*)^T g_B(x, \beta) \geq (x - x^*)^T g(x, \beta). \tag{29}$$

So, if in the PC methods we set

$$x^{k+1} = \Pi_X[x^k - \gamma_k \rho_k g_B(x^k, \beta_k)], \tag{30}$$

where

$$\rho_k = E(x^k, \beta_k)^T g(x^k, \beta_k)/\|g_B(x^k, \beta_k)\|^2,$$

then the convergence Theorems 3.1 and 3.2 hold for the improved PC method. In practice, we use the iterative form (30) whenever X is of the form (26). When X is a general convex polyhedral set, we cannot give similar definitions of g_N and g_B.

4. Search Directions

In this section, under some conditions, we give two forms of search directions which satisfy the assumptions of Theorems 3.1 and 3.2.

For any $\beta > 0$, define

$$g(x, \beta) = F(\Pi_X[x - \beta F(x)]) \tag{31}$$

or

$$g(x, \beta) = F(\Pi_X[x - \beta F(x)]) - F(x) + E(x, \beta)/\beta. \tag{32}$$

The form (31) is a modification of the extragradient (Ref. 5) and was used as a search direction by Sun (Refs. 16 and 17). The form (32) first appeared in Sun (Ref. 17) as an auxiliary vector function to obtain a new stepsize for the algorithm proposed by Sun (Ref. 16). Recently, Solodov and Tseng (Ref. 15) and He (Ref. 23) also considered the form (32) (including its extension) as a search direction to obtain a new globally convergent method for monotone variational inequalities. An Armijo-type inexact line search was also introduced in both Solodov and Tseng (Ref. 15) and He (Ref. 23), but no global convergence theorems were provided without assuming the Lipschitz continuity of F.[3]

Theorem 4.1. Suppose that F is continuous over X, that X^* is nonempty, and that $g(x, \beta)$ is of the form (31) or (32). If F is pseudomonotone over X relative to $x^* \in X^*$, and if there exists $\beta > 0$ such that (9) holds for some $x \in X \setminus X^*$, then

$$(x - x^*)^T g(x, \beta) \geq E(x, \beta)^T g(x, \beta) \geq \psi(x, \beta). \tag{33}$$

Furthermore, if F is pseudomonotone over X relative to X^*, then (33) holds for all $x^* \in X^*$.

[3] Such theorem was given by Solodov and Tseng in a revised version of their paper (Ref. 15).

Proof. Since F is pseudomonotone over X relative to $x^* \in X^*$, for all $z \in X$ we have

$$(z-x^*)^T F(z) \geq 0.$$

In particular, we have

$$\{\Pi_X[x-\beta F(x)]-x^*\}^T F(\Pi_X[x-\beta F(x)]) \geq 0. \tag{34}$$

First, we consider the case where $g(x, \beta)$ takes the form (31). Considering (34), we have

$$\begin{aligned}
&(x-x^*)^T g(x, \beta)\\
&=(x-x^*)^T F(\Pi_X[x-\beta F(x)])\\
&=E(x, \beta)^T g(x, \beta)+\{\Pi_X[x-\beta F(x)]-x^*\}^T F(\Pi_X[x-\beta F(x)])\\
&\geq E(x, \beta)^T g(x, \beta)\\
&=E(x, \beta)^T\{F(\Pi_X[x-\beta F(x)])-F(x)\}+E(x, \beta)^T F(x)\\
&\geq -[1-\eta(x)]\|E(x, \beta)\|^2/\beta+E(x, \beta)^T F(x),
\end{aligned}$$

where the last inequality follows from (9). By taking $z=x-\beta F(x)$ and $y=x$ in (i) of Lemma 2.2, we have

$$\beta E(x, \beta)^T F(x) \geq \|E(x, \beta)\|^2,$$

which means that

$$\begin{aligned}
(x-x^*)^T g(x, \beta) &\geq E(x, \beta)^T g(x, \beta)\\
&\geq -[1-\eta(x)]\|E(x, \beta)\|^2/\beta+\|E(x, \beta)\|^2/\beta\\
&=\eta(x)\|E(x, \beta)\|^2/\beta\\
&=\psi(x, \beta).
\end{aligned}$$

Next, we consider the case where $g(x, \beta)$ takes the form (32). By taking

$$z=x-\beta F(x) \quad \text{and} \quad y=x^*$$

in (i) of Lemma 2.2, we have

$$\{x^*-\Pi_X[x-\beta F(x)]\}^T\{x-\beta F(x)-\Pi_X[x-\beta F(x)]\} \leq 0.$$

By rearrangement, we have

$$(x-x^*)^T E(x, \beta) \geq \beta\{\Pi_X[x-\beta F(x)]-x^*\}^T F(x)+\|E(x, \beta)\|^2.$$

