JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 112, No. 1, pp. 111–128, January 2002 (©2002)

Improvements of Some Projection Methods for Monotone Nonlinear Variational Inequalities[1]

B. S. HE[2] AND L. Z. LIAO[3]

Communicated by R. Glowinski

Abstract. In this paper, we study the relationship of some projection-type methods for monotone nonlinear variational inequalities and investigate some improvements. If we refer to the Goldstein–Levitin–Polyak projection method as the explicit method, then the proximal point method is the corresponding implicit method. Consequently, the Korpelevich extragradient method can be viewed as a prediction-correction method, which uses the explicit method in the prediction step and the implicit method in the correction step. Based on the analysis in this paper, we propose a modified prediction-correction method by using better prediction and correction stepsizes. Preliminary numerical experiments indicate that the improvements are significant.

Key Words. Monotone variational inequalities, explicit methods, implicit methods, prediction-correction methods.

1. Introduction

Let Ω be a nonempty closed convex subset of R^n, and let F be a continuous monotone mapping from R^n into itself. The variational inequality problem is to determine a vector $u^* \in \Omega$ such that

$$(\mathrm{VI}(\Omega, F)) \quad (u - u^*)^T F(u^*) \geq 0, \qquad \forall u \in \Omega. \tag{1}$$

Problem VI(Ω, F) includes nonlinear complementarity problems (when $\Omega = R^n_+$) and system of nonlinear equations (when $\Omega = R^n$); thus, it has many important applications (Refs. 1–2). For solving variational inequality

[1]This research was supported in part by NSFC Grant 19971040, by Jiangsu NSF Grant BK99041, and by Hong Kong Baptist University Grant FRG/99-00/II-23.
[2]Professor, Department of Mathematics, Nanjing University, Nanjing, China.
[3]Associate Professor, Department of Mathematics, Hong Kong Baptist University, Kowloon, Hong Kong, China.

problems, besides Newton-like methods (e.g., Refs. 3–6), there are projection-type methods for linear variational inequalities (Refs. 7–10) and nonlinear variational inequalities (e.g., Refs. 11–17).

Our interest in this paper is to study the relationships between some projection-type methods and then investigate some improvements. Even though most projection methods are gradient-type methods and are sometimes inefficient, they are considerably simpler and well suited for large sparse problems. Motivated by the Goldstein–Levitin–Polyak method (Refs. 11–12) and the proximal method (Ref. 18), which can be viewed as explicit and implicit methods respectively, the extragradient method proposed by Korpelevich (Refs. 14–15) can be referred to as a prediction-correction method. From this point of view, we make some improvements to such projection methods. For convenience, we consider only the projection under the Euclidean norm and do not consider the projection under the general G-norm. However, the general case can be added easily once the basic ideas are clear.

The paper is organized as follows. In Section 2, we summarize some basic concepts and consequent results. In Section 3–4, we study the explicit and implicit methods, respectively. In Section 5, we analyze the convergence property of the prediction-correction method. A modified prediction-correction method is given in Section 6. In Section 7, we present some numerical results to indicate that the improvements in the modified prediction-correction method are significant. Some concluding remarks are given in Section 8.

Throughout this paper, we assume that the operator F is monotone and Lipschitz continuous on Ω, and that the solution set of VI(Ω, F), denoted by Ω^*, is nonempty.

2. Preliminaries

Now, let us summarize first some basic concepts and their properties that will be used in the subsequent sections. We use the concept of projection under the Euclidean norm; this will be denoted by $P_\Omega(\cdot)$, i.e.,

$$P_\Omega(v) = \operatorname{argmin}\{\|v-u\| \,|\, u\in\Omega\}.$$

From the above definition, it follows that

$$\{v - P_\Omega[v]\}^T\{w - P_\Omega[v]\} \le 0, \qquad \forall v\in R^n, \forall w\in\Omega, \tag{2}$$

$$(v-w)^T\{P_\Omega(v) - P_\Omega(w)\} \ge \|P_\Omega(v) - P_\Omega(w)\|^2, \qquad \forall v, w\in R^n. \tag{3}$$

Consequently, we have

$$\|P_\Omega(v) - P_\Omega(w)\| \leq \|v - w\|, \qquad \forall v, w \in R^n, \tag{4}$$

$$\|P_\Omega(v) - u\|^2 \leq \|v - u\|^2 - \|v - P_\Omega(v)\|^2, \qquad \forall u \in \Omega. \tag{5}$$

Lemma 2.1. See Ref. 19, p. 267. Let $\beta > 0$. Then, u^* solves VI(Ω, F) if and only if

$$u^* = P_\Omega[u^* - \beta F(u^*)].$$

Denote

$$e(u, \beta) := u - P_\Omega[u - \beta F(u)]. \tag{6}$$

Then, solving VI(Ω, F) is equivalent to finding a zero point of $e(u, \beta)$. The next lemma states that $\|e(u, \beta)\|$ is a nondecreasing function for $\beta > 0$.

