An additional projection step to He and Liao’s method for solving variational inequalities

Abdellah Bnouhachem
School of Management Science and Engineering, Nanjing University, Nanjing 210093, PR China

Received 1 April 2006; received in revised form 28 June 2006

Abstract
In this paper, we proposed a modified extragradient method for solving variational inequalities. The method can be viewed as an extension of the method proposed by He and Liao [Improvement of some projection methods for monotone variational inequalities, J. Optim. Theory Appl. 112 (2002) 111–128], by performing an additional projection step at each iteration and another optimal step length is employed to reach substantial progress in each iteration. We used a self-adaptive technique to adjust parameter \( \rho \) at each iteration. Under certain conditions, the global convergence of the proposed method is proved. Preliminary numerical experiments are included to compare our method with some known methods.

MSC: 49J40; 65N30

Keywords: Variational inequalities; Self-adaptive rules; Projection method; Pseudomonotone operators

1. Introduction
Variational inequality and complementarity problems are very powerful tools of the current mathematical technology. A large number of problems arising in various branches of pure and applied sciences can be studied in the unified framework of variational inequalities. In recent years, classical variational inequality and complementarity problems have been extended and generalized to study a wide range of problems arising in mechanics, physics, optimization and applied sciences, see [1–24]. We now have a variety of techniques to suggest and analyze various iterative algorithms for solving variational inequalities and the related optimization problems. The fixed-point theory has played an important role in the development of various algorithms for solving variational inequalities. The basic idea is very simple. Using the projection operator technique, one usually establishes an equivalence between the variational inequalities and the fixed-point problem. This alternative equivalent formulation was used by Lions and Stampacchia [13] to study the existence of a solution of the variational inequalities. Projection methods and its variants forms represent important tools for finding the approximate solution of variational inequalities. It is well-known that the convergence of the projection method requires that the operator must be strongly monotone and Lipschitz continuous. Unfortunately, these strict conditions rule out many applications of this method. This fact motivated to modify the projection method or to...
develop other methods. The extragradient method \cite{8,10–12,14,17,18,20,21,23} overcomes this difficulty by performing an additional forward step and a projection at each iteration according to double projection. Its convergence requires only that a solution exists and the monotone operator is Lipschitz continuous. When the operator is not Lipschitz continuous or when the Lipschitz continuous is not known, the extragradient method and its variant forms require an Armijo-like line search procedure to compute the step size with a new projection need for each trial, which leads to expensive computation. To overcome these difficulties, several modified projection and extragradient-type methods \cite{1,7–10,17,18,20,21,23,24} have been suggested and developed for solving variational inequality problems. Using essentially the idea and technique of He and Liao \cite{8}, we suggest and analyze a new method for solving variational inequalities. The proposed method consists of three steps, using step 1 and step 2 of \cite{8} and we propose a new third step. We use a self-adaptive technique. The main contribution of this technique is that we allow elements of the penalty sequence to either increase or decrease in iterations, not necessarily monotone. Under certain conditions, the global convergence is proved. Some preliminary computational results are given to illustrate the efficiency of the proposed method.

2. Preliminaries

A classical variational inequality problem, denoted by VI\((T, K)\), is to find a vector \(u^* \in K\) such that

\[
(T(u^*), v - u^*) \geq 0, \quad \forall v \in K, \quad (2.1)
\]

where \(K \subset R^n\) is a nonempty closed convex subset of \(R^n\) and \(T\) is a mapping from \(R^n\) into itself.

First, we give some results which will be used in latter analysis.

**Lemma 2.1.** For a given \(u \in K, z \in R^n\) satisfy the inequality

\[
\langle u - z, v - u \rangle \geq 0, \quad \forall v \in K
\]

holds if and only if \(u = P_K(z)\).