Therefore,

$$
\begin{aligned}
&(x-x^*)^T g(x,\beta)\\
&=(x-x^*)^T F(\Pi_X[x-\beta F(x)])\\
&-(x-x^*)^T F(x)+(x-x^*)^T E(x,\beta)/\beta\\
&\geq (x-x^*)^T F(\Pi_X[x-\beta F(x)])-(x-x^*)^T F(x)\\
&+\{\Pi_X[x-\beta F(x)]-x^*\}^T F(x)+\|E(x,\beta)\|^2/\beta\\
&=E(x,\beta)^T F(\Pi_X[x-\beta F(x)])\\
&+\{\Pi_X[x-\beta F(x)]-x^*\}F(\Pi_X[x-\beta F(x)])\\
&-E(x,\beta)^T F(x)+\|E(x,\beta)\|^2/\beta.
\end{aligned}
\tag{35}
$$

Substituting (34) into (35) gives

$$
\begin{aligned}
&(x-x^*)^T g(x,\beta)\\
&\geq E(x,\beta)^T F(\Pi_X[x-\beta F(x)])-E(x,\beta)^T F(x)+\|E(x,\beta)\|^2/\beta\\
&=E(x,\beta)^T g(x,\beta).
\end{aligned}
$$

Substituting (9) into the above formulas, we have

$$
\begin{aligned}
(x-x^*)^T g(x,\beta)\geq E(x,\beta)^T g(x,\beta)&\\
&=E(x,\beta)^T\{F(\Pi_X[x-\beta F(x)])-F(x)\}\\
&+\|E(x,\beta)\|^2/\beta\\
&\geq -[1-\eta(x)]\|E(x,\beta)\|^2/\beta+\|E(x,\beta)\|^2/\beta\\
&=\eta(x)\|E(x,\beta)\|^2/\beta\\
&=\psi(x,\beta).
\end{aligned}
$$

□

Remark 4.1. Assume that $F(x)=Mx+c$ and that M is skew-symmetric, i.e., $M^T=-M$. If $g(x,\beta)$ takes the form (31), then

$$\beta_k=1 \quad \text{and} \quad g(x_k,\beta_k)=M^T E(x^k,1)+(Mx^k+c),$$

which means that, for linear programming (translated into an equivalent linear complementarity problem), our method reduces to the same discussed by He (Ref. 10). If $g(x,\beta)$ takes the form (32), then

$$\beta_k=1 \quad \text{and} \quad g(x^k,\beta_k)=M^T E(x^k,1)+E(x^k,1),$$

which also appeared in He (Ref. 12).

Remark 4.2. Assume that F is Lipschitz continuous over X. Then, in the PC method, by taking the form (32), we can easily prove that ρ_k is bounded away from zero (under the Lipschitz continuity assumption, β_k is bounded away from zero); by taking the form (31), we cannot prove such conclusion.

Remark 4.3. From the last part of the proof of Theorem 4.1 we know that, under the conditions of Theorem 4.1, (33) holds for all $x \notin X^*$ if we take form (32). So, if in this case we set

$$x^{k+1} = x^k - \gamma_k \rho_k g(x^k, \beta_k), \tag{36}$$

then the convergence theorems also hold. Considering the forms (15) and (16), we know that, in (36), one projection step is saved per iteration. So, from the theoretical point of view, it is more suitable to use form (36). But according to our computational experience, we suggest to use the iterative forms (15) and (16) in practice.

5. Existence of Solutions

When F is continuous and pseudomonotone over X, there exist some results on the existence of the solutions of Eq. (1); see Harker and Pang (Ref. 2). Here, we give a necessary and sufficient condition on the existence of solutions.

Theorem 5.1. Suppose that $g(x, \beta)$ takes the form (31) or (32). If F is continuous and pseudomonotone over X, then $X^* \neq \emptyset$ if and only if some or any sequence $\{x^k\}$ generated by PC methods is bounded.

Proof. We just discuss the case where $g(x, \beta)$ takes the form (31). The proof on taking the form (32) is similar.

When $X^* \neq \emptyset$, then from Theorems 3.2 and 4.1, any sequence $\{x^k\}$ generated by the PC method is bounded.