Lemma 2.2. For all $u \in R^n$ and $\tilde{\beta} \geq \beta > 0$, it holds that

$$\|e(u, \tilde{\beta})\| \geq \|e(u, \beta)\|. \tag{7}$$

Proof. We need to prove only that

$$e(u, \beta)^T(e(u, \tilde{\beta}) - e(u, \beta)) \geq 0, \qquad \forall \tilde{\beta} \geq \beta > 0. \tag{8}$$

Substituting

$$w := P_\Omega[u - \tilde{\beta} F(u)] \quad \text{and} \quad v := u - \beta F(u)$$

into (2) and using

$$P_\Omega[u - \beta F(u)] - P_\Omega[u - \tilde{\beta} F(u)] = e(u, \tilde{\beta}) - e(u, \beta),$$

we get

$$\{[u - \beta F(u)] - P_\Omega[u - \beta F(u)]\}^T\{e(u, \tilde{\beta}) - e(u, \beta)\} \geq 0. \tag{9}$$

It follows from (9) that

$$e(u, \beta)^T\{e(u, \tilde{\beta}) - e(u, \beta)\} \geq \beta F(u)^T\{e(u, \tilde{\beta}) - e(u, \beta)\}. \tag{10}$$

Setting

$$v := u - \beta F(u) \quad \text{and} \quad w := u - \tilde{\beta} F(u)$$

in (3), we get

$$(\tilde{\beta} - \beta)F(u)^T\{e(u, \tilde{\beta}) - e(u, \beta)\} \geq \|e(u, \tilde{\beta}) - e(u, \beta)\|^2. \tag{11}$$

Using (10)–(11), it follows that inequality (8) is true and the lemma is proved. □

Now, let us present a convergence theorem which is useful for various methods studied in this paper.

Theorem 2.1. Let $c_0 > 0$ be a constant, let $l \in \{0, 1\}$ be a given integer, let $\{\beta_k\}$ be a positive sequence, and let $\inf\{\beta_k\} = \beta_{\min} > 0$. If the sequence $\{u^k\}$ generated by a method satisfies

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - c_0 \|e(u^{k+l}, \beta_k)\|^2, \qquad \forall u^* \in \Omega^*, \tag{12}$$

then $\{u^k\}$ converges to a solution point of $\mathrm{VI}(\Omega, F)$.

Proof. Let $\hat{u}$ be a solution of $\mathrm{VI}(\Omega, F)$. First, from (12), we get

$$\sum_{k=0}^{\infty} c_0 \|e(u^{k+l}, \beta_k)\|^2 \leq \|u^0 - \hat{u}\|^2,$$

and it follows from Lemma 2.2 that

$$\lim_{k \to \infty} e(u^k, \beta_{\min}) = 0.$$

Again, it follows from (12) that the sequence $\{u^k\}$ is bounded. Let $\tilde{u}^*$ be a cluster point of $\{u^k\}$, and let the subsequence $\{u^{k_j}\}$ converge to $\tilde{u}^*$. Because $e(u, \beta_{\min})$ is continuous,

$$e(\tilde{u}^*, \beta_{\min}) = \lim_{j \to \infty} e(u^{k_j}, \beta_{\min}) = 0,$$

and thus $\tilde{u}^*$ is a solution of $\mathrm{VI}(\Omega, F)$. In the following, we prove that the sequence $\{u^k\}$ has exactly one cluster point. Assume that $\tilde{u}$ is another cluster point, and denote

$$\delta := \|\tilde{u} - \tilde{u}^*\| > 0.$$

Because $\tilde{u}^*$ is a cluster point of the sequence $\{u^k\}$, there is a $k_0 > 0$ such that

$$\|u^{k_0} - \tilde{u}^*\| \leq \delta/2.$$

On the other hand, since $\tilde{u}^* \in \Omega^*$ and thus

$$\|u^k - \tilde{u}^*\| \leq \|u^{k_0} - \tilde{u}^*\|, \qquad \forall k \geq k_0,$$

it follows that

$$\|u^k - \tilde{u}\| \geq \|\tilde{u} - \tilde{u}^*\| - \|u^k - \tilde{u}^*\| \geq \delta/2, \qquad \forall k \geq k_0.$$

This contradicts the assumption; thus, the sequence $\{u^k\}$ converges to $\tilde{u}^* \in \Omega^*$. □

Let

$$\mathrm{dist}(u, \Omega^*) = \inf\{\|u - u^*\| \,|\, u^* \in \Omega^*\}.$$

Note that (12) implies

$$\operatorname{dist}(u^{k+1}, \Omega^*) \le \operatorname{dist}(u^{(k)}, \Omega^*) - c_0 \|e(u^{k+l}, \beta_k)\|^2.$$

$\|e(u, \beta_{\min})\|$ is the error bound of $\mathrm{VI}(\Omega, F)$. If there exists a $\mu > 0$ such that

$$\operatorname{dist}(u, \Omega^*) \le \mu \|e(u, \beta_{\min})\|, \qquad \forall u, \tag{13a}$$

with

$$\operatorname{dist}(x, \Omega^*) \le \operatorname{dist}(u^0, \Omega^*), \tag{13b}$$

we can obtain the linear convergence from Theorem 2.1 immediately. In the rest of this paper, for convergence analysis, we need only to pay our attention on the conditions of Theorem 2.1 for the generated sequence.