It follows from Lemma 2.1 that

\[
\langle z - P_K(z), P_K(z) - v \rangle \geq 0, \quad \forall z \in R^n, \quad v \in K. \quad (2.2)
\]

It follows that

\[
\|P_K(z) - v\| \leq \|z - v\|, \quad \forall z \in R^n, \quad v \in K, \quad (2.3)
\]

\[
\|P_K(z) - v\|^2 \leq \|z - v\|^2 - \|z - P_K(z)\|^2, \quad \forall z \in R^n, \quad v \in K. \quad (2.4)
\]

It is well-known that the projection operator \(P_K\) is nonexpansive, that is,

\[
\|P_K(u) - P_K(v)\| \leq \|u - v\|, \quad \forall u, v \in R^n. \quad (2.5)
\]

**Lemma 2.2.** \(u^*\) is a solution of problem (2.1) if and only if \(u^*\) satisfies the relation:

\[
u^* = P_K[u^* - \rho T(u^*)] \quad \text{where } \rho > 0. \quad (2.6)
\]

From Lemma 2.2, it is clear that \(u\) is a solution of (2.1) if and only if \(u\) is a zero point of the function

\[
r(u, \rho) := u - P_K[u - \rho T(u)].
\]

The next lemma shows that \(\|r(u, \rho)\|\) is a nondecreasing function with respect to \(\rho\).

**Lemma 2.3** \((\text{Calamai and Moré [2], Gafni and Bertsekas [4], Peng and Fukushima [19]}). For all \(u \in H and \rho'/ \rho > 0\), it holds that

\[
\|r(u, \rho')\| \geq \|r(u, \rho)\|. \quad (2.7)
\]


In what follows, we always assume that the underlying function is continuous and pseudomonotone on \( \mathbb{R}^n \), i.e.,
\[
\langle T(u), u' - u \rangle \geq 0 \Rightarrow \langle T(u'), u' - u \rangle \geq 0 \quad \forall u', u \in \mathbb{R}^n,
\]
and the solution set of problem (2.1), denoted by \( S^* \), is nonempty.

3. Basic results

In this section, we suggest and consider the new modified extragradient method for solving variational inequality (2.1). For a given \( u^k \in K \), each iteration of the proposed method consists of three steps.

**Algorithm 3.1.**

**Step 1.** Compute
\[
\tilde{u}^k = P_K[u^k - \rho_k T(u^k)],
\]
where \( \rho_k > 0 \) satisfies
\[
\|\rho_k (T(u^k) - T(\tilde{u}^k))\| \leq \delta \|u^k - \tilde{u}^k\|, \quad 0 < \delta < 1.
\]

**Step 2.**
\[
\bar{u}^k = P_K[u^k - \alpha_k \rho_k T(\tilde{u}^k)],
\]
where
\[
\epsilon^k := \rho_k (T(\tilde{u}^k) - T(u^k)),
\]
\[
d(u^k, \rho_k) := u^k - \tilde{u}^k + \epsilon^k,
\]
\[
\phi(u^k, \rho_k) := \langle u^k - \tilde{u}^k, d(u^k, \rho_k) \rangle
\]
and
\[
\alpha_k := \frac{\phi(u^k, \rho_k)}{\|d(u^k, \rho_k)\|^2}.
\]

**Step 3.** For \( \tau > 0 \), the new iterate \( u^{k+1}(\tau) \) is defined by
\[
u^{k+1}(\tau) = P_K[u^k - \tau (u^k - \bar{u}^k)].
\]

How to choose a suitable step length \( \tau > 0 \) to force convergence will be discussed latter.

**Remark 3.1.** (3.2) implies that
\[
|\langle u^k - \bar{u}^k, \epsilon^k \rangle| \leq \delta \|u^k - \tilde{u}^k\|^2, \quad 0 < \delta < 1.
\]

For the convergence analysis of the proposed method, we need the following results.

