PERGAMON Computers and Mathematics with Applications 42 (2001) 969–979

An International Journal
computers & mathematics with applications

www.elsevier.nl/locate/camwa

A New Version of Extragradient Method for Variational Inequality Problems

YIJU WANG
Institute of Operations Research, Qufu Normal University
Qufu, Shandong 273165, China
and
Institute of Applied Mathematics, Chinese Academy of Sciences
Beijing 100080, China

NAIHUA XIU
Department of Applied Mathematics, Northern Jiaotong University
Beijing 100044, China

CHANGYU WANG
Institute of Operations Research, Qufu Normal University
Qufu, Shandong 273165, China
and
Institute of Applied Mathematics, Chinese Academy of Sciences
Beijing 100080, China

(Received June 2000; accepted July 2000)

Abstract—In this paper, we propose a new version of extragradient method for the variational inequality problem. The method uses a new searching direction which differs from any one in existing projection-type methods, and is of a better step-size rule. Under a certain generalized monotonicity condition, it is proved to be globally convergent. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords—Variational inequality, Extragradient method, Convergence.

1. INTRODUCTION

Let C be a closed convex set in R^n and F be a continuous mapping from R^n to itself. The variational inequality problem, denoted by $\mathrm{VI}(F,C)$, is to find a vector $x^* \in C$ such that

$$\langle F(x^*), x - x^* \rangle \geq 0, \qquad \forall x \in C,$$

where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in R^n. We denote the solution set of $\mathrm{VI}(F,C)$ by C^*. Variational inequality problem plays a significant role in economics, engineering mechanics, mathematical programming, transportation, etc. It has received considerable attention and many numerical algorithms for solving it have been constructed, see e.g., [1–3].

This work is supported by the Natural Science Foundation of Shandong Province, China (Grant No. Q99A11) and the Natural Science Foundation of China (Grant No. 19971002,19871049). The authors would like to thank Associate Editor R. Glowinski for his helpful comments and suggestions on this paper.

PII: S0898-1221(01)00213-9

On account of projected gradient method for constrained optimization, Korpelevich [4] proposed an extragradient method with iterative scheme

$$\bar{x}^k = P_C\left(x^k - \alpha_k F\left(x^k\right)\right),$$
$$x^{k+1} = P_C\left(x^k - \alpha_k F\left(\bar{x}^k\right)\right),$$

where $P_C(\cdot)$ is an orthogonal projection onto C, and $\alpha_k > 0$ is step-size. But its convergence requires Lipschitz continuity of F. When F is not Lipschitz continuous or the Lipschitz constant is not known, the extragradient method requires an Armijo-type line-search procedure to obtain step-size with a new projection needed for each trial point (see e.g., [5–10]), and this can be very computationally expensive. To overcome this defect, Iusem and Svaiter [11] proposed a modified extragradient method which requires only one projection to obtain step-size α_k. The condition under which global convergence is guaranteed is not enhanced. Solodov and Svaiter [12] gave an improvement of such a method by modifying the projection region. They also reported an encouraging computational experience. Recently, the authors [13] showed an interesting fact that the projection-type method in [12] has the same searching direction as the modified extragradient method in [11], but it uses a better step-size rule.

In this paper, we propose a new version of extragradient method for the solution of problem VI(F, C). The searching direction in this method is a combination of the projection residue and the modified extra-gradient direction in [11,12], and differs from any one in existing projection-type methods (such as [14–18]). The step-size in this method is chosen so that the distance between the new iterative point and the solution set has a larger decrease. In Section 3, we state our new method and prove that under a weaker condition than the monotonicity, the new method is globally convergent.

2. PRELIMINARIES

Let Ω be a subset in R^n, projection from $x \in R^n$ onto Ω is defined by

$$P_\Omega(x) = \arg\min\{\|y - x\| \mid y \in \Omega\},$$

where $\|\cdot\|$ is l_2-norm in R^n. The projection operator has been extensively studied, and we here list some properties of it.

LEMMA 2.1. *Let Ω be a nonempty closed convex subset in R^n, then for any $x, y \in R^n$ and $z \in \Omega$,*

(1) $\langle P_\Omega(x) - x, z - P_\Omega(x)\rangle \geq 0$;
(2) $\langle P_\Omega(x) - P_\Omega(y), x - y\rangle \geq 0$;
(3) $\|P_\Omega(x) - P_\Omega(y)\| \leq \|x - y\|$;
(4) $\|P_\Omega(x) - z\|^2 \leq \|x - z\|^2 - \|P_\Omega(x) - x\|^2$.

