A self-adaptive projection method with improved step-size for solving variational inequalities

Xihong Yan, Deren Han*, Wenyu Sun

Institute of Mathematics, School of Mathematics and Computer Science, Nanjing Normal University, Nanjing 210097, China

Abstract

In this paper, we propose a new projection method for solving variational inequality problems, which can be viewed as an improvement of the method of Han and Lo [D.R. Han, Hong K. Lo, Two new self-adaptive projection methods for variational inequality problems, Computers & Mathematics with Applications 43 (2002) 1529–1537], by adopting a new step-size rule. The method is as simple as Han and Lo’s methods [D.R. Han, Hong K. Lo, Two new self-adaptive projection methods for variational inequality problems, Computers & Mathematics with Applications 43 (2002) 1529–1537] and other extra-gradient-type methods, which uses only function evolutions and projections onto the feasible set. We prove that under the condition that the underlying function is co-coercive, the sequence generated by the method converges to a solution of the variational inequality problem globally. Some preliminary computational results are reported, which illustrated that the new method is more efficient than Han and Lo’s method [D.R. Han, Hong K. Lo, Two new self-adaptive projection methods for variational inequality problems, Computers & Mathematics with Applications 43 (2002) 1529–1537].

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Variational inequalities; Self-adaptive; Projection methods; Co-coercive mappings; Global convergence

1. Introduction

Let $\Omega$ be a nonempty closed convex subset of $\mathbb{R}^n$, and let $F$ be a mapping from $\mathbb{R}^n$ into itself. The variational inequality problem, denoted by $\text{VI}(F, \Omega)$, is to find a vector $u^* \in \Omega$, such that

$$F(u^*)^T(u - u^*) \geq 0, \quad \forall u \in \Omega.$$  \hspace{1cm} (1)

Problem $\text{VI}(F, \Omega)$ includes nonlinear complementarity problems (when $\Omega = \mathbb{R}^n_+$) and system of nonlinear equations (when $\Omega = \mathbb{R}^n$); it has many applications in the fields such as mathematical programming, network economics, transportation research, game theory and regional sciences; see the excellent monograph of Facchinei and Pang [2] and the references therein.

---

This research was supported by the NSFC grant 10501024 and NSF of Jiangsu Province at Grant No. BK2006214.

* Corresponding author.  
E-mail address: handeren@njnu.edu.cn (D. Han).

0898-1221/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.  
doi:10.1016/j.camwa.2007.05.008
One of the simplest methods for solving variational inequality problem is the projection-type methods, proposed by Goldstein [3], and Levitin and Polyak [4]. This projection method generally updates the iterations according to the following formula: given an arbitrary initial point $u^0 \in \mathbb{R}^n$,

$$u^{k+1} = P_\Omega[u^k - \beta_k F(u^k)],$$  
(2)

where $P_\Omega(\cdot)$ denotes the orthogonal projection map onto $\Omega$ and $\beta_k$ is a judiciously chosen positive step size. Under suitable assumptions, e.g. that $F$ is Lipschitz continuous with a constant $L > 0$

$$\|F(x) - F(y)\| \leq L\|x - y\|,$$  
(3)

strongly monotone with a constant modulus $\alpha > 0$

$$(x - y)^T(F(x) - F(y)) \geq \alpha\|x - y\|^2,$$  
(4)

and the step size $\beta_k$ satisfies

$$0 < \beta_L \leq \beta_k \leq \beta_U < \frac{2\alpha}{L^2}.$$  
(5)

this projection method is globally convergent. It is worthwhile to point out that the efficiency of this approach depends heavily on the estimations of the Lipschitz constant $L$ and the strongly monotone modulus $\alpha$. In fact, it might be difficult to estimate the modulus $L$ and $\alpha$ even if $F$ is an affine mapping. As a novel modification, He et al. [5] proposed to choose $\beta_k$ self-adaptively and Han and Sun [6] gave another self-adaptive rule.

The condition for the convergence of the method, i.e. the assumption of strong monotonicity, is stringent, which precludes the application of the method in reality. To overcome it, Korpelevich [7] proposed a new projection-type method, which is called extra-gradient method. For any given initial point $u^0 \in \Omega$, it generates a sequence of iterates according to the following recursion:

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

Under the conditions that the underlying mapping $F$ is monotone and Lipschitz continuous, and $0 < \beta_L \leq \beta_k \leq \beta_U < 1/L$, the method converges globally.

Recently, He [8], Sun [9] and Solodov and Tseng [10] gave a new projection and contraction method with the following recursion:

$$u^{k+1} = u^k - \gamma \rho(u^k, \beta_k)g(u^k, \beta_k),$$  
(6)

where

$$g(u, \beta) = e(u, \beta) - \beta[F(u) - F(u - e(u, \beta))],$$  
(7)

$$\rho(u, \beta) = e(u, \beta)^T \frac{g(u, \beta)}{\|g(u, \beta)\|^2},$$  
(8)

and

$$e(u, \beta) = u - P_\Omega[u - \beta F(u)],$$  
(9)

is the residual function. Under the condition that the underlying mapping $F$ is monotone, the method converges globally, for suitably chosen parameter $\beta_k$.