For the converse part of the theorem, we suppose that there exists a bounded sequence $\{x^k\}$ generated by the PC method. From the boundedness of $\{x^k\}$ and the continuity of F, there exists a positive constant r such that

$$\|x^k\| \leq r, \quad \|F(x^k)\| \leq r, \qquad \text{for all } k.$$

From (ii) of Lemma 2.2, for all k and $\beta \in [0, 1]$, we have

$$\|\Pi_X[x^k - \beta F(x^k)]\| \leq 2r.$$

Choosing an arbitrary fixed vector $v \in X$, define

$$Y = \{x \in R^n | \, \|x\| \leq 2r + \|v\|\} \cap X.$$

Then, Y is a nonempty compact convex set and, for all k and $\beta \in [0, 1]$, we have

$$\Pi_Y[x^k - \beta F(x^k)] = \Pi_X[x^k - \beta F(x^k)], \qquad \text{for all } \beta \in [0, 1], \tag{37}$$

$$x^{k+1} = \Pi_X[x^k - \gamma_k \rho_k g(x^k, \beta_k)] = \Pi_Y[x^k - \gamma_k \rho_k g(x^k, \beta_k)]. \tag{38}$$

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

$$\bar{\eta}(x) = \begin{cases} \max\{\eta, 1 - \bar{t}(x)/\|E_Y(x, 1)\|^2\}, & \text{if } \bar{t}(x) > 0, \\ 1, & \text{otherwise,} \end{cases}$$

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

and

$$\bar{\psi}(x, \beta) = \bar{\eta}(x)\|E_Y(x, \beta)\|^2/\beta,$$

where

$$\bar{t}(x) = \{F(x) - F(\Pi_Y[x - F(x)])\}^T E_Y(x, 1).$$

For each k, if $\bar{s}(x^k) = 1$, let $\bar{\beta}_k = 1$; otherwise, determine $\bar{\beta}_k = \bar{s}(x^k)\alpha^{m_k}$, where m_k is the smallest nonnegative integer m such that

$$\{F(x^k) - F(\Pi_Y[x^k - \bar{s}(x^k)\alpha^m F(x^k)])\}^T E_Y(x^k, \bar{s}(x^k)\alpha^m)$$
$$\leq [1 - \bar{\eta}(x^k)]\|E_Y(x^k, \bar{s}(x^k)\alpha^m)\|^2/(\bar{s}(x^k)\alpha^m).$$

From (37), we know that

$$\bar{\eta}(x^k) = \eta(x^k), \qquad \bar{s}(x^k) = s(x^k), \tag{39}$$

and for all $\beta \in [0, 1]$,

$$E_Y(x^k, \beta) = E_X(x^k, \beta). \tag{40}$$

Therefore, for all k, we have

$$\bar{\beta}_k = \beta_k. \tag{41}$$

Define

$$\bar{g}(x, \beta) = F(\Pi_Y[x - \beta F(x)]),$$

$$\bar{\rho}_k = E_Y(x^k, \bar{\beta}_k)^T \bar{g}(x^k, \bar{\beta}_k)/\|\bar{g}(x^k, \bar{\beta}_k)\|^2.$$

Then from (37) and (39)–(41), we have

$$\bar{g}(x^k, \bar{\beta}_k) = g(x^k, \beta_k) \quad \text{and} \quad \bar{\rho}_k = \rho_k. \tag{42}$$

Hence, from (38) and (42), we have

$$\begin{aligned} x^{k+1} &= \Pi_X[x^k - \gamma_k \rho_k g(x^k, \beta_k)] \\ &= \Pi_Y[x^k - \gamma_k \rho_k g(x^k, \beta_k)] \\ &= \Pi_Y[x^k - \gamma_k \bar{\rho}_k \bar{g}(x^k, \bar{\beta}_k)], \end{aligned}$$

which means that $\{x^k\}$ can be regarded as a sequence generated by applying the PC method to solve

$$E_Y(x, 1) = 0. \tag{43}$$

Since Y is a nonempty compact convex subset of R^n, from Lemma 2.1 and Eaves (Ref. 1), we know that the solution set,

$$Y^* = \{y \in Y \mid y \text{ is a solution of } (43)\},$$

is nonempty. According to Theorems 3.2 and 4.1, there exists $x^* \in Y^*$ such that

$$x^k \to x^*, \qquad \text{as } k \to \infty.$$

Since $x^* \in Y^*$ and $v \in Y$, from Eaves (Ref. 1) we know that

$$F(x^*)^T(v - x^*) \geq 0.$$

Since v is an arbitrary fixed point of X and x^* is the limit point of $\{x^k\}$, we have

$$F(x^*)^T(x - x^*) \geq 0, \qquad \text{for all } x \in X.$$

which again from Eaves (Ref. 1) means that $E_X(x^*, 1) = 0$; i.e., X^* is nonempty and $x^* \in X^*$. □

Remark 5.1. When X is of the form (26), Theorem 5.1 also holds for the improved PC methods. The proof is similar and the details are omitted.