3. Goldstein–Levitin–Polyak Projection Method (Explicit Method)

Among the existing methods for nonlinear variational inequality problems, the simplest is the Goldstein–Levitin–Polyak projection method (Refs. 11–12), which starts with any $u^0 \in \Omega$ and updates iteratively u^{k+1} according to the formula

$$\text{(EM)} \quad u^{k+1} = P_\Omega[u^k - \beta_k F(u^k)], \tag{14}$$

where β_k is a chosen positive stepsize. This projection method can be viewed as an explicit method because u^{k+1} occurs only on the left-hand side of Eq. (14). Under the assumptions that F is Lipschitz continuous with a constant $L > 0$,

$$\|F(u) - F(v)\| \le L\|u - v\|, \tag{15}$$

that F is uniformly strong monotone with a constant modulus $\tau > 0$,

$$(u - v)^T [F(u) - F(v)] \ge \tau \|u - v\|^2, \tag{16}$$

and that the stepsize β_k satisfies

$$0 < \beta_L \le \beta_k \le \beta_U < 2\tau / L^2, \tag{17}$$

the explicit method (14) is convergent. Namely, we have

$$1 - 2\tau\beta_k + \beta_k^2 L^2 < 1,$$

and using (15)–(16),

$$\begin{aligned}\|u^{k+1}-u^*\|^2 &= \|P_\Omega[u^k-\beta_k F(u^k)]-P_\Omega[u^*-\beta_k F(u^*)]\|^2\\ &\le \|(u^k-u^*)-\beta_k[F(u^k)-F(u^*)]\|^2\\ &\le \|u^k-u^*\|^2-2\beta_k(u^k-u^*)^T[F(u^k)-F(u^*)]\\ &\quad+\beta_k^2\|F(u^k)-F(u^*)\|^2\\ &\le (1-2\tau\beta_k+\beta_k^2L^2)\|u^k-u^*\|^2. \end{aligned} \tag{18}$$

Moreover, if $\{\beta_k\}$ is monotonically nonincreasing, then it follows from Lemma 2.2 that

$$\begin{aligned}\|e(u^{k+1},\beta_{k+1})\|^2 &\le \|e(u^{k+1},\beta_k)\|^2\\ &= \|u^{k+1}-P_\Omega[u^{k+1}-\beta_k F(u^{k+1})]\|^2\\ &= \|P_\Omega[u^k-\beta_k F(u^k)]-P_\Omega[u^{k+1}-\beta_k F(u^{k+1})]\|^2\\ &\le \|(u^k-u^{k+1})-\beta_k[F(u^k)-F(u^{k+1})]\|^2\\ &\le \|u^k-u^{k+1}\|^2-2\beta_k(u^k-u^{k+1})^T[F(u^k)-F(u^{k+1})]\\ &\quad+\beta_k^2\|F(u^k)-F(u^{k+1})\|^2\\ &\le (1-2\tau\beta_k+\beta_k^2L^2)\|u^k-u^{k+1}\|^2\\ &= (1-2\tau\beta_k+\beta_k^2L^2)\|e(u^k,\beta_k)\|^2. \end{aligned} \tag{19}$$

In other words, for fixed β satisfying (17), both sequences $\{\|u^k-u^*\|\}$ and $\{\|e(u^k,\beta)\|\}$ are linearly convergent to zero globally. However, the efficiency of this method depends on the estimations of the Lipschitz constant L and the uniform strong monotone modulus τ. It is very expensive to estimate the modulus τ and the Lipschitz constant L, even if F is an affine mapping. Hence, in practice, the explicit method (14) is used only for well-conditioned problems. Moreover, in general, a simple explicit projection method can be used only for solving uniformly strong monotone variational inequalities. This can be seen from the following example.

Example 3.1. Let

$$\Omega=R^2, \qquad F(u)=Mu, \qquad M=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}. \tag{20}$$

This linear variational inequality has a unique solution $u^*=0$. When the problem is solved by the explicit method (14), the recursion is

$$u^{k+1}=u^k-\beta_k Mu^k. \tag{21}$$

Since M is skew-symmetric and $\|Mu\| = \|u\|$, from (21) we have

$$\|u^{k+1}\| = \sqrt{1+\beta_k^2}\|u^k\|;$$

thus, the sequence $\{u^k\}$ generated by (21) is divergent when $u^0 \neq 0$.

4. Proximal-Point Algorithm (Implicit Method)

Another popular formulation for VI(Ω, F) is the multivalued equation

$$0 \in T(u) =: F(u) + N_\Omega(u), \tag{22}$$

where $N_\Omega(\cdot)$ is the normal cone operator to Ω, i.e.,

$$N_\Omega(u) := \begin{cases} \{w \mid (v-u)^T w \leq 0, \ \forall v \in \Omega\}, & \text{if } u \in \Omega, \\ \varnothing, & \text{otherwise.} \end{cases} \tag{23}$$

Note that $N_\Omega(u)$ is a cone and hence

$$cN_\Omega(u) = N_\Omega(u), \qquad \text{for all } u \in R^n \text{ and } c > 0.$$

A classical method to solve this problem is the proximal point algorithm (18), which starts with any vector $u^0 \in \Omega$ and updates iteratively u^{k+1} according to the following recursion:

$$0 \in (u^{k+1} - u^k) + \alpha_k T(u^{k+1}),$$

where $\{\alpha_k\} \subset [\alpha_{\min}, \infty)$, $\alpha_{\min} > 0$, is a sequence of scalars. This is equivalent to

$$-[(u^{k+1} - u^k) + \alpha_k F(u)] \in N_\Omega(u^{k+1}).$$

In other words, for given $u^k \in \Omega$, the new iterate u^{k+1} is obtained by finding

$$u \in \Omega, \qquad (u' - u)^T[(u - u^k) + \alpha_k F(u)] \geq 0, \qquad \forall u' \in \Omega. \tag{24}$$

According to Lemma 1.1, u^{k+1} is the solution of

$$u = P_\Omega\{u - [(u - u^k) + \alpha_k F(u)]\};$$

thus,

$$\text{(IM)} \quad u^{k+1} = P_\Omega[u^k - \alpha_k F(u^{k+1})], \tag{25}$$

Because u^{k+1} occurs on both sides of Eq. (25), we call the method implicit.