The direction $g(u^k, \beta_k)$ tends to zero when $u^k$ tends to the solution $u^*$, leading to a slow convergence behaviour of their methods (6)–(9). To avoid this, Han and Lo [1] suggest a new extra-gradient-type method: Let

$$d(u^k, \beta_k) = e(u^k, \beta_k) + \beta_k F(u^k - e(u^k, \beta_k)),$$  
(10)

$$\rho(u^k, \beta_k) = e(u^k, \beta_k)^T g(u^k, \beta_k)/\|d(u^k, \beta_k)\|^2,$$  
(11)
and
\[ u^{k+1} = P_\Omega[u^k - \gamma \rho(u^k, \beta_k)d(u^k, \beta_k)]. \tag{12} \]

From (10) and (11), it follows that when \( u^k \) converges to the solution \( u^* \), the search direction \( d(u^k, \beta_k) \) does not converge to zero; however, \( \rho(u^k, \beta_k) \to 0 \). In other words the step is too small. Therefore, there is no essential improvement of their method over the methods of (6)–(9) [8–10].

The motivation of this paper is to design a new practical and robust step-size choice rule for variational inequality problems by using search direction (10), which does not converges to zero. The new step-size rule is reminiscent of He and Liao’s rule [11]. Then we modify the extra-gradient-type method (10)–(12) by embodying this step-size rule and prove that this modified method has the global convergence property under the condition that the underlying mapping \( F(\cdot) \) is co-coercive. Our preliminary computational experience shows that the new algorithm is efficient for variational inequality problems.

The paper is organized as following: in the next section, we give some useful preliminaries. In Section 3, we describe the method formally and show its global convergence. We report our preliminary computational results in Section 4 and give some final conclusions in the last section.

2. Preliminaries

Now, let us summarize some basic properties and concepts that will be used in the subsequent sections.

First, we denote \( \|x\| = \sqrt{x^\top x} \) as the Euclidean norm. Let \( \Omega \) be a nonempty closed convex subset of \( \mathbb{R}^n \) and \( \Omega^* \) be the solution set of VI\((F, \Omega)\). Throughout the paper, we assume that \( \Omega^* \) is nonempty. Let \( P_\Omega(\cdot) \) denote the projection mapping from \( \mathbb{R}^n \) onto \( \Omega \), i.e.
\[ P_\Omega(v) = \arg\min\{\|v - u\| : u \in \Omega\}. \]

From the above definition, it follows that the projection mapping \( P_\Omega(\cdot) \) has the following two properties:
\[ \{v - P_\Omega(v)\}^\top\{w - P_\Omega(v)\} \leq 0, \quad \forall v \in \mathbb{R}^n, \forall w \in \Omega, \tag{13} \]
and
\[ (v - w)^\top(P_\Omega(v) - P_\Omega(w)) \geq \|P_\Omega(v) - P_\Omega(w)\|^2, \quad \forall v \in \mathbb{R}^n, \forall w \in \mathbb{R}^n. \tag{14} \]

Consequently, we have
\[
\begin{align*}
\|P_\Omega(v) - P_\Omega(w)\| &\leq \|v - w\|, \quad \forall v \in \mathbb{R}^n, \forall w \in \mathbb{R}^n, \tag{15} \\
\|P_\Omega(v) - u\|^2 &\leq \|v - u\|^2 - \|v - P_\Omega(v)\|^2, \quad \forall u \in \Omega. \tag{16}
\end{align*}
\]

**Lemma 2.1** ([12]). Let \( \beta > 0 \), then \( u^* \) solves VI\((F, \Omega)\) if and only if
\[ u^* = P_\Omega[u^* - \beta F(u^*)]. \]

It can be seen easily from the above lemma that solving VI\((F, \Omega)\) is equivalent to finding a zero point of \( e(u, \beta) \).

The following two lemmas give the properties of \( \|e(u, \beta)\| \) which are needed in our later convergence analysis.

**Lemma 2.2.** For all \( u \in \mathbb{R}^n \) and \( \bar{\beta} \geq \beta > 0 \), it holds that
\[ \|e(u, \bar{\beta})\| \geq \|e(u, \beta)\|. \tag{17} \]

**Proof.** See [13–15].

**Lemma 2.3.** For any \( u \in \Omega \) and \( \bar{\beta} \geq \beta > 0 \), we have
\[ \frac{\|e(u, \beta)\|}{\beta} \geq \frac{\|e(u, \bar{\beta})\|}{\bar{\beta}}. \tag{18} \]

**Proof.** See [13].
**Definition 2.1.** Let $c_0 > 0$ be a constant and $\varphi(u) : R^n \to R$ be a continuous function. We call $\varphi(u)$ an error measure function of VI($F$, $\Omega$) on $\Omega$ (or $R^n$) if it satisfies

$$\varphi(u) \geq c_0\|e(u, \beta)\|^2, \quad \forall u \in \Omega (or R^n),$$

and

$$\varphi(u) = 0 \Leftrightarrow e(u, \beta) = 0. \quad \square$$

**Definition 2.2.** Let $\Pi(u)$ be a function from $R^n$ into itself. We call $\Pi(u)$ a profitable direction of VI($F$, $\Omega$) if

$$(u - u^*)^T \Pi(u) \geq \varphi(u), \quad \forall u \in \Omega,$$

where $\varphi(u)$ is an error measure function and $u^*$ is a solution of VI($F$, $\Omega$). \quad \square

In fact, if $u$ is not a solution of VI($F$, $\Omega$), then

$$\left(\nabla \left(\frac{1}{2}\|u - u^*\|^2\right)\right)^T \Pi(u) \geq \varphi(u) \geq c_0\|e(u, \beta)\|^2 > 0.$$ 

In other words, $-\Pi(u)$ is a descent direction of the function $\frac{1}{2}\|u - u^*\|^2$, although $u^*$ is unknown.

In the following, we give the definitions of the underlying mapping $F(\cdot)$.