LEMMA 2.2. *Let Ω be a nonempty closed convex subset in R^n. For any $x, d \in R^n$ and $\alpha \geq 0$, define $x(\alpha) = P_\Omega(x + \alpha d)$. Then $\langle d, x - x(\alpha)\rangle$ is nonincreasing for $\alpha > 0$.*

The following property was proved in [19].

LEMMA 2.3. *Let Ω be a nonempty closed convex subset in R^n. For any $x \in \Omega$, $d \in R^n$, and $\alpha \geq 0$, define*

$$\psi(\alpha) = \min\left\{\|y - x - \alpha d\|^2 \mid y \in \Omega\right\}.$$

Then $\psi'(\alpha) = 2\langle d, x + \alpha d - x(\alpha)\rangle$.

Throughout the paper, we assume that

($\mathbf{A}_1$) C^* is nonempty;
($\mathbf{A}_2$) for each $x^* \in C^*$, $\langle F(x), x - x^*\rangle \geq 0$, $\forall x \in C$.

It is easy to see that when F is monotone or pseudo-monotone, (A_2) holds. So, this is a weaker assumption.

For $x \in C$ and $\alpha > 0$, define $r(x, \alpha) = [x - P_C(x - \alpha F(x))]/\alpha$, and $r(x) = r(x, 1)$. They are called the scaled projection residue and the projection residue for VI(F, C), respectively. We also use the following well-known result.

LEMMA 2.4. *For VI(F, C), $x^* \in C^*$ if and only if $r(x^*, \alpha) = 0$ for some $\alpha > 0$.*

3. ALGORITHMS AND CONVERGENCE

In this section, we will give two modified extragradient algorithms for solving the variational inequality problem VI(F, C). They are the same in the sense that the same iterative sequence is generated but the second algorithm is easier to implement than the first one. Their convergence properties are developed under the Assumptions (A_1) and (A_2).

ALGORITHM NVE-1.

Initial Step: Select any $\sigma, \gamma \in (0, 1)$, $x^0 \in C$. $k = 0$.

Iterative Step: For $x^k \in C$, define

$$z^k = P_C\left(x^k - F\left(x^k\right)\right).$$

If $r(x^k) = 0$, then stop. Otherwise, compute

$$y^k = (1 - \eta_k)x^k + \eta_k z^k,$$

where $\eta_k = \gamma^{m_k}$ with m_k being the smallest nonnegative integer m satisfying

$$\left\langle F\left(x^k\right) - F\left(x^k - \gamma^m r\left(x^k\right)\right), r\left(x^k\right)\right\rangle \le \sigma \|r(x^k)\|^2. \tag{3.1}$$

Let

$$d_k = -\left(\frac{r\left(x^k\right) + F\left(y^k\right)}{\eta_k}\right),$$

$$\alpha_k^1 = \frac{\left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle}{\|d_k\|^2}.$$

Select $\alpha_k \ge \alpha_k^1$ such that

$$\left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left\langle x^k - x^k(\alpha_k), d_k\right\rangle = 0, \tag{3.2}$$

where $x^k(\alpha) := P_C(x^k + \alpha d_k)$. Let

$$x^{k+1} = P_C\left(x^k + \alpha_k d_k\right).$$

In order to get a better understanding of Algorithm NVE-1, we give the following analysis.

First, we analyze the step-size rule given in (3.1). If x^k is not a solution of VI(F, C), then $r(x^k) \neq 0$. By continuity of F, η_k satisfying (3.1) must exist. From (3.1) and $\langle F(x^k), r(x^k)\rangle \ge \|r(x^k)\|^2$, we know that

$$\begin{aligned} \left\langle F\left(y^k\right), r\left(x^k\right)\right\rangle &\ge \left\langle F\left(x^k\right), r\left(x^k\right)\right\rangle - \sigma\left\|r\left(x^k\right)\right\|^2 \\ &\ge (1 - \sigma)\left\|r\left(x^k\right)\right\|^2, \end{aligned} \tag{3.3}$$

which shows that the step-size rule (3.1) is related closely to the ones in [11,12]. Also, if $F(x)$ is Lipschitz continuous with Lipschitz constant $L > 0$ on C, then η_k for all $k \ge 1$ have a positive bound from below. In fact, by (3.1) we have for any $k \ge 1$, if $\eta_k \neq 1$,

$$\frac{L}{\gamma}\eta_k\left\|r\left(x^k\right)\right\|^2 \ge \left\langle F\left(x^k\right) - F\left(x^k - \left(\frac{\eta_k}{\gamma}\right) r\left(x^k\right)\right), r(x^k)\right\rangle > \sigma\left\|r\left(x^k\right)\right\|^2,$$

i.e., $\eta_k \ge (\sigma\gamma)/L > 0$.