Recall that u^{k+1} is the solution of variational inequality problem (24). Compared with the original variational inequality (1), problem (24) is well conditioned. By denoting

$$\bar{u}^{k+1} := u^k - \alpha_k F(u^{k+1}),$$

we have

$$u^{k+1} = P_\Omega(\bar{u}^{k+1}).$$

Using (5), we obtain

$$\begin{aligned}\|u^{k+1} - u^*\|^2 &\le \|\bar{u}^{k+1} - u^*\|^2 - \|\bar{u}^{k+1} - u^{k+1}\|^2 \\ &= \|(u^k - u^*) - \alpha_k F(u^{k+1})\|^2 \\ &\quad - \|(u^k - u^{k+1}) - \alpha_k F(u^{k+1})\|^2 \\ &= \|u^k - u^*\|^2 - 2\alpha_k (u^{k+1} - u^*)^T F(u^{k+1}) \\ &\quad - \|u^k - u^{k+1}\|^2. \end{aligned} \tag{26}$$

But from the monotonicity of F and (1), we have

$$(u^{k+1} - u^*)^T F(u^{k+1}) \ge (u^{k+1} - u^*)^T F(u^*) \ge 0.$$

This and (26) imply

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - \|u^k - u^{k+1}\|^2. \tag{27}$$

Using (25) and (4), we have

$$\begin{aligned}\|e(u^{k+1}, \alpha_k)\| &= \|u^{k+1} - P_\Omega[u^{k+1} - \alpha_k F(u^{k+1})]\| \\ &= \|P_\Omega[u^k - \alpha_k F(u^{k+1})] - P_\Omega[u^{k+1} - \alpha_k F(u^{k+1})]\| \\ &\le \|u^k - u^{k+1}\|, \end{aligned} \tag{28}$$

and (27) becomes

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - \|e(u^{k+1}, \alpha_k)\|^2, \qquad \forall u^* \in \Omega^*. \tag{29}$$

By setting $\beta = \alpha$, $l = 1$, and $c_0 = 1$ it follows from Theorem 2.1 that the sequence $\{u^k\}$ produced by this implicit method converges to a solution point u^*. If the illustrative problem (20) is solved by the implicit method (25), the recursion becomes

$$u^{k+1} = u^k - \alpha_k M u^{k+1}, \tag{30}$$

and we have

$$\sqrt{1 + \alpha_k^2}\, \|u^{k+1}\| = \|u^k\|;$$

thus, the sequence $\{u^k\}$ generated by (30) converges to the solution point $u^* = 0$ from any starting point u^0.

This implicit method is convergent whenever F is monotone and Ω^* is nonempty; however, it has to solve a well-conditioned variational inequality of the form (24) at least approximately in each iteration.

5. Extragradient Method

Among the numerical solution methods for ordinary differential equations, besides the explicit methods (e.g., explicit Euler method) and implicit methods (e.g., Adams–Moulton method), there are also prediction-correction methods, which use explicit methods in the prediction step and implicit methods in the correction step. For $\mathrm{VI}(\Omega, F)$, the extragradient method introduced by Korpelevich (Ref. 14) can be viewed as a prediction-correction method. It uses the Goldstein–Levitin–Polyak explicit method (14) to make a prediction,

$$\text{(P)} \quad \bar{u}^{k+1} = P_\Omega[u^k - \beta_k F(u^k)], \tag{31}$$

and then uses the implicit scheme (24) to make a correction,

$$\text{(C)} \quad u^{k+1} = P_\Omega[u^k - \beta_k F(\bar{u}^{k+1})]. \tag{32}$$

Substituting (31) into (32), the recursion of the Korpelevich method can be rewritten as

$$\text{(PC)} \quad u^{k+1} = P_\Omega\{u^k - \beta_k F(P_\Omega[u^k - \beta_k F(u^k)])\}. \tag{33}$$

Under the assumption

$$\begin{aligned} &\|\beta_k F(u^k) - \beta_k F(P_\Omega[u^k - \beta_k F(u^k)])\| \\ &\le \nu \|u^k - P_\Omega[u^k - \beta_k F(u^k)]\|, \qquad \nu \in (0, 1), \end{aligned} \tag{34}$$

it was proved in Ref. 14 that the sequence $\{u^k\}$ generated by the Korpelevich method satisfies

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - (1 - \nu^2)\|e(u^k, \beta_k)\|^2. \tag{35}$$

Later, the extragradient method was extended by Khobotov (Ref. 15). The framework of the Korpelevich–Khobotov method is outlined below.

Korpelevich–Khobotov Method (KK Method)

Step 0. Set $\beta_0 = 1$, $\nu \in (0, 1)$, $u^0 \in \Omega$, and $k = 0$.

Step 1. Compute $\bar{u}^{k+1} = P_\Omega[u^k - \beta_k F(u^k)]$.

Step 2. If $r_k := \beta_k \|F(u^k) - F(\bar{u}^{k+1})\| / \|u^k - \bar{u}^{k+1}\| \le \nu$,
then set $u^{k+1} = P_\Omega[u^k - \beta_k F(\bar{u}^{k+1})]$,
$\beta_{k+1} = \beta_k$, and $k = k+1$; go to Step 1.

Step 3. Reduce the value of β_k via $\beta_k := (2/3)\beta_k \min\{1, 1/r_k\}$;
set $\bar{u}^{k+1} = P_\Omega[u^k - \beta_k F(u^k)]$; go to Step 2.