**Definition 2.3.** Let $F$ be a mapping from a set $\Omega \subseteq R^n \to R^n$, then

(a) $F$ is said to be monotone on $\Omega$, if

$$(u - v)^T (F(u) - F(v)) \geq 0 \quad \forall u, v \in \Omega;$$

(b) $F$ is said to be strictly monotone on $\Omega$, if

$$(u - v)^T (F(u) - F(v)) > 0 \quad \forall u, v \in \Omega, u \neq v;$$

(c) $F$ is said to be strongly monotone on $\Omega$ with modulus $\mu > 0$, if

$$(u - v)^T (F(u) - F(v)) \geq \mu\|u - v\|^2 \quad \forall u, v \in \Omega;$$

(d) $F$ is said to be co-coercive on $\Omega$ with modulus $\tau > 0$, if

$$(u - v)^T (F(u) - F(v)) \geq \tau\|F(u) - F(v)\|^2 \quad \forall u, v \in \Omega;$$

(e) $F$ is said to be Lipschitz continuous on $\Omega$ with modulus $L > 0$, if

$$\|F(u) - F(v)\| \leq L\|u - v\| \quad \forall u, v \in \Omega.$$

From Definition 2.3, it is clear that co-coercive mappings are monotone but not necessarily strictly or strongly monotone. Conversely, strongly monotone and Lipschitz continuous mappings are co-coercive. Thus, co-coercive is an intermediate concept between simple and strong monotonicity.

Now, we present a convergence theorem which is useful for the method studied in this paper.

**Theorem 2.1.** Let $C_0 > 0$ be a constant, $k \in \{0, 1\}$ be a given integer, $\{\beta_k\}$ be a positive sequence, and $\beta_k = \beta_{\min} > 0$. If the sequence $\{u^k\}$ generated by a method satisfies

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - C_0\|e(u^{k+1}, \beta_k)\|^2, \quad \forall u^* \in \Omega^*,$$

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

**Proof.** Let $u^*$ be a solution of VI($F$, $\Omega$). First, from (19), we get

$$\sum_{k=0}^{\infty} C_0\|e(u^{k+1}, \beta_k)\|^2 \leq \|u^0 - u^*\|^2.$$
which means that
\[ \lim_{k \to \infty} e(u^k, \beta_k) = 0. \]
Since \( \beta_k \geq \beta_{\text{min}} \), it follows from Lemma 2.2 that
\[ \lim_{k \to \infty} e(u^k, \beta_{\text{min}}) = 0. \]
Again, it follows from (19) that the sequence \( \{u^k\} \) is bounded and therefore it has at least one cluster point. Let \( u^* \) be a cluster point of \( \{u^k\} \), and let the subsequence \( \{u^{k_j}\} \) converge to \( u^* \). Because \( e(u, \beta_{\text{min}}) \) is continuous, taking limit along the subsequence,
\[ e(u^*, \beta_{\text{min}}) = \lim_{j \to \infty} e(u^{k_j}, \beta_{\text{min}}) = 0. \]
Thus, it follows from Lemma 2.1 that \( u^* \) is a solution of VI(\( F, \Omega \)).

In the following, we prove the sequence \( \{u^k\} \) has exactly one cluster point. Assume that \( \tilde{u} \) is another cluster point, and denote \( \delta := \|\tilde{u} - u^*\| > 0 \). Because \( u^* \) is a cluster point of the sequence \( \{u^k\} \), there is a \( k_0 > 0 \) such that
\[ \|u^{k_0} - u^*\| \leq \frac{\delta}{2}. \]
On the other hand, since \( u^* \in \Omega^* \) and thus
\[ \|u^k - u^*\| \leq \|u^{k_0} - u^*\|, \quad \forall k \geq k_0. \]
It follows that
\[ \|u^k - \tilde{u}\| \geq \|\tilde{u} - u^*\| - \|u^{k_0} - u^*\| \geq \frac{\delta}{2}, \quad \forall k \geq k_0. \]
This contradicts the assumption; thus the sequence \( \{u^k\} \) converges to \( u^* \in \Omega^* \).

In the rest of this paper, for convergence analysis, it is important to make the iteration sequence generated by the algorithm satisfies (19) in the Theorem 2.1.

3. The algorithms and convergence analysis

Based on the direction computed by (10), we propose a new step size along it, and thus come up with the following new method for solving VI(\( F, \Omega \)).

**Algorithm 3.1 (A new self-adaptive projection method).**

S0. Given \( \varepsilon > 0 \), choose \( u^0 \in \Omega, \mu \in (0, 1), \beta_0 = \beta = 1, \tau, \ L \in (0, 1), \) and \( \gamma \in (0, 2) \). Set \( k := 0 \).

S1. Compute \( e(u^k, \beta_k) \) by (9). If \( \|e(u^k, \beta_k)\| < \varepsilon \), stop; otherwise

S2. Find the smallest nonnegative integer \( m_k, \beta_k = \beta \mu^m \) satisfying
\[ \beta_k \|F(u^k) - F(u^k - e(u^k, \beta_k))\| \leq L \|e(u^k, \beta_k)\|. \]

S3. Calculate \( \alpha_k \) by
\[ \alpha_k = \gamma \min \left\{ \left( 1 - \frac{\beta_k}{4\tau} \right), \frac{g(u^k, \beta_k)^T e(u^k, \beta_k)}{\|g(u^k, \beta_k)\|^2 + 2e(u^k, \beta_k)^T g(u^k, \beta_k)} \right\}, \]
where \( g(u, \beta) \) is defined by (7).