Second, we observe the searching direction d_k. We recall the searching directions that appear in existing projection-type methods for solving problem VI(F, C). They are

(i) the direction $-\alpha_k F(\bar{x}^k)$ by Korpelevich [4];
(ii) the direction $-\{x^k - \bar{x}^k - \alpha_k[F(x^k) - F(\bar{x}^k)]\}$ by He [15], Solodov and Tseng [17], Sun [10], and Tseng [18];
(iii) the direction $-\{x^k - \bar{x}^k + F(\bar{x}^k)\}$ by Noor [16];
(iv) the direction $-F(y^k)$ by Iusem and Svaiter [11] and Solodov and Svaiter [12].

In our algorithm, the searching direction is taken as

$$-\left(r\left(x^k\right) + \frac{F\left(y^k\right)}{\eta_k}\right) = -\frac{1}{\eta_k}\left(x^k - y^k + F\left(y^k\right)\right).$$

It is a combination of the projection residue and the modified extra-gradient direction in [11,12], and differs from any one of the above four directions. (Note that y^k is not different from $\bar{x}^k$.) I wonder whether other searching direction could be generated if we combine some searching directions which have good behavior. This is a topic for further research.

Finally, we discuss the feasibility of the step-size rule given in (3.2). From the iterative procedure of Algorithm NVE-1, we know that $x^k, y^k, z^k \in C$, for all k. For any $x^* \in C^*$, by (A_2) and Lemma 2.1, we have

$$\left\langle F\left(x^k\right), x^k - x^*\right\rangle \geq 0$$

and

$$\left\langle x^k - F\left(x^k\right) - z^k, z^k - x^*\right\rangle \geq 0.$$

Adding the above two inequalities, we obtain

$$\left\langle x^k - F\left(x^k\right) - z^k, z^k - x^k\right\rangle + \left\langle x^k - z^k, x^k - x^*\right\rangle \geq 0.$$

Since

$$\begin{aligned}\left\langle x^k - x^*, \frac{F\left(y^k\right)}{\eta_k}\right\rangle &= \left\langle x^k - y^k + y^k - x^*, \frac{F\left(y^k\right)}{\eta_k}\right\rangle \\ &\geq \left\langle x^k - y^k, \frac{F\left(y^k\right)}{\eta_k}\right\rangle \\ &= \left\langle x^k - z^k, F\left(y^k\right)\right\rangle,\end{aligned}$$

we have

$$\begin{aligned}\left\langle x^k - x^*, r\left(x^k\right) + \frac{1}{\eta_k} F\left(y^k\right)\right\rangle &= \left\langle x^k - x^*, r\left(x^k\right)\right\rangle + \left\langle x^k - x^*, \frac{1}{\eta_k} F\left(y^k\right)\right\rangle \\ &\geq \left\langle x^k - F\left(x^k\right) - z^k, x^k - z^k\right\rangle + \left\langle x^k - z^k, F\left(y^k\right)\right\rangle \\ &= \left\langle r\left(x^k\right), r\left(x^k\right) + F\left(y^k\right) - F\left(x^k\right)\right\rangle.\end{aligned} \tag{3.4}$$

For $\alpha \geq 0$, by Lemma 2.1 and (3.4), we have

$$\begin{aligned}\left\|x^k(\alpha) - x^*\right\|^2 &= \left\|P_C\left(x^k + \alpha d_k\right) - x^*\right\|^2 \\ &\leq \left\|x^k - x^* + \alpha d_k\right\|^2 - \left\|x^k - x^k(\alpha) + \alpha d_k\right\|^2 \\ &= \left\|x^k - x^*\right\|^2 + \alpha^2\|d_k\|^2 + 2\alpha\left\langle d_k, x^k - x^*\right\rangle - \left\|x^k - x^k(\alpha) + \alpha d_k\right\|^2 \\ &\leq \left\|x^k - x^*\right\|^2 + \alpha^2\|d_k\|^2 - 2\alpha\left\langle r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right), r\left(x^k\right)\right\rangle \\ &\quad - \left\|x^k - x^k(\alpha) + \alpha d_k\right\|^2.\end{aligned}$$

Denote

$$\phi_k(\alpha) = -\alpha^2\|d_k\|^2 + 2\alpha\left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left\|x^k - x^k(\alpha) + \alpha\, d_k\right\|^2, \quad (3.5)$$

then

$$\left\|x^k(\alpha) - x^*\right\|^2 \le \left\|x^k - x^*\right\|^2 - \phi_k(\alpha). \quad (3.6)$$

By Lemma 2.3, we have

$$\phi_k'(\alpha) = 2\left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + 2\left\langle x^k - x^k(\alpha), d_k\right\rangle.$$

To show that the step-size rule for α_k is implementable, it is sufficient to show that there exists $\alpha_k \ge \alpha_k^1$ such that $\phi_k'(\alpha_k) = 0$. Since

$$\left\langle x^k(\alpha) - x^k, d_k\right\rangle \le \alpha\|d_k\|^2,$$

so if $\alpha < \alpha_k^1 = (\langle r(x^k), r(x^k) - F(x^k) + F(y^k)\rangle)/(\|d_k\|^2)$, then $\phi_k'(\alpha) > 0$. That is to say, the smallest nonnegative solution to the equation $\phi_k'(\alpha) = 0$, say $\overline{\alpha}_k$, satisfies $\overline{\alpha}_k \ge \alpha_k^1$.