**Lemma 3.1.** For given \( u^k \in \mathbb{R}^n \) and \( \rho_k > 0 \), let \( \tilde{u}^k \) and \( \epsilon^k \) satisfy (3.1) and (3.3), then
\[
\phi(u^k, \rho_k) \geq (1 - \delta) \|u^k - \tilde{u}^k\|^2
\]
and
\[
\alpha_k \geq \frac{1}{\sigma}.
\]
Proof. It follows from (3.4) and (3.8) that
\[ \phi(u^k, \rho_k) := \langle u^k - \tilde{u}^k, d(u^k, \rho_k) \rangle \]
\[ = \|u^k - \tilde{u}^k\|^2 + \langle u^k - \tilde{u}^k, \varepsilon_k \rangle \]
\[ \geq (1 - \delta)\|u^k - \tilde{u}^k\|^2. \]

Otherwise, we have
\[ \langle u^k - \tilde{u}^k, d(u^k, \rho_k) \rangle = \|u^k - \tilde{u}^k\|^2 + \langle u^k - \tilde{u}^k, \varepsilon_k \rangle \]
\[ \geq \frac{1}{2}\|u^k - \tilde{u}^k\|^2 + \langle u^k - \tilde{u}^k, \varepsilon_k \rangle + \frac{1}{2}\|\varepsilon_k\|^2 \]
\[ = \frac{1}{2}\|d(u^k, \rho_k)\|^2, \]
where the inequality follows from (3.2) and
\[ \varepsilon_k \geq \frac{1}{2}, \]
we can get the assertion of this lemma. □

Remark 3.2. Since \( \tilde{u}^k \in K \), it follows from (2.1) that
\[ \langle T(u^*), \tilde{u}^k - u^* \rangle \geq 0. \quad (3.11) \]

Under the assumption that \( T \) is pseudomonotone we have
\[ \langle T(\tilde{u}^k), \tilde{u}^k - u^* \rangle \geq 0. \quad (3.12) \]

It follows from (3.12) that
\[ \langle T(\tilde{u}^k), u^k - u^* \rangle \geq \langle T(\tilde{u}^k), u^k - \tilde{u}^k \rangle. \quad (3.13) \]

We now consider the criteria of \( \tau \), which ensures that \( u^{k+1}(\tau) \) is closer to the solution set than \( u^k \). For this purpose, we define
\[ \Gamma(\tau) := \|u^k - u^*\|^2 - \|u^{k+1}(\tau) - u^*\|^2. \quad (3.14) \]

Lemma 3.2. Let \( u^* \in S^* \). Then we have
\[ \Gamma(\tau) \geq \tau\{\|u^k - \tilde{u}^k\|^2 + \Theta_k(\varepsilon_k)\} - \tau^2\|u^k - \tilde{u}^k\|^2, \quad (3.15) \]
where
\[ \Theta_k(\varepsilon_k) := \|u^k - u^*\|^2 - \|\tilde{u}^k - u^*\|^2. \quad (3.16) \]

Proof. It follows from (2.3) and (3.7) that
\[ \Gamma(\tau) \geq \|u^k - u^*\|^2 - \|u^k - \tau(u^k - \tilde{u}^k) - u^*\|^2 \]
\[ = 2\tau\langle u^k - u^*, u^k - \tilde{u}^k \rangle - \tau^2\|u^k - \tilde{u}^k\|^2 \]
\[ = 2\tau\{\|u^k - \tilde{u}^k\|^2 - \langle u^* - \tilde{u}^k, u^k - \tilde{u}^k \rangle \} - \tau^2\|u^k - \tilde{u}^k\|^2. \quad (3.17) \]
Using the following identity
\[ \langle u^* - \tilde{u}^k, u^k - \tilde{u}^k \rangle = \frac{1}{2}\{\|\tilde{u}^k - u^*\|^2 - \|u^k - u^*\|^2\} + \frac{1}{2}\|u^k - \tilde{u}^k\|^2, \]
and the notation of \( \Theta_k(\varepsilon_k) \), we obtain (3.15), the required result. □
Next theorem has already been studied in [8]. For the sake of completeness and to convey an idea of the technique involved, we include its proof.