S4. Compute
\[ u^{k+1} = P_{\Omega}[u^k - \alpha_k d(u^k, \beta_k)], \]
where \( d(u, \beta) \) is defined by (10).

S5. If
\[ \beta_k \|F(u^k) - F(u^k - e(u^k, \beta_k))\| \leq 0.4 \|e(u^k, \beta_k)\|, \]
\( \beta = \beta_k / 0.7; \) else \( \beta = \beta_k \). Set \( k := k + 1; \) go to Step 1.
Remark 3.1. It is clear that the proposed algorithm is a modification of the first method proposed by Han and Lo in [1] in the sense that we adopted a new stepsizes rule, which makes the method more efficient. If the iteration terminates after finite steps, then from Lemma 2.1, the current iterative point \( u^k \) is an approximate solution of VI. So we suppose in the following that the algorithm does not stop in finite many steps and an infinite sequence \( \{u^k\} \) is generated. In fact, from Lemma 3.1, the parameter sequence \( \{\beta_k\} \) is bounded below from zero, meaning that after a few trial steps, we can find \( \beta_k \) satisfying the condition in Step 2; Step 3 only involves some function evaluations; Step 4 needs function evaluation and a projection onto \( \Omega \) and Step 5 needs simple comparison. Thus, we can see that the whole algorithm is well defined.

In the following, we give convergence analysis of the Algorithm 3.1, beginning with a series of lemmas.

Lemma 3.1. If \( \|e(u, 1)\| \neq 0 \), then there exist \( 0 < L < 1 \) and \( \tilde{\beta} > 0 \), such that for all \( 0 < \beta \leq \tilde{\beta} \)

\[
\beta \|F(u) - F(u - e(u, \beta))\| \leq L\|e(u, \beta)\|.
\]

Proof. Suppose that (20) is not true, i.e.,

\[
\beta \|F(u) - F(u - e(u, \beta))\| > L\|e(u, \beta)\|, \quad \forall \beta > 0.
\]

Since \( F(\cdot) \) is continuous, \( e(u, \beta) \) is continuous. Let \( \beta \rightarrow 0 \) in the above inequality, we have

\[
\frac{L\|e(u, \beta)\|}{\beta} \leq 0.
\]

From Lemma 2.1 it follows that

\[
L\|e(u, 1)\| \leq 0.
\]

This implies

\[
\|e(u, 1)\| = 0,
\]

which contradict the assumption that \( \|e(u, 1)\| \neq 0 \). This completes the proof. \( \square \)

The following lemma give the bound of \( e(u, \beta)^T g(u, \beta) \), showing that the step size in the proposed algorithm is bounded away from zero.

Lemma 3.2. Under the assumption (20), we have

\[
e(u, \beta)^T g(u, \beta) \geq (1 - L)\|e(u, \beta)\|^2,
\]

\[
e(u, \beta)^T g(u, \beta) \geq \frac{1}{2}\|g(u, \beta)\|^2.
\]

Proof. From (7) and (20), we have

\[
e(u, \beta)^T g(u, \beta) = e(u, \beta)^T [e(u, \beta) - \beta(F(u) - F(u - e(u, \beta)))]
\]

\[
= \|e(u, \beta)\|^2 - \beta e(u, \beta)^T [F(u) - F(u - e(u, \beta))]
\]

\[
\geq \|e(u, \beta)\|^2 - L\|e(u, \beta)\|^2
\]

\[
= (1 - L)\|e(u, \beta)\|^2,
\]

and

\[
e(u, \beta)^T g(u, \beta) = e(u, \beta)^T [e(u, \beta) - \beta(F(u) - F(u - e(u, \beta)))]
\]

\[
= \|e(u, \beta)\|^2 - \beta e(u, \beta)^T [F(u) - F(u - e(u, \beta))]
\]

\[
\geq \frac{1}{2}\|e(u, \beta)\|^2 - \beta e(u, \beta)^T [F(u) - F(u - e(u, \beta))] + \frac{1}{2}\beta^2\|F(u) - F(u - e(u, \beta))\|^2
\]

\[
\geq \frac{1}{2}\|g(u, \beta)\|^2.
\]

This completes the proof. \( \square \)
From (22), it follows that
\[
\frac{e(u, \beta)^T g(u, \beta)}{2e(u, \beta)^T g(u, \beta) + \|g(u, \beta)\|^2} \geq \frac{e(u, \beta)^T g(u, \beta)}{4e(u, \beta)^T g(u, \beta)} = \frac{1}{4}.
\] (23)

The following lemma is from [1], which states that for $\beta$ satisfying (20), $g(u, \beta)$ is a profitable direction:

Lemma 3.3. Suppose that (20) is satisfied. Then for any $u \notin \Omega^*$,
\[
g(u, \beta)^T (u - u^*) \geq e(u, \beta)^T g(u, \beta) > 0. \quad \Box
\]

From the assumption that $F$ is co-coercive, we have that
\[
F(u)^T (u - u^*) \geq 0, \quad \forall u \in \Omega, u^* \in \Omega^*.
\] (24)

The following lemma states that if $g(u, \beta)$ is a profitable direction, then $d(u, \beta)$ defined by (10) is also a profitable direction.

Lemma 3.4. Under the assumption that $F(\cdot)$ is co-coercive on $\Omega$, $d(u, \beta)$ is a profitable direction.