Consider the optimization problem

$$\max\{\phi_k(\alpha) \mid \alpha \ge 0\}.$$

Since $\phi_k'(\alpha)$ is nonincreasing for $\alpha > 0$ by Lemma 2.2, and for $x^k \notin C^*$,

$$\phi_k(0) = 2\left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle \ge 2(1-\sigma)\left\|r\left(x^k\right)\right\|^2 > 0,$$

we know that if the equation $\phi_k'(\alpha) = 0$ is solvable, then every solution to this equation is a solution to the problem $\max\{\phi_k(\alpha) \mid \alpha \ge 0\}$. The following lemma shows that the equation $\phi_k'(\alpha) = 0$ is solvable, which implies that α_k in Algorithm NVE-1 is well defined.

LEMMA 3.1. *For function $\phi_k(\alpha)$ defined by (3.5), the equation $\phi_k'(\alpha) = 0$ is solvable.*

PROOF. Define the hyperplane

$$H_k = \left\{x \in R^n \mid \left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left\langle x^k - x, d_k\right\rangle = 0\right\}.$$

For $x^k \notin C^*$, we have

$$\left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle \ge (1-\sigma)\left\|r\left(x^k\right)\right\|^2 > 0$$

and from (3.3), we obtain

$$\begin{aligned}
&\left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left\langle r\left(x^k\right), d_k\right\rangle \\
&= \left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left\langle -r\left(x^k\right), r\left(x^k\right) + \frac{F\left(y^k\right)}{\eta_k}\right\rangle \\
&= \left\langle r\left(x^k\right), -F\left(x^k\right) + F\left(y^k\right) - \frac{F\left(y^k\right)}{\eta_k}\right\rangle \\
&= -\left\langle r\left(x^k\right), F\left(x^k\right)\right\rangle - \left\langle r\left(x^k\right), \frac{1-\eta_k}{\eta_k} F\left(y^k\right)\right\rangle \\
&\le -\left\|r\left(x^k\right)\right\|^2 - \frac{1-\eta_k}{\eta_k}(1-\sigma)\left\|r\left(x^k\right)\right\|^2 \\
&\le -\left\|r\left(x^k\right)\right\|^2 < 0,
\end{aligned}$$

which implies that $x^k \in \{x \in R^n \mid \langle r(x^k), r(x^k) - F(x^k) + F(y^k)\rangle + \langle x^k - x, d_k\rangle \geq 0\}$ and $z^k \in \{x \in R^n \mid \langle r(x^k), r(x^k) - F(x^k) + F(y^k)\rangle + \langle x^k - x, d_k\rangle \leq 0\}$. From the convexity of C, we know that $H_k \cap C \neq \phi$. It is easy to verify that $x^k + \alpha_k^1 d_k \in H_k$ and

$$x^k + \alpha\, d_k \in \left\{x \in R^n \mid \left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left\langle x^k - x, d_k\right\rangle < 0\right\}, \qquad \forall\, \alpha > \alpha_k^1.$$

Obviously, $d_k \perp H_k$.

Let P be any point in $C \cap \{x \in R^n \mid \langle r(x^k), r(x^k) - F(x^k) + F(y^k)\rangle + \langle x^k - x, d_k\rangle \geq 0\}$, $Q = z^k$, and $O(\alpha) = x^k + \alpha\, d_k$ where $\alpha > 0$. In the triangle composed by points $O(\alpha)$, Q, and P, the inner corner at points P, Q are denoted by β_P and β_Q, respectively. By the knowledge of geometry, we know that when $\alpha > \alpha_k^1$ is sufficiently large, then $\beta_P < \beta_Q$, and so $\|O(\alpha)Q\| < \|O(\alpha)P\|$. From the arbitrariness of $P \in C \cap \{x \in R^n \mid \langle r(x^k), r(x^k) - F(x^k) + F(y^k)\rangle + \langle x^k - x, d_k\rangle \geq 0\}$ and by the definition of orthogonal projection, there exists $\alpha_k' > \alpha_k^1$ such that

$$P_C\left(x^k + \alpha_k' d_k\right) \in \left\{x \in R^n \mid \left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left\langle x^k - x, d_k\right\rangle \leq 0\right\}.$$