The reducing rule for β_k in Step 3 is of the Armijo type. If the mapping F is Lipschitz continuous, then we can find a $\beta_k > 0$ such that $r_k \le \nu$ in finite reduction substeps. In addition, under the Lipschitz continuity, it is easy to prove that

$$\inf\{\beta_k\} = \beta_{\min} > 0.$$

Thus, it follows from (35) and Theorem 2.1 that the sequence generated by the KK method converges to a solution point of VI(Ω, F).

6. Improvements In Prediction-Correction Methods

Now, we consider how to modify the KK method and improve the efficiency of the prediction-correction methods. Our modifications include the following two aspects:

Improvement 1: Selection of $\{\alpha_k\}$, correction stepsize.
Improvement 2: Selection of $\{\beta_k\}$, prediction stepsize.

The detailed analysis for the two improvements is given below. First, for a given $u^k \in \Omega$, we use the prediction step,

$$\text{(P)} \quad \bar{u}^{k+1} = P_\Omega[u^k - \beta_k F(u^k)], \tag{36}$$

in which β_k satisfies

$$\|\beta_k[F(u^k) - F(\bar{u}^{k+1})]\| \le \nu \|u^k - \bar{u}^{k+1}\|, \qquad \nu \in (0, 1), \tag{37}$$

as in the KK method. Then, a correction step,

$$\text{(C)} \quad u^{k+1} = P_\Omega[u^k - \alpha_k F(\bar{u}^{k+1})], \tag{38}$$

is taken. It seems that the stepsize α_k should depend on the current point u^k and the stepsize β_k in the prediction step. This motivates the following analysis which aims at finding a best α for given u^k and β_k.

For the convenience of the coming analysis, we replace u^k and β_k by u and β, respectively. In addition, we use the notation

$$K(u, \beta) := F(P_\Omega[u - \beta F(u)]), \tag{39}$$

and then the new iterate of the prediction-correction method can be written as

$$u(\alpha) := P_\Omega[u - \alpha K(u, \beta)]. \tag{40}$$

Another useful notation in the coming analysis is

$$d(u, \beta) := e(u, \beta) - \beta[F(u) - K(u, \beta)]. \tag{41}$$

Under Assumption (37), we have

$$\begin{aligned} e(u, \beta)^T d(u, \beta) &= \|e(u, \beta)\|^2 - \beta\, e(u, \beta)^T[F(u) - K(u, \beta)] \\ &\geq (1 - \nu)\|e(u, \beta)\|^2, \end{aligned} \tag{42}$$

$$\begin{aligned} e(u, \beta)^T d(u, \beta) &= \|e(u, \beta)\|^2 - \beta\, e(u, \beta)^T[F(u) - K(u, \beta)] \\ &> (1/2)\|e(u, \beta)\|^2 - \beta\, e(u, \beta)^T[F(u) - K(u, \beta)] \\ &\quad + (1/2)\beta^2\|F(u) - K(u, \beta)\|^2 \\ &= (1/2)\|d(u, \beta)\|^2. \end{aligned} \tag{43}$$

Now, let us observe the difference between $\|u - u^*\|^2$ and $\|u(\alpha) - u^*\|^2$. Denote

$$\Theta(\alpha) := \|u - u^*\|^2 - \|u(\alpha) - u^*\|^2. \tag{44}$$

First, it follows from (5) and (40) that

$$\|u(\alpha) - u^*\|^2 \leq \|u - \alpha K(u, \beta) - u^*\|^2 - \|u - \alpha K(u, \beta) - u(\alpha)\|^2.$$

Substituting this into (44), we obtain

$$\begin{aligned} \Theta(\alpha) \geq\ & 2\alpha(u - u^*)^T, K(u, \beta) \\ & + \|u - u(\alpha)\|^2 - 2\alpha[u - u(\alpha)]^T K(u, \beta). \end{aligned} \tag{45}$$

Now, we observe the first term on the right-hand side of the above inequality. Since $P_\Omega[\cdot] \in \Omega$, it follows from the definition of $\mathrm{VI}(\Omega, F)$ that

$$\{P_\Omega[u - \beta F(u)] - u^*\}^T F(u^*) \geq 0. \tag{46}$$

Under the assumption that F is monotone with respect to Ω, we obtain

$$\{P_\Omega[u - \beta F(u)] - u^*\}^T \{F(P_\Omega[u - \beta F(u)]) - F(u^*)\} \geq 0. \tag{47}$$

Adding (46)–(47), and using the notation $e(u, \beta)$ and $K(u, \beta)$, we get

$$\{(u - u^*) - e(u, \beta)\}^T K(u, \beta) \geq 0, \qquad \forall u \in R^n,\ u^* \in \Omega^*.$$

Hence, we have

$$(u - u^*)^T K(u, \beta) \geq e(u, \beta)^T K(u, \beta). \tag{48}$$

Substituting (48) into (45), we get

$$\Theta(\alpha) \geq 2\alpha\, e(u,\beta)^T K(u,\beta) + \|u-u(\alpha)\|^2 - 2\alpha[u-u(\alpha)]^T K(u,\beta). \tag{49}$$

From (49) and using the notation $d(u,\beta)$ [see (41)], we obtain

$$\begin{aligned}\Theta(\alpha) \geq\ & 2\alpha\, e(u,\beta)^T K(u,\beta) - (\alpha^2/\beta^2)\|d(u,\beta)\|^2 \\ & + \|[u-u(\alpha)] - (\alpha/\beta)d(u,\beta)\|^2 \\ & + (2\alpha/\beta)\,[u-u(\alpha)]^T\{e(u,\beta) - \beta F(u)\}.\end{aligned} \tag{50}$$