Proof. From (7) and (10), we get
\[
d(u, \beta) = g(u, \beta) + \beta F(u).
\]
If $F(\cdot)$ is co-coercive on $\Omega$, then (24) holds and from Lemmas 3.2 and 3.3, it follows that
\[
(u - u^*)^T d(u, \beta) \geq e(u, \beta)^T g(u, \beta) \geq (1 - L)\|e(u, \beta)\|^2.
\]
Thus, $d(u, \beta)$ can be viewed as a profitable direction with $e(u, \beta)^T g(u, \beta)$ being the associated error measure function. \quad \Box

For the convenience of the analysis, in the following, we replace the generated point $u^k$, the step-size $\alpha_k$ and the parameter $\beta_k$ by $u$, $\alpha$ and $\beta$ respectively, and the new iterate of the method can be written as
\[
u(\alpha) = P_{\Omega}[u - \alpha d(u, \beta)].
\] (25)

In addition, we denote
\[
\Gamma(u, \beta) = F(u - e(u, \beta)),
\] (26)
and
\[
\Theta(\alpha) = \|u - u^*\|^2 - \|u(\alpha) - u^*\|^2.
\] (27)

It is clear that $\Theta(\alpha)$ is the difference between $\|u - u^*\|^2$ and $\|u(\alpha) - u^*\|^2$, let us observe $\Theta(\alpha)$. It follows from (16) and (25) that
\[
\|u(\alpha) - u^*\|^2 - \|u - \alpha d(u, \beta) - u^*\|^2
\leq \|u - u^*\|^2 - 2\alpha(u - u^*)^T d(u, \beta) - \alpha^2\|d(u, \beta)\|^2
- \|u - u(\alpha)\|^2 + 2\alpha(u - u(\alpha))^T d(u, \beta) - \alpha^2\|d(u, \beta)\|^2
\leq \|u - u^*\|^2 - 2\alpha(u - u^*)^T d(u, \beta) - \|u - u(\alpha)\|^2 + 2\alpha(u - u(\alpha))^T d(u, \beta).
\]
Substituting this into (27), we obtain
\[
\Theta(\alpha) \geq 2\alpha(u - u^*)^T d(u, \beta) + \|u - u(\alpha)\|^2 - 2\alpha(u - u(\alpha))^T d(u, \beta).
\] (28)

From (10) and (26), we can get
\[
\Theta(\alpha) \geq 2\alpha(u - u^*)^T [e(u, \beta) + \beta \Gamma(u, \beta)] + \|u - u(\alpha)\|^2
- 2\alpha(u - u(\alpha))^T e(u, \beta) - 2\alpha\beta(u - u(\alpha))^T \Gamma(u, \beta)
\] (29)
\[
2\alpha \beta (u - u^*)^T \Gamma (u, \beta) + \| u - u(\alpha) \|^2 - 2\alpha \beta (u - u(\alpha))^T \Gamma (u, \beta) + 2\alpha (u(\alpha) - u^*)^T e(u, \beta)
\]
\[
= 2\alpha \beta (u - u^*)^T \Gamma (u, \beta) + \| u - u(\alpha) - g(u, \beta) \|^2 + 2\alpha (u - u(\alpha))^T g(u, \beta)
\]
\[
- \alpha^2 \| g(u, \beta) \|^2 - 2\alpha \beta (u - u(\alpha))^T \Gamma (u, \beta) + 2\alpha (u(\alpha) - u^*)^T e(u, \beta).
\] (30)

It follows from (7) and (29) that
\[
\Theta(\alpha) \geq 2\alpha \beta (u - u^*)^T \Gamma (u, \beta) + 2\alpha (u - u(\alpha))^T [e(u, \beta) - \beta F(u)]
\]
\[
- \alpha^2 \| g(u, \beta) \|^2 + \| u - u(\alpha) - \alpha g(u, \beta) \|^2 + 2\alpha (u(\alpha) - u^*)^T e(u, \beta).
\] (31)

We have the following lemma, which gives a bound of the first term on the right-hand side of the inequality (31). For the proof, the reader is referred to [11].

Lemma 3.5. Under the assumption that \( F(\cdot) \) is co-coercive with respect to \( \Omega \), we have
\[
(u - u^*)^T \Gamma (u, \beta) \geq e(u, \beta)^T \Gamma (u, \beta), \quad \forall u \in \mathbb{R}^n, \ u^* \in \Omega^*.
\] (32)

The following lemma, gives a bound for the second term on right-hand side of inequality (31).

Lemma 3.6. We have the following inequality:
\[
(u - u(\alpha))^T [e(u, \beta) - \beta F(u)] \geq e(u, \beta)^T [e(u, \beta) - \beta F(u)], \quad \forall u \in \mathbb{R}^n.
\] (33)

Proof. Using the notation \( e(u, \beta) \), we get
\[
u - u(\alpha) = e(u, \beta) + P_{\Omega}[u - \beta F(u)] - u(\alpha).
\] (34)

From (13), we have
\[
\{ P_{\Omega}[u - \beta F(u)] - u(\alpha) \}^T [u - \beta F(u) - P_{\Omega}[u - \beta F(u)]] \geq 0.
\] (35)

Then
\[
\{ P_{\Omega}[u - \beta F(u)] - u(\alpha) \}^T [e(u, \beta) - \beta F(u)] \geq 0.
\] (36)

From (34) and (36), we obtain
\[
(u - u(\alpha) - e(u, \beta))^T (e(u, \beta) - \beta F(u)) \geq 0.
\] (37)

Thus,
\[
(u - u(\alpha))^T [e(u, \beta) - \beta F(u)] \geq e(u, \beta)^T [e(u, \beta) - \beta F(u)]. \quad \square
\]