Remark 5.2. The procedure introduced here can be used to give a positive answer to an open problem proposed by He and Stoer (Ref. 14).

6. Numerical Experiments

In the following examples, we take $\eta = \alpha = 0.5$ and $\Delta_1 = \Delta_2 = 1.95$ (the algorithms behave better when γ_k approaches 2.0). We use $\varphi(x, 1) = F(x)^T E(x, 1) \leq \epsilon^2$ [note that $\varphi(x, 1) \geq \|E(x, 1)\|^2$, for all $x \in X$] as a stopping criterion, where ϵ is a small nonnegative number. The projection and contraction method for solving nonlinear projection equations with forms (31) and (32) is abbreviated as NPC1 and NPC2 respectively. The projection and contraction method for solving linear projection equations by He (Ref. 10) is abbreviated as LPC. In the above algorithms, we use the improved search direction $g_B(x, \beta)$ instead of $g(x, \beta)$. Here, we do not list the results of EG and IEG methods, since we know from Sun (Ref. 16) that in general EG and IEG methods cannot compete with PC methods.

Example 6.1. This example, discussed by Ahn (Ref. 24), is of the form $F(x)=Dx+c$, where c is an n-vector, D is an $n\times n$ nonsymmetric matrix,

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

Take

$$X=[l, u], \qquad l=(0, 0, \ldots, 0)^T, \qquad u=(1, 1, \ldots, 1)^T.$$

Take $\epsilon^2=n10^{-14}$, where n is the dimension of the problem. See Table 1.

Example 6.2. This problem was discussed by Sun (Ref. 16). Consider

$$F(x)=F_1(x)+F_2(x), \qquad x=(x_1, \ldots, x_n)^T, \qquad x_0=x_{n+1}=0,$$

$$F_1(x)=[f_1(x), \ldots, f_n(x)]^T, \qquad F_2(x)=Dx+c,$$

$$f_i(x)=x_{i-1}^2+x_i^2+x_{i-1}x_i+x_ix_{i+1}, \qquad i=1, \ldots, n,$$

and D and c are the same as those of Example 6.1. Take

$$X=[l, u], \qquad l=(0, 0, \ldots, 0)^T, \qquad u=(1, 1, \ldots, 1)^T.$$

Take $\epsilon^2=\eta 10^{-14}$, where n is the dimension of the problem. See Table 2.

Table 1. Results for Example 6.1, starting point (0, 0, . . . , 0).

Algorithms	Number of iterations (left) and inner iterations (right)									
	$n=10$		$n=50$		$n=100$		$n=200$		$n=500$	
LPC	39	—	39	—	39	—	39	—	39	—
NPC1	19	13	16	6	15	5	17	9	16	11
NPC2	16	8	17	11	14	4	14	4	13	4

Table 2. Results for Example 6.2, starting point (0, 0, . . . , 0).

Algorithms	Number of iterations (left) and inner iterations (right)									
	$n=10$		$n=50$		$n=100$		$n=200$		$n=500$	
NPC1	9	0	9	0	9	0	9	0	10	2
NPC2	9	0	9	0	9	0	10	0	10	0

7. Discussion

In this paper, a class of globally convergent algorithms for solving nonlinear projection equations (1) is provided. Here, the convergence rate of the given methods is not discussed, since we think that the best convergence rate is Q-linear. The basic reason for this is that the derivative of F is not assumed. However, the methods given here can converge to the neighborhood of the solution set very fast. In practice, when the iterative point is far away from the solution set, the PC methods can be used to make the iterative sequence reach the neighborhood of the solution set; when the iterative sequence approaches the solution set close enough, more rapid locally convergent methods, such as the Newton and quasi-Newton methods, can be used. For the Newton and quasi-Newton methods for solving Eqs. (1), see Ref. 4 and references therein for details.

In Section 4, two forms of search directions are given to satisfy the requirements. In fact, more search directions can be given. For example, a convex combination of the forms (31) and (32) is also a suitable choice. For various forms of the search directions for solving linear projection equations, see He (Refs. 10–13) and He and Stoer (Ref. 14). As an extension of the search directions used in this paper, we can set

$$g^{\text{new}}(x, \beta) = G^{-1}g(x, \beta),$$

where G is an arbitrary symmetric positive-definite matrix. To choose a suitable G is useful, but difficult in theory.

From the computational experiments presented here, there is not too much difference between choosing (31) and (32). But, when F is Lipschitz continuous over X, the stepsize is bounded away from zero if we take the form (32), and this result does not hold for the form (31).