Now, we consider the last term on the right-hand side of (50). Notice that

$$u - u(\alpha) = e(u,\beta) + \{P_\Omega[u-\beta F(u)] - u(\alpha)\}. \tag{51}$$

By using

$$v := u - \beta F(u) \quad \text{and} \quad w := u(\alpha)$$

in the basic inequality of the projection mapping (2), we get

$$\{P_\Omega[u-\beta F(u)] - u(\alpha)\}^T\{e(u,\beta) - \beta F(u)\} \geq 0. \tag{52}$$

Inequalities (51)–(52) imply

$$[u-u(\alpha)]^T\{e(u,\beta) - \beta F(u)\} \geq e(u,\beta)^T\{e(u,\beta) - \beta F(u)\};$$

thus, from (50) we have

$$\Theta(\alpha) \geq (2\alpha/\beta)\, e(u,\beta)^T d(u,\beta) - (\alpha^2/\beta^2)\|d(u,\beta)\|^2. \tag{53}$$

The right-hand side of (53) is a quadratic function of α and it reaches its maximum at

$$\alpha^* = \beta\tau(u,\beta), \tag{54}$$

where

$$\tau(u,\beta) := e(u,\beta)^T d(u,\beta)/\|d(u,\beta)\|^2. \tag{55}$$

Thus, from (53) and (54), we have

$$\begin{aligned}\Theta(\alpha^*) &\geq (\alpha^*/\beta)\, e(u,\beta)^T d(u,\beta) \\ &= \tau(u,\beta)\cdot e(u,\beta)^T d(u,\beta).\end{aligned} \tag{56}$$

Let $\gamma \in (0,2)$ be a relaxation factor, and let $\alpha = \gamma\alpha^*$. Using (53), by a simple manipulation we get

$$\Theta(\gamma\alpha^*) \geq \gamma(2-\gamma)\tau(u,\beta)\cdot e(u,\beta)^T d(u,\beta). \tag{57}$$

Note that it follows from (42)–(43) that

$$\tau(u,\beta)\cdot e(u,\beta)^T d(u,\beta) > (1/2)(1-\nu)\|e(u,\beta)\|^2.$$

This and (57) imply

$$\Theta(\gamma\alpha^*) > [\gamma(2-\gamma)(1-\nu)/2]\|e(u,\beta)\|^2, \qquad \gamma\in(0,2). \tag{58}$$

We summarize the analytical results of this section in the following theorem.

Theorem 6.1 Let

$$d(u,\beta) := e(u,\beta) - \{\beta F(u) - \beta F(P_\Omega[u-\beta F(u)])\},$$

$$\tau(u,\beta) = e(u,\beta)^T d(u,\beta)/\|d(u,\beta)\|^2.$$

For given $u^k\in\Omega$, β_k is chosen such that

$$\|\beta_k F(u^k) - \beta_k F(P_\Omega[u^k-\beta_k F(u^k)])\| \le \nu\|e(u^k,\beta_k)\|, \qquad \nu\in(0,1).$$

Then, the prediction-correction method

$$\text{(PC)} \quad u^{k+1} = P_\Omega\{u^k - \alpha_k F(P_\Omega[u^k-\beta_k F(u^k)])\},$$

with correction stepsize

$$\alpha_k = \gamma\beta_k\tau(u^k,\beta_k), \qquad \gamma\in(0,2),$$

produces a new iterate which satisfies

$$\|u^{k+1}-u^*\|^2 \le \|u^k-u^*\|^2 - [\gamma(2-\gamma)(1-\nu)/2]\|e(u^k,\beta_k)\|^2, \qquad \forall u^*\in\Omega^*. \tag{59}$$

Note that the convergence assertion (35) of the Korpelevich method can be derived directly from (57). Since $\tau(u,\beta) > 1/2$ [see (43) and (55)], we have

$$\gamma^* := 1/\tau(u,\beta) = \|d(u,\beta)\|^2/e(u,\beta)^T d(u,\beta) \in (0,2). \tag{60}$$

Let

$$\alpha := \gamma^*\alpha^* = \beta;$$

we get the Korpelevich method. It follows from (57), (41), (37) that

$$\begin{aligned}\Theta(\beta) &\ge (2-\gamma^*)e(u,\beta)^T d(u,\beta)\\ &= [2-\|d(u,\beta)\|^2/e(u,\beta)^T d(u,\beta)]\, e(u,\beta)^T d(u,\beta)\\ &= 2e(u,\beta)^T d(u,\beta) - \|d(u,\beta)\|^2\\ &= \|e(u,\beta)\|^2 - \|\beta[F(u)-K(u,\beta)]\|^2\\ &\ge (1-\nu^2)\|e(u,\beta)\|^2.\end{aligned} \tag{61}$$

This is the main convergence result (35), which was proved by Korpelevich in Ref. 14.

Based on the above analysis, for given u and β, the ideal stepsize in the correction step α_k^* is given by (54). To ensure a faster convergence, we take a relaxation factor $\gamma \in (0, 2)$, but close to 2. Hence, instead of $\alpha_k = \beta_k$, our first improvement is that the stepsize in the correction step is

$$\alpha_k = \gamma_k \beta_k \tau_k, \qquad \gamma_k \equiv 1.8.$$

Now, we discuss Improvement 2. The parameter sequence $\{\beta_k\}$ in the KK method is monotonically nonincreasing. However, this may cause a slow convergence if β_k is taken too small. To avoid this situation, sometimes increasing the stepsize in Step 2 is necessary.