Lemma 3.7. Suppose that \( F(\cdot) \) is co-coercive with constant \( \tau \) on \( \Omega \). Then,
\[
(u - u^*)^T e(u, \beta) \geq \left( 1 - \frac{\beta}{4\tau} \right) \| e(u, \beta) \|^2, \quad \forall u \in \Omega, \ u^* \in \Omega^*.
\] (38)

Proof. It follows from the definition of \( VI(F, \Omega) \), that
\[
(u - e(u, \beta) - u^*)^T F(u^*) \geq 0.
\] (39)

From (13), we have
\[
(u - u^* - e(u, \beta))^T (e(u, \beta) - \beta F(u)) \geq 0.
\] (40)

Inequalities (39) and (40) imply that
\[
[e(u, \beta) - \beta F(u) - F(u^*)]^T (u - u^*) - e(u, \beta)] \geq 0.
\] (41)
From (41), we get
\[(u - u^*)^T e(u, \beta) \geq \|e(u, \beta)\|^2 + \beta [F(u) - F(u^*)]^T (u - u^*) - \beta [F(u) - F(u^*)]^T e(u, \beta).\]

It follows from the Definition 2.3,
\[(u - u^*)^T e(u, \beta) \geq \|e(u, \beta)\|^2 - \beta [F(u) - F(u^*)]^T e(u, \beta) + \beta \|F(u) - F(u^*)\|^2 + \frac{\beta}{4\tau} \|e(u, \beta)\|^2 \geq \left(1 - \frac{\beta}{4\tau}\right) \|e(u, \beta)\|^2.

This completes the proof. □

Now, we observe the last two terms on the right-hand side of inequality (31) in the following lemma.

**Lemma 3.8.** Under the same assumption as Lemma 3.7, we have
\[
\|u - u(\alpha) - \alpha g(u, \beta)\|^2 + 2\alpha(u(\alpha) - u^*)^T e(u, \beta) \geq 2\alpha \left(1 - \frac{\beta}{4\tau}\right) - \alpha^2 \|e(u, \beta)\|^2 - 2\alpha^2 g(u, \beta)^T e(u, \beta).
\]

**Proof.** Let
\[
\Phi(\alpha) := \|u - u(\alpha) - \alpha g(u, \beta)\|^2 + 2\alpha(u(\alpha) - u^*)^T e(u, \beta).
\]
Then,
\[
\Phi(\alpha) = \|u - u(\alpha) - \alpha g(u, \beta)\|^2 + 2\alpha[u(\alpha) - u + \alpha g(u, \beta)]^T e(u, \beta) - 2\alpha^2 g(u, \beta)^T e(u, \beta) - 2\alpha g(u, \beta)^T e(u, \beta) - \alpha^2 \|e(u, \beta)\|^2.
\]
From Lemma 3.7, we get
\[
\Phi(\alpha) \geq \left[2\alpha \left(1 - \frac{\beta}{4\tau}\right) - \alpha^2 \right] \|e(u, \beta)\|^2 - 2\alpha^2 g(u, \beta)^T e(u, \beta).
\]
This completes the proof. □

From (7), Lemmas 3.5 and 3.6, it follows that
\[
\Theta(\alpha) \geq 2\alpha g(u, \beta)^T e(u, \beta) - \alpha^2 \|g(u, \beta)\|^2 + \|u - u(\alpha) - \alpha g(u, \beta)\|^2 + 2\alpha(u(\alpha) - u^*)^T e(u, \beta). \tag{44}
\]
From Lemma 3.8, we get
\[
\Theta(\alpha) \geq 2\alpha g(u, \beta)^T e(u, \beta) - \alpha^2 [\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta)] + \alpha \left[2 \left(1 - \frac{\beta}{4\tau}\right) - \alpha\right] \|e(u, \beta)\|^2. \tag{45}
\]
Let
\[
\Psi(\alpha) := 2\alpha g(u, \beta)^T e(u, \beta) - \alpha^2 [\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta)]
\]
Then, \(\Psi(\alpha)\) is a quadratic function of \(\alpha\), and it reaches its maximum at
\[
\alpha^* = \frac{g(u, \beta)^T e(u, \beta)}{\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta)}
\]
and
\[
\Phi(\alpha^*) = \frac{(g(u, \beta)^T e(u, \beta))^2}{\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta)}.
\]
From (21) and (23), it follows that

$$\Psi(\alpha^*) \geq \frac{(1 - L)}{4} \|e(u, \beta)\|^2. \quad (46)$$

When

$$0 < \alpha \leq 2 \left(1 - \frac{\beta}{4\tau}\right),$$

we have

$$\alpha \left[ 2 \left(1 - \frac{\beta}{4\tau}\right) - \alpha \right] \|e(u, \beta)\|^2 \geq 0.$$

Now, in the following lemma, we give our step size:

**Lemma 3.9.** Let $\beta_u$ is upper bound of $\beta$, let $\gamma \in (0, 2)$, $\beta_u \in (0, 4\tau)$ and

$$\alpha = \gamma \min \left\{ \left(1 - \frac{\beta}{4\tau}\right), \frac{g(u, \beta)^T e(u, \beta)}{\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta)} \right\}, \quad (47)$$

then, there exists a constant $C_0 > 0$, satisfying

$$\Theta(\alpha) \geq C_0 \|e(u, \beta)\|^2.$$

**Proof.** From (47), we get

$$\alpha = \gamma \left(1 - \frac{\beta}{4\tau}\right),$$

or

$$\alpha = \gamma \frac{g(u, \beta)^T e(u, \beta)}{\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta)}.$$