On the other hand,

$$P_C\left(x^k + 0 \cdot d_k\right) = x^k \in \left\{x \in R^n \mid \left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left\langle x^k - x, d_k\right\rangle > 0\right\}.$$

By continuity of orthogonal projection operator, we know that there exists $\alpha_k \in (0, \alpha_k')$ such that $x^k(\alpha_k) = P_C(x^k + \alpha_k\, d_k) \in H_k \cap C$, which implies that the equation $\phi_k'(\alpha) = 0$ is solvable. ∎

Based on the above analysis, we know that Algorithm NVE-1 is implementable. Next, we state the convergence result and its proof.

THEOREM 3.1. *Suppose (A_1) and (A_2) hold. If the sequence $\{x^k\}$ generated by Algorithm NVE-1 is infinite, then $\{x^k\}$ globally converges to a solution x^* of VI(F, C).*

PROOF. Since α_k is a solution of $\max\{\phi_k(\alpha) \mid \alpha \geq 0\}$, by (3.6), we know that

$$\begin{aligned}
\left\|x^{k+1} - x^*\right\|^2 &\leq \left\|x^k - x^*\right\|^2 - \phi_k(\alpha_k) \\
&\leq \left\|x^k - x^*\right\|^2 - \phi_k\left(\alpha_k^1\right) \\
&= \left\|x^k - x^*\right\|^2 - 2\alpha_k^1 \left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle \\
&\quad + \left(\alpha_k^1\right)^2 \|d_k\|^2 - \left\|x^k - x^k(\alpha_k^1) - \alpha_k^1 d_k\right\|^2 \\
&\leq \left\|x^k - x^*\right\|^2 - 2\alpha_k^1 \left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle + \left(\alpha_k^1\right)^2 \|d_k\|^2 \\
&= \left\|x^k - x^*\right\|^2 - \frac{\left\langle r\left(x^k\right), r\left(x^k\right) - F\left(x^k\right) + F\left(y^k\right)\right\rangle^2}{\|d_k\|^2} \\
&\leq \left\|x^k - x^*\right\|^2 - \frac{(1-\sigma)\left\|r\left(x^k\right)\right\|^4}{\|d_k\|^2}.
\end{aligned}$$

So $\{x^k\}$ is a Fejér sequence with respect to C^*, and $\{x^k\}$ is a bounded sequence, so are $\{y^k\}$, $\{z^k\}$, $\{F(x^k)\}$, $\{F(y^k)\}$, respectively. There exist infinite subsets N_1 and N_2 in $\{0, 1, 2, \dots\}$ such that

$$\lim_{k \in N_1, k \to \infty} \left\|r\left(x^k\right)\right\| = 0 \qquad \text{or} \qquad \lim_{k \in N_2, k \to \infty} \eta_k = 0.$$

If $\lim_{k \in N_1, k \to \infty} \|r(x^k)\| = 0$, by Lemma 2.4, we know that any cluster $\tilde{x}$ of $\{x^k : k \in N_1\}$ is a solution to VI(F, C). Since $\{x^k\}$ is a Fejér sequence with respect to C^*, if we take $x^* = \tilde{x}$, then we know that $\{x^k\}$ globally converges to $\tilde{x}$. Otherwise, by (3.1) we know that

$$\left\langle F\left(x^k\right) - F\left(x^k - \left(\frac{\eta_k}{\gamma}\right) r\left(x^k\right)\right), r\left(x^k\right)\right\rangle > \sigma \left\|r\left(x^k\right)\right\|^2, \qquad \text{for } k \in N_2.$$

Therefore,

$$\left\| F\left(x^k\right) - F\left(x^k - \left(\frac{\eta_k}{\gamma}\right) r\left(x^k\right)\right) \right\| > \sigma \left\| r\left(x^k\right) \right\|, \qquad \text{for } k \in N_2.$$

This plus $\lim_{k\in N_2, k\to\infty} \eta_k/\gamma = 0$, yield $\lim_{k\in N_2,k\to\infty} \|r(x^k)\| = 0$. Similar discussion leads to that any cluster of $\{x^k : k \in N_2\}$ is a solution to VI(F, C). Replacing x^* by this cluster point yields the desired result. ∎

Obviously, α_k given in Algorithm NVE-1 is a long step (by $\alpha_k \geq \alpha_k^1$) and guarantees that the distance between the new iterative point and the solution set has a larger decrease. However, in practice, if C does not possess any special structure, it is difficult to give an explicit formula of α_k. That is to say, we need to find a simple way to compute the projection $P_C(x^k + \alpha_k d_k)$. The following lemma gives an answer to this question.