References

1. EAVES, B. C., *On the Basic Theorem of Complementarity*, Mathematical Programming, Vol. 1, pp. 68–75, 1971.
2. HARKER, P. T., and PANG, J. S., *Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms, and Applications,* Mathematical Programming, Vol. 48, pp. 161–220, 1990.
3. PANG, J. S., and CHAN, D., *Iterative Methods for Variational and Complementarity Problems*, Mathematical Programming, Vol. 24, pp. 284–313, 1982.
4. PANG, J. S., and QI, L., *Nonsmooth Equations: Motivation and Algorithms,* SIAM Journal on Optimization, Vol. 3, pp. 443–465, 1993.
5. KORPLEVICH, G. M., *Ekstragradientnyi Metod dlia Otyskaniia Sedlovykh Tchek i Drugikh Zadach*, Ekonomica i Matematicheski Metody, Vol. 12, pp. 947–956, 1976.
6. KHOBOTOV, E. N., *Modification of the Extragradient Method for the Solution of Variational Inequalities and Some Optimization Problems*, Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki, Vol. 27, pp. 1162–1473, 1987.
7. MARCOTTE, P., *Application of Khobotov's Algorithm to Variational Inequalities and Network Equilibrium Problems*, Information Systems and Operations Research, Vol. 29, pp. 258–270, 1991.
8. SUN, D., *An Iterative Method for Solving Variational Inequality Problems and Complementarity Problems*, Numerical Mathematics: Journal of Chinese Universities, Vol. 16, pp. 145–153, 1994.
9. FUKUSHIMA, M., *Equivalent Differentiable Optimization Problems and Descent Methods for Asymmetric Variational Inequality Problems*, Mathematical Programming, Vol. 53, pp. 99–110, 1992.
10. HE, B., *A Projection and Contraction Method for a Class of Linear Complementarity Problems and Its Application in Convex Quadratic Programming*, Applied Mathematics and Optimization, Vol. 25, pp. 247–262, 1992.
11. HE, B., *On a Class of Iterative Projection and Contraction Methods for Linear Programming*, Journal of Optimization Theory and Applications, Vol. 78, pp. 247–266, 1993.
12. HE, B., *Solving a Class of Linear Projection Equations*, Numerische Mathematik, Vol. 69, pp. 71–80, 1994.
13. HE, B., *A New Method for a Class of Linear Variational Inequalities*, Mathematical Programming, Vol. 66, pp. 137–144, 1994.
14. HE, B., and STOER, J., *Solutions of Projection Problems over Polytopes*, Numerische Mathematik, Vol. 61, pp. 73–90, 1992.
15. SOLODOV, M. V., and TSENG, P., *Modified Projection-Type Methods for Monotone Variational Inequalities*, SIAM Journal on Control and Optimization, Vol. 34, 1996 (to appear).

16. SUN, D., *A Projection and Contraction Method for the Nonlinear Complementarity Problem and Its Extensions*, Mathematica Numerica Sinica, Vol. 16, pp. 183–194, 1994.
17. SUN, D., *A New Stepsize Skill for Solving a Class of Nonlinear Projection Equations*, Journal of Computational Mathematics, Vol. 13, pp. 357–368, 1995.
18. LUO, Z. Q., and TSENG, P., *On the Linear Convergence of Descent Methods for Convex Essentially Smooth Minimization*, SIAM Journal on Control and Optimization, Vol. 30, pp. 408–425, 1992.
19. MORÉ, J. J., *Coercivity Conditions in Nonlinear Complementarity Problems*, SIAM Review, Vol. 16, pp. 1–16, 1974.
20. ZARANTONELLO, E. H., *Projections on Convex Sets in Hilbert Space and Spectral Theory*, Contributions to Nonlinear Functional Analysis, Edited by E. H. Zarantonello, Academic Press, New York, pp. 237–424, 1971.
21. GAFNI, E. H., and BERTSEKAS, D. P., *Two-Metric Projection Methods for Constrained Optimization*, SIAM Journal on Control and Optimization, Vol. 22, pp. 936–964, 1984.
22. CALAMAI, P. H., and MORÉ, J. J., *Projected Gradient Method for Linearly Constrained Problems*, Mathematical Programming, Vol. 39, pp. 93–116, 1987.
23. HE, B., *A Class of Projection and Contraction Methods for Monotone Variational Inequalities*, Applied Mathematics and Optimization (to appear).
24. AHN, B. H., *Iterative Methods for Linear Complementarity Problem with Upper Bounds and Lower Bounds*, Mathematical Programming, Vol. 26, pp. 295–315, 1983.