When $\alpha = \gamma \left(1 - \frac{\beta}{4\tau}\right)$, it follows from (45) that

$$\Theta(\alpha) \geq \gamma (2 - \gamma) \left(1 - \frac{\beta}{4\tau}\right)^2 \|e(u, \beta)\|^2 \geq \gamma (2 - \gamma) \left(1 - \frac{\beta_u}{4\tau}\right)^2 \|e(u, \beta)\|^2.$$

Denote $C_1 = \gamma (2 - \gamma) (1 - \frac{\beta_u}{4\tau})^2 > 0$, we have

$$\Theta(\alpha) \geq C_1 \|e(u, \beta)\|^2.$$

When $\alpha = \gamma \frac{g(u, \beta)^T e(u, \beta)}{\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta)}$, from (45), we obtain

$$\Theta(\alpha) \geq \alpha \left\{ 2g(u, \beta)^T e(u, \beta) - \gamma \frac{g(u, \beta)^T e(u, \beta)}{\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta)} \|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta) \right\}$$

$$= \alpha (2 - \gamma) g(u, \beta)^T e(u, \beta)$$

$$= \gamma (2 - \gamma) (g(u, \beta)^T e(u, \beta))^2$$

$$\|g(u, \beta)\|^2 + 2g(u, \beta)^T e(u, \beta).$$

From (46), it follows that:

$$\Theta(\alpha) \geq \frac{\gamma (2 - \gamma) (1 - L)}{4} \|e(u, \beta)\|^2.$$
Denote $C_2 = \frac{\gamma(2-\gamma)(1-L)}{4} > 0$, we have
\[
\Theta(\alpha) \geq C_2 \|e(u, \beta)\|^2,
\]
let $C_0 = \min\{C_1, C_2\} > 0$, We obtain
\[
\Theta(\alpha) \geq C_0 \|e(u, \beta)\|^2
\]
This completes the proof. \(\square\)

**Remark 1.** From (23) and (47), we have
\[
\alpha \geq \gamma \min \left\{ \left(1 - \frac{\beta_u}{4 \tau}\right), \frac{1}{4} \right\}.
\]
It is clear that our step size is larger than a constant.

We summarize the analytical result of this section in the following lemma:

**Theorem 3.1.** Let $C_0 > 0$ be a constant, let
\[
g(u, \beta) = e(u, \beta) - \beta[F(u) - F(u - e(u, \beta))],
\]
\[
d(u, \beta) = e(u, \beta) + \beta[F(u - e(u, \beta))].
\]
For given $u^k \in \Omega$, $\beta_k$ is chosen such that
\[
\beta_k \|F(u^k) - F(u^k - e(u^k, \beta_k))\| \leq L\|e(u^k, \beta_k)\|, \quad L \in (0, 1),
\]
then under the assumption that $F(\cdot)$ is co-coercive on $\Omega$, the method
\[
u^{k+1} = P_\Omega[u^k - \alpha_k d(u^k, \beta_k)],
\]
with step-size
\[
\alpha_k = \gamma \min \left\{ \left(1 - \frac{\beta_k}{4 \tau}\right), \frac{g(u^k, \beta_k)^T e(u^k, \beta_k)}{\|g(u^k, \beta_k)\|^2 + 2g(u^k, \beta_k)^T e(u^k, \beta_k)} \right\},
\]
produce a new iterate which satisfies
\[
\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - C_0 \|e(u^k, \beta_k)\|^2, \quad \forall u^* \in \Omega^*.
\]

From Theorems 2.1 and 3.1, we obtain the following theorem that states the global convergence of the Algorithm 3.1.

**Theorem 3.2.** Suppose $F(\cdot)$ is co-coercive on $\Omega$, and the solution set $\Omega^*$ is nonempty, then the sequence $\{u^k\} \subset \mathbb{R}^n$ generated by Algorithm 3.1 convergence to a solution of $VI(F, \Omega)$.

## 4. Computational results

In this section, we give some preliminary computational results. We implement our Algorithm 3.1 in MATLAB to solve some complementarity problems. Our main purpose is to show the advantages of the proposed step-size strategy over the old one. To this end, we also code Algorithm 3.1 of Han and Lo [1].

The first problem under consideration is $F(u) = Mu + q$, where
\[
M = \begin{bmatrix}
1 & 2 & \cdots & \cdots & 2 \\
0 & 1 & 2 & \cdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots \\
0 & \cdots & \cdots & 0 & 1
\end{bmatrix}, \quad q = (-1, -1, \ldots, -1)^T.
\]
Table 1
Computational results for \( u^0 = (0, 0, \ldots, 0) \)