LEMMA 3.2. *For $\alpha_k \geq \alpha_k^1$ determined by (3.2), $x^k(\alpha_k) = P_{C\cap H_k}(x^k + \alpha_k^1 d_k)$.*

PROOF. Denote $\overline{x}^k(\alpha_k^1) = P_{C\cap H_k}(x^k + \alpha_k^1 d_k)$. It can easily be verified that

$$P_{H_k}\left(x^k\right) = x^k + \alpha_k^1 d_k,$$

i.e., $(x^k + \alpha_k^1 d_k) \in H_k$. From $x^k(\alpha_k) \in H_k$, we know that

$$x^k(\alpha_k) = P_{C\cap H_k}\left(x^k + \alpha_k d_k\right).$$

Since

$$\left(x^k + \alpha_k d_k\right) - \left(x^k + \alpha_k^1 d_k\right) = \left(\alpha_k - \alpha_k^1\right) d_k \perp H_k,$$

by Pythagoras Theorem, we have

$$\left\|x^k + \alpha_k^1 d_k - \overline{x}^k\left(\alpha_k^1\right)\right\|^2 + \left\|\left(\alpha_k - \alpha_k^1\right) d_k\right\|^2 = \left\|x^k + \alpha_k d_k - \overline{x}^k\left(\alpha_k^1\right)\right\|^2, \tag{3.7}$$

$$\left\|x^k + \alpha_k^1 d_k - x^k(\alpha_k)\right\|^2 + \left\|\left(\alpha_k - \alpha_k^1\right) d_k\right\|^2 = \left\|x^k + \alpha_k d_k - x^k(\alpha_k)\right\|^2. \tag{3.8}$$

From (3.7) and (3.8), we have

$$\begin{aligned} &\left\|x^k + \alpha_k^1 d_k - \overline{x}^k\left(\alpha_k^1\right)\right\|^2 - \left\|x^k + \alpha_k^1 d_k - x^k(\alpha_k)\right\|^2 \\ &= \left\|x^k + \alpha_k d_k - \overline{x}^k\left(\alpha_k^1\right)\right\|^2 - \left\|x^k + \alpha_k d_k - x^k(\alpha_k)\right\|^2. \end{aligned} \tag{3.9}$$

By the definition of orthogonal projection, we know that

$$\left\|x^k + \alpha_k d_k - \overline{x}^k\left(\alpha_k^1\right)\right\| \geq \left\|x^k + \alpha_k d_k - x^k(\alpha_k)\right\| \tag{3.10}$$

and

$$\left\|x^k + \alpha_k^1 d_k - x^k(\alpha_k)\right\| \geq \left\|x^k + \alpha_k^1 d_k - \overline{x}^k\left(\alpha_k^1\right)\right\|. \tag{3.11}$$

Combining (3.9)–(3.11), and by the uniqueness of orthogonal projection, the desired result is obtained. ∎

Thus, Algorithm NVE-1 can be rewritten as the following form.

ALGORITHM NVE-2.

Iterative Step: For $x_k \in C$, define

$$z^k = P_C\left(x^k - F\left(x^k\right)\right).$$

If $r(x^k) = 0$, then stop. Otherwise, compute

$$y^k = (1 - \eta_k)x^k + \eta_k z^k,$$

where $\eta_k = \gamma^{m_k}$ with m_k being the smallest nonnegative integer m satisfying

$$\langle F(x^k) - F(x^k - \gamma^m r(x^k)), r(x^k)\rangle \le \sigma \|r(x^k)\|^2.$$

Let

$$x^{k+1} = P_{C\cap H^k}(x^k + \alpha_k^1 d_k),$$

where

$$d_k = -\left(r(x^k) + \frac{F(y^k)}{\eta_k}\right),$$

$$\alpha_k^1 = \frac{\langle r(x^k), r(x^k) - F(x^k) + F(y^k)\rangle}{\|d_k\|^2}.$$

It is easy to see that Algorithms NVE-1 and NVE-2 share the same convergence result.

THEOREM 3.2. *Suppose that* (A_1) *and* (A_2) *hold. If the sequence* $\{x^k\}$ *generated by Algorithm NVE-2 is infinite, then it converges to a solution of VI*(F, C).