By considering the above two improvements, we arrive at the following improved prediction-correction method.

Improved Prediction-Correction Method (PC Method M_{1+2})

Step 0. Let $\beta_0 = 1, 0 < \mu < \nu < 1, u^0 \in \Omega, \gamma = 1.8$, and $k = 0$.

Step 1. Compute $\bar{u}^{k+1} = P_\Omega[u^k - \beta_k F(u^k)]$.

Step 2. If $r_k := \beta_k \|F(u^k) - F(\bar{u}^{k+1})\| / \|u^k - \bar{u}^{k+1}\| \leq \nu$, then set

$$e(u^k, \beta_k) = u^k - \bar{u}^{k+1},$$

$$d(u^k, \beta_k) = e(u^k, \beta_k) - \beta_k[F(u^k) - F(\bar{u}^{k+1})],$$

$$\alpha_k = \gamma \beta_k [e(u^k, \beta_k)^T d(u^k, \beta_k) / \|d(u^k, \beta_k)\|^2], \quad \text{(Improvement 1)}$$

$$u^{k+1} = P_\Omega[u^k - \alpha_k F(\bar{u}^{k+1})],$$

$$\beta_k := \begin{cases} (3/2)\beta_k, & \text{if } r_k \leq \mu, \\ \beta_k, & \text{otherwise,} \end{cases} \qquad \text{(Improvement 2)}$$

$$\beta_{k+1} = \beta_k, \quad \text{and} \quad k = k + 1; \qquad \text{go to Step 1.}$$

Step 3. Reduce the value of β_k via $\beta_k := (2/3)\beta_k * \min\{1, 1/r_k\}$; set $\bar{u}^{k+1} = P_\Omega[u^k - \beta_k F(u^k)]$; go to Step 2.

Remark 6.1. In comparison with the Korpelevich–Khobotov method, we need some extra computation in Improvements 1 and 2 of the improved method. Instead of $\alpha_k \equiv \beta_k$, we determine α_k via Improvement 1 and this requires only $O(n)$ flops. The computational cost in Improvement 2 of each iteration is tiny and thus can be ignored.

7. Numerical Experiments

Our main interest is in showing the advantages of the improved prediction-correction method. In our numerical test, we considered the following nonlinear complementarity problem:

(NCP) Find $u \in R^n$ such that

$$u \geq 0, \qquad F(u) \geq 0, \qquad u^T F(u) = 0. \tag{62}$$

This is a special case of $\mathrm{VI}(\Omega, F)$ with $\Omega = R^n_+$ and the projection on Ω in the sense of the Euclidean norm is very easy to carry out. For any $v \in R^n$, $P_\Omega[v]$ is defined componentwise as

$$(P_\Omega[v])_j = \begin{cases} v_j, & \text{if } v_j \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

In our test problem, we take

$$F(u) = D(u) + Mu + q,$$

where $D(u)$ and $Mu + q$ are the nonlinear part and the linear part of $F(u)$, respectively. We form the linear part $Mu + q$ similarly as in Ref. 20.[4] The matrix $M = A^T A + B$, where A is an $n \times n$ matrix whose entries are randomly generated in the interval $(-5, +5)$ and a skew-symmetric matrix B is generated in the same way. The vector q is generated from a uniform distribution in the interval $(-500, +500)$ or $(-500, 0)$. In $D(u)$, the nonlinear part of $F(u)$, the components are $D_j(u) = a_j * \arctan(u_j)$ and a_j is a random variable in $(0, 1)$. A similar type of the problem was tested in Refs. 21–22.[5]

In order to see the effects of Improvements 1 and 2, respectively, we test the same problem with the following different methods:

KK Method: original Korpelevich–Khobotov method.
PC Method M_1: KK method with Improvement 1.
PC Method M_2: KK method with Improvement 2.
PC Method M_{1+2}: KK method with Improvements 1 and 2, the improved prediction-correction method described in Section 6.

In all tests, we choose $\nu = 0.9$ and the parameter $\mu = 0.4$ in Improvement 2. All codes were written in Matlab and run on a PIII-600 Acer notebook computer. We tested problems with dimensions $n = 100, 200, 500$. For

[4] In the paper by Harker and Pang (Ref. 20), the matrix $M = A^T A + B + D$, where A and B are the same matrices as here, and D is a diagonal matrix with uniformly distributed random variable $d_{jj} \in (0.0, 0.3)$.

[5] In Refs. 21–22, the components of nonlinear mapping $D(u)$ are $D_j(u) = \mathrm{const} * \arctan(u_j)$.

our test problem, all methods started at the same $u^0 \in (0, 10)$ and stopped as soon as $\|e(u^k, 1)\|_\infty \le 10^{-7}$. To obtain more stable results, we did run each test case 5 times. The average numbers of iterations and the computation times of these methods for problem with different sizes are given in the following tables.

Table 1. Numerical results for NCP easy problems, $q \in (-500, 500)$.

	KK Method		PC Method M_1		PC Method M_2		PC Method M_{1+2}	
n	Nit	CPU (s)	Nit	CPU (s)	Nit	CPU (s)	Nit	CPU (s)
100	893	1.14	476	0.67	616	0.81	342	0.48
200	1116	5.56	579	3.09	726	3.86	408	2.26
500	1154	30.41	625	16.46	731	19.70	413	11.27

Table 2. Numerical results for NCP hard problems, $q \in (-500, 0)$.