<table>
<thead>
<tr>
<th>Dim ((n))</th>
<th>(\text{Algorithm 3.1 in [1]})</th>
<th>(\text{Cpu})</th>
<th>(\text{Our Algorithm 3.1})</th>
<th>(\text{Cpu})</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>73</td>
<td>0.02</td>
<td>17</td>
<td>0.01</td>
</tr>
<tr>
<td>200</td>
<td>74</td>
<td>0.0701</td>
<td>17</td>
<td>0.02</td>
</tr>
<tr>
<td>300</td>
<td>62</td>
<td>0.1662</td>
<td>17</td>
<td>0.0601</td>
</tr>
<tr>
<td>400</td>
<td>62</td>
<td>0.4907</td>
<td>17</td>
<td>0.2003</td>
</tr>
<tr>
<td>500</td>
<td>46</td>
<td>0.6209</td>
<td>18</td>
<td>0.3205</td>
</tr>
<tr>
<td>600</td>
<td>55</td>
<td>1.0515</td>
<td>17</td>
<td>0.4206</td>
</tr>
<tr>
<td>700</td>
<td>57</td>
<td>1.3920</td>
<td>18</td>
<td>0.5808</td>
</tr>
<tr>
<td>800</td>
<td>66</td>
<td>2.3033</td>
<td>17</td>
<td>0.7311</td>
</tr>
<tr>
<td>900</td>
<td>55</td>
<td>2.2132</td>
<td>18</td>
<td>0.9313</td>
</tr>
<tr>
<td>1000</td>
<td>67</td>
<td>12.4679</td>
<td>18</td>
<td>1.1116</td>
</tr>
<tr>
<td>2000</td>
<td>64</td>
<td>14.5209</td>
<td>14</td>
<td>4.0558</td>
</tr>
<tr>
<td>3000</td>
<td>79</td>
<td>37.5840</td>
<td>16</td>
<td>10.0044</td>
</tr>
</tbody>
</table>

Table 2
Computational results for \( u^0 \) generated uniformly in \((0, 1)\)

<table>
<thead>
<tr>
<th>Dim ((n))</th>
<th>(\text{Algorithm 3.1 in [1]})</th>
<th>(\text{Cpu})</th>
<th>(\text{Our Algorithm 3.1})</th>
<th>(\text{Cpu})</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>69</td>
<td>0.02</td>
<td>16</td>
<td>0.01</td>
</tr>
<tr>
<td>200</td>
<td>71</td>
<td>0.0701</td>
<td>14</td>
<td>0.02</td>
</tr>
<tr>
<td>300</td>
<td>78</td>
<td>0.1602</td>
<td>15</td>
<td>0.0401</td>
</tr>
<tr>
<td>400</td>
<td>53</td>
<td>0.3906</td>
<td>14</td>
<td>0.1903</td>
</tr>
<tr>
<td>500</td>
<td>64</td>
<td>0.8412</td>
<td>14</td>
<td>0.2403</td>
</tr>
<tr>
<td>600</td>
<td>47</td>
<td>0.8813</td>
<td>12</td>
<td>0.3805</td>
</tr>
<tr>
<td>700</td>
<td>71</td>
<td>1.6323</td>
<td>15</td>
<td>0.4807</td>
</tr>
<tr>
<td>800</td>
<td>65</td>
<td>2.0129</td>
<td>15</td>
<td>0.6610</td>
</tr>
<tr>
<td>900</td>
<td>57</td>
<td>2.2532</td>
<td>15</td>
<td>0.8212</td>
</tr>
<tr>
<td>1000</td>
<td>87</td>
<td>4.0859</td>
<td>15</td>
<td>1.0215</td>
</tr>
<tr>
<td>2000</td>
<td>75</td>
<td>14.5209</td>
<td>14</td>
<td>4.0558</td>
</tr>
<tr>
<td>3000</td>
<td>79</td>
<td>37.5840</td>
<td>16</td>
<td>10.0044</td>
</tr>
</tbody>
</table>

In our test, we take \( \gamma = 1.8, L = 0.95, \mu = 0.7, \tau = 0.9 \). The stop criterion is that \( \| e(u, \beta) \| \leq 10^{-6} \). Tables 1 and 2 list the computational results with the initial point \( u^0 = (0, 0, \ldots, 0) \) and \( u^0 \) generated uniformly in \((0, 1)\), respectively.

From Tables 1 and 2, we can see that our Algorithm 3.1 is more efficient than Algorithm 3.1 in [1]. For all scale of the problem (the number of variables), the number of iterations is much less than that of [1] and the CPU time needed in our algorithm is about \( 1/3 \) of that needed for Algorithm 3.1 of [1]. The reason is that the step-size becomes larger. The results confirm that the new step size is useful to improve the efficiency.

We now consider a nonlinear complementarity problem

\[
\begin{align*}
  u & \geq 0, & F(u) & \geq 0, & \langle u, F(u) \rangle = 0,
\end{align*}
\]

where

\[
F(u) = D(u) + Mu + q,
\]

\(D(u)\) and \(Mu + q\) are the nonlinear part and the linear part of \( F(u) \), respectively. We form \( F(u) \) similarly as in [11]. 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-symmetrical matrix \( B \) is generated in the same way. The vector \( q \) is generated from a uniform distribution in the interval \((-500, 500)\). 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 [16–18].
We solve this problem with our Algorithm 3.1 and the original algorithm of Han and Lo [1] with different starting points. The parameters in the algorithms are $\gamma = 1.8$, $L = 0.8$, $\mu = 0.7$, $\tau = 0.9$. Tables 3–5 give the computational results.

From Tables 3–5 we can also observe that improvement strategy is effective. In addition, for a set of similar problems, it seems that the number of iterations is not very sensitive to the problem size and starting point too.
5. Conclusion

In this paper, we gave a self-adaptive projection method with improved step size for variational inequality problems. Under the condition that $F$ is co-coercive, the convergence of the algorithm is proved and our preliminary computational results indicated the efficiency introduced by the new strategy.

In our implementation, we set $\tau = 0.9$ in both examples. This does not mean that both examples have $\tau = 0.9$ as their co-coercivity constants. This is just a guess and they may have larger co-coercivity constants. In fact, choosing a suitable parameter $\tau$ is difficult in practice, as choosing the strong monotonicity modulus and the Lipschitz continuous constant. Thus, it is important to find a self-adaptive scheme to choose such parameters self-adaptively. This is one of our ongoing research topics.

References