APPENDIX
PRELIMINARY NUMERICAL EXPERIMENTS

To give some insight into the behavior of our new algorithm, we implemented it in MATLAB to solve linear constrained variational inequality problems (by solving the quadratic program to perform the projection). We compared the performance of this implementation with analogous implementations of the methods described in [7–10,12,17]. By contrast, our algorithm seems to perform better than the alternatives in many cases. In Algorithm NVE-2, two projections onto C and $C \cap H_k$ are needed at each iteration, respectively. To decrease the computation cost of initial point in projecting, in the following examples, we use the following iterative procedure which needs two projections onto C:

$$x^{k+1} = P_C(x^k - \alpha_k d_k),$$

$$\alpha_k = \rho(1-\sigma)\frac{\|r(x^k)\|^2}{\|d_k\|^2},$$

where ρ is a positive constant. We select $\sigma = 0.4$ and $\gamma = 0.8$.

Though our experience is limited in scope, it suggests that our method (NVE method for short) is a valuable alternative to the extra-gradient methods in [7–10,12,17]. We describe the detailed tests below.

EXAMPLE 1. This example was considered in [8,9], where $F(x) = Dx + c$, D is a nonsymmetric matrix of the form

$$\begin{pmatrix} 4 & -2 & & & & & \\ 1 & 4 & -2 & & & & \\ & 1 & 4 & -2 & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & & \cdot & \cdot & \cdot \\ & & & & \cdot & 4 & -2 \\ & & & & & 1 & 4 \end{pmatrix},$$

$c = (-1, -1, \dots, -1)^\top$ is a vector. The feasible set $C = [l, u]$, where $l = (0, 0, \dots, 0)^\top$ and $u = (1, 1, \dots, 1)^\top$. We choose $x^0 = (0, 0, \dots, 0)^\top$, take $\|r(x^k)\|^2 \le n10^{-14}$ as the termination criterion, where n is the dimension of the problem. The numerical results of Algorithm D in [9] and NVE algorithm are given in Table 1.

Table 1.

	Number of iterations (inner iterations)				
Algorithm	$n = 10$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
Alg. D	18 (9)	19 (10)	15 (4)	16 (6)	13 (2)
NVE ($\rho = 2$)	24	25	25	25	25
NVE ($\rho = 3$)	13	13	13	13	13
NVE ($\rho = 4$)	18	18	17	16	16

EXAMPLE 2. This example was considered in [8], where

$$\begin{aligned} F(x) &= F_1(x) + F_2(x), \\ F_1(x) &= (f_1(x), f_2(x), \dots, f_n(x))^\top, \\ F_2(x) &= Dx + c, \\ f_i(x) &= x_{i-1}^2 + x_i^2 + x_{i-1}x_i + x_i x_{i+1}, \qquad i = 1, 2, \dots, n, \\ x_0 &= x_{n+1} = 0, \end{aligned}$$

and D and c are the same as those in Example 1, respectively. The feasible set $C = R_+^n$. We choose $x^0 = (0, 0, \dots, 0)^\top$ and take $\|r(x^k)\|^2 \le n10^{-14}$ as the termination criterion, where n is the dimension of the problem. The numerical results of Algorithm PC in [8] and NVE algorithm are given in Table 2.

Table 2.

	Number of iterations (inner iterations)			
Algorithm	$n = 10$	$n = 20$	$n = 50$	$n = 100$
Alg. PC	14 (13)	14 (13)	13 (12)	13 (11)
NVE ($\rho = 2$)	21	20	20	20
NVE ($\rho = 3$)	12	12	12	11
NVE ($\rho = 4$)	19	19	19	18

EXAMPLE 3. The Kojima-Shindo Nonlinear Complementarity problem (NCP) (with $n = 4$) was considered in [20], where the function $F(x)$ is defined by

$$F(x_1, x_2, x_3, x_4) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}.$$

Let the feasible set be the simplex

$$C = \left\{ x \in R_+^n \mid x_1 + x_2 + x_3 + x_4 = 4 \right\}.$$

For all methods, $x^0 = (1, 1, 1, 1)^\top$, the termination criterion is $\|r(x)\| \le 10^{-4}$. The numerical results (iterative number) of extra-gradient algorithms in [7,9,12,17] and NVE algorithm are given in Table 3.

Table 3.

NVE ($300 \le \rho \le 570$)	Extra. in [12]	Extra. in [17]	Extra. in [7]	Extra. in [9]
5	7	38	16	78

The following example is a nonlinear complementarity problem whose defining function is taken from Nash equilibrium problem.

EXAMPLE 4. The defining function $F : R^n_+ \to R^n$ is of the form

$$F_i(q) = c_i(q_i) - p\left(\sum_{j=1}^{n} q_i\right) - q_i p\left(\sum_{j=1}^{n} q_i\right),$$

for $i = 1, 2, \ldots, n$, where

$$c_i(q_i) = \alpha_i q_i + \frac{\beta_i}{1+\beta_i} L_i^{-1/\beta_i} q_i^{1+1/\beta_i}, \qquad p(Q) = 5000^{1/\gamma} Q^{-1/\gamma},$$

with $Q = \sum_{j=1}^{n} q_i$. The data α_i, L_i, β_i, and γ are positive scalars which are taken from [21]

Table 4.