	KK Method		PC Method M_1		PC Method M_2		PC Method M_{1+2}	
n	Nit	CPU (s)	Nit	CPU (s)	Nit	CPU (s)	Nit	CPU (s)
100	1762	2.24	1026	1.43	1362	1.82	776	1.10
200	1980	9.67	1028	5.45	1400	7.41	786	4.31
500	2354	60.84	1328	32.21	1728	47.77	1003	27.68

From the numerical results, we find that both Improvements 1 and 2 are effective. The computational costs in each iteration of the methods are almost equal. For such quasi-randomly constructed test problem, we can observe from the above two tables that

$$\frac{\text{total computational load of PC Method } M_1}{\text{total computational load of KK Method}} < 0.60,$$

$$\frac{\text{total computational load of PC Method } M_2}{\text{total computational load of KK Method}} < 0.80,$$

$$\frac{\text{total computational load of PC Method } M_{1+2}}{\text{total computational load of KK Method}} < 0.45.$$

With a tiny extra computation, the efficiency of the prediction-correction methods is improved significantly. In addition, for a set of similar problems, it seems that the iteration numbers are not very sensitive to the problem size.

8. Conclusions

For solving $\mathrm{VI}(\Omega, F)$ problems, we have studied the relationship of some projection-type methods and suggested how to improve the efficiency of these methods. The essence of the relationship is similar to those in computational schemes for ordinary differential equations. From this point of view, all methods discussed in this paper can be viewed as one-step methods for VI(Ω, F) problems. A further question is whether there exist multistep methods for VI(Ω, F) problems as in numerical methods for ordinary differential equations. This question remains to be explored.

References

1. Glowinski, R., *Numerical Methods for Nonlinear Variational Problems*, Springer Verlag, New York, NY, 1984.
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. Xiao, B., and Harker, P. T., *A Nonsmooth Newton Method for Variational Inequalities, I: Theory*, Mathematical Programming, Vol. 65, pp. 151–194, 1994.
4. Xiao, B., and Harker, P. T., *A Nonsmooth Newton Method for Variational Inequalities, II: Numerical Results*, Mathematical Programming, Vol. 65, pp. 195–216, 1994.
5. Pang, J. S., *Inexact Newton Methods for the Nonlinear Complementarity Problem*, Mathematical Programming, Vol. 36, pp. 54–71, 1986.
6. Pang, J. S., and Qi, L., *Nonsmooth Equations: Motivation and Algorithms*, SIAM Journal on Optimization, Vol. 3, pp. 443–465, 1993.
7. Mangasarian, O. L., *Solution of Symmetric Linear Complementarity Problems by Iterative Methods*, Journal of Optimization Theory and Applications, Vol. 22, pp. 465–485, 1979.
8. He, B. S., *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.
9. He, B. S., *A New Method for a Class of Linear Variational Inequalities*, Mathematical Programming, Vol. 66, pp. 137–144, 1994.
10. He, B. S., *Solving a Class of Linear Projection Equations*, Numerische Mathematik, Vol. 68, pp. 71–80, 1994.
11. Goldstein, A. A., *Convex Programming in Hilbert Space*, Bulletin of the American Mathematical Society, Vol. 70, pp. 709–710, 1964.
12. Levitin, E. S., and Polyak, B. T., *Constrained Minimization Problems*, USSR Computational Mathematics and Mathematical Physics, Vol. 6, pp. 1–50, 1966.
13. Bertsekas, D. P., and Gafni, E. M., *Projection Methods for Variational Inequalities with Application to the Traffic Assignment Problem*, Mathematical Programming Study, Vol. 17, pp. 139–159, 1982.
14. Korpelevich, G. M., *The Extragradient Method for Finding Saddle Points and Other Problems*, Ekonomika i Matematicheskie Metody, Vol. 12, pp. 747–756, 1976.

15. Khobotov, E. N., *Modification of the Extragradient Method for Solving Variational Inequalities and Certain Optimization Problems*, USSR Computational Mathematics and Mathematical Physics, Vol. 27, pp. 120–127, 1987.
16. He, B. S., *A Class of Projection and Contraction Methods for Monotone Variational Inequalities*, Applied Mathematics and Optimization, Vol. 35, pp. 69–76, 1997.
17. Sun, D. F., *A Class of Iterative Methods for Solving Nonlinear Projection Equations*, Journal of Optimization Theory and Applications, Vol. 91, pp. 123–140, 1996.
18. Rockafellar, R. T., *Augmented Lagrangians and Applications of the Proximal Point Algorithm in Convex Programming*, Mathematics of Operations Research, Vol. 1, pp. 97–116, 1976.
19. Bertsekas, D. P., and Tsitsiklis, J. N., *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, New Jersey, 1989.
20. Harker, P. T., and Pang, J. S., *A Damped Newton Method for the Linear Complementarity Problem*, Lectures in Applied Mathematics Vol. 26, pp. 265–284, 1990.
21. Marcotee, P., and Dussault, J. P., *A Note on a Globally Convergent Newton Method for Solving Variational Inequalities*, Operation Research Letters, Vol. 6, pp. 35–42, 1987.
22. Taji, K., Fukushima, M., and Ibaraki, T., *A Globally Convergent Newton Method for Solving Strongly Monotone Variational Inequalities*, Mathematical Programming, Vol. 58, pp. 369–383, 1993.