Firm i	c_i	L_i	β_i
1	10.0	5.0	1.2
2	8.0	5.0	1.1
3	6.0	5.0	1.0
4	4.0	5.0	.9
5	2.0	5.0	.8

We take $n = 5$, $\gamma = 1.1$. The termination criterion is $\|r(x)\|^2 \le 10^{-8}$. The numerical result (iterative number) of NVE algorithm is given in Table 5.

Table 5.

initial point	$\rho = 2$	$\rho = 2.5$	$\rho = 3$	$\rho = 3.5$	$\rho = 4$
$(10, 10, 10, 10, 10)^{\mathrm{T}}$	19	14	15	31	49
$(1, 1, 1, 1, 1)^{\mathrm{T}}$	20	14	10	15	45

The numerical results show that although the optimal step-size α_k is not used in Algorithm NVE, the NVE type projection method has good behavior if a suitable constant ρ is selected, which implies that d_k is a good direction. From Section 3, we know that the optimal step-size α_k is used in Algorithm NVE-1 or NVE-2, it should also have good behavior.

REFERENCES

1. R. Glowinski, *Numerical Methods for Nonlinear Variational Inequality Problems*, Springer Verlag, New York, (1984).
2. P.T. Harker and J.S. Pang, Finite-dimensional variation inequality and nonlinear complementarity problems: A survey of theory, algorithm and applications, *Math. Programming* **48**, 161–220, (1990).
3. M.C. Ferris and C. Kanzow, Complementarity and related problems, In *Handbook on Applied Optimization*, (Edited by P.M. Pardalos and M.G.C. Resende), Oxford University Press (to appear).
4. G.M. Korpelevich, The extragradient method for finding saddle points and other problems, *Matecon* **12**, 747–756, (1976).
5. A.N. Iusem, An iterative algorithm for the variational inequality problem, *Math. Appl. Compu.* **13**, 103-114, (1994).
6. E.N. Khobotov, Modification of the extragradient method for solving variational inequalities and certain optimization problem, *USSR Comput. Math. Phys.* **27**, 120–127, (1987).
7. P. Marcotte, Application of Khobotov's algorithm to variational inequalities and network equilibrium problems, *Inform. Systems Oper. Res.* **29**, 258–270, (1991).
8. D. Sun, A projection and contraction method for the nonlinear complementarity problems and its extensions, *Mathematica Numerica Sinica* **16**, 183–194, (1994).
9. D. Sun, A new step-size skill for solving a class of nonlinear projection equations, *J. Comput. Math.* **13**, 357–368, (1995).

10. D. Sun, A class of iterative method for solving nonlinear projection equations, *J. Optim. Theory Appl.* **91**, 123–140, (1996).
11. A.N. Iusem and B.F. Svaiter, A variant of Korpelevich's method for variational inequalities with a new search strategy, *Optimization* **42**, 309–321, (1997).
12. M.V. Solodov and B.F. Svaiter, A new projection method for variational inequality problems, *SIAM J. Control Optim.* **37**, 765–776, (1999).
13. Y.J. Wang, N.H. Xiu and C.Y. Wang, Unified framework fo extragradient-type methods for pseudomonotone variational inequalities, *Journal of Optimization Theory and Applications* **111** (3), (December 2001).
14. B.S. He, A new method for a class of linear variational inequalities, *Math. Programming* **66**, 137–144, (1994).
15. B.S. He, A class of projection and contraction methods for monotone variational inequalities, *Appl. Math. Optim.* **35**, 69–76, (1997).
16. M.A. Noor, A modified extragradient method for general monotone variational inequalities, *Computers Math. Applic.* **38** (1), 19–24, (1999).
17. M.V. Solodov and P. Tseng, Modified projection-type methods for monotone variational inequalities, *SIAM J. Control Optim.* **34**, 1814–1830, (1996).
18. P. Tseng, A modified forward-backward splitting method for maximal monotone mapping, *SIAM J. Control Optim.* **38**, 431–446, (2000).
19. N.H. Xiu and C.Y. Wang, On the step-size rule of extragradient method for monotone variational inequalities, *Mathematica Numerica Sinica* **22** (2), 197-208, (2000).
20. J.S. Pang and S.A. Gabriel, NE/SQP: A robust algorithm for the nonlinear complementarity problem, *Math. Programming* **60**, 295–337, (1993).
21. P.T. Harker, Accelerating the convergence of the diagonalization and projection algorithms for finite-dimensional variational inequalities, *Math. Programming* **41**, 29–59, (1988).