**Theorem 3.1 (He and Liao’s [8]).** For given \( u^k \in K, u^* \in S^* \) and \( \rho_k > 0 \), let \( \tilde{u}^k \) and \( \varepsilon^k \) satisfy (3.1) and (3.3). Then

\[
\Theta_k(z_k) \geq \gamma_k(z_k),
\]

where

\[
\gamma_k(z_k) := 2\varepsilon_k(u^k - \tilde{u}^k, d(u^k, \rho_k)) - \varepsilon_k^2\|d(u^k, \rho_k)\|^2
\]

and \( d(u^k, \rho_k) \) is defined in (3.4).

**Proof.** Since \( u^* \in K \) and \( \tilde{u}^k = P_K[u^k - \varepsilon_k \rho_k T(\tilde{u}^k)] \), it follows from (2.4) that

\[
\|\tilde{u}^k - u^*\|^2 \leq \|u^k - \tilde{u}^k - z_k \rho_k T(\tilde{u}^k)\|^2 - \|u^k - u^* - z_k \rho_k T(\tilde{u}^k)\|^2.
\]

Consequently, using the definition of \( \Theta_k(z_k) \), we get

\[
\Theta_k(z_k) \geq \|u^k - u^*\|^2 + \|u^k - \tilde{u}^k - z_k \rho_k T(\tilde{u}^k)\|^2 - \|u^k - u^* - z_k \rho_k T(\tilde{u}^k)\|^2
\]

\[
= \|u^k - \tilde{u}^k\|^2 + 2\varepsilon_k \rho_k \langle \tilde{u}^k - u^k, T(\tilde{u}^k) \rangle + 2\varepsilon_k \rho_k \langle u^k - \tilde{u}^k, T(\tilde{u}^k) \rangle.
\]

Applying (3.13) to the last term in the right side of the above inequality, we obtain

\[
\Theta_k(z_k) \geq \|u^k - \tilde{u}^k\|^2 + 2\varepsilon_k \rho_k \langle \tilde{u}^k - u^k, T(\tilde{u}^k) \rangle + 2\varepsilon_k \rho_k \langle u^k - \tilde{u}^k, T(\tilde{u}^k) \rangle
\]

\[
= \|u^k - \tilde{u}^k\|^2 + 2\varepsilon_k \rho_k \langle \tilde{u}^k - u^k, d(u^k, \rho_k) \rangle.
\]

Using \( \|a\|^2 \geq 2(a, b) - \|b\|^2 \) and the definition of \( \gamma_k(z_k) \), we have

\[
\|u^k - \tilde{u}^k\|^2 \geq 2\varepsilon_k (u^k - \tilde{u}^k, d(u^k, \rho_k)) - \varepsilon_k^2\|d(u^k, \rho_k)\|^2
\]

\[
= 2\varepsilon_k (u^k - \tilde{u}^k, \tilde{u}^k - \tilde{u}^k, d(u^k, \rho_k)) - \varepsilon_k^2\|d(u^k, \rho_k)\|^2
\]

\[
= \gamma_k(z_k) + 2\varepsilon_k \langle \tilde{u}^k - u^k, d(u^k, \rho_k) \rangle.
\]

Substituting (3.22) into (3.21), we obtain

\[
\Theta_k(z_k) \geq \gamma_k(z_k) + 2\varepsilon_k \langle \tilde{u}^k - u^k, d(u^k, \rho_k) - \rho_k T(\tilde{u}^k) \rangle.
\]

Now we consider the last term in the right side of (3.23). Setting \( z := u^k - \rho_k T(\tilde{u}^k) + \varepsilon^k \) and \( v := \tilde{u}^k \) in (2.2), we get

\[
\langle u^k - \rho_k T(\tilde{u}^k) + \varepsilon^k - P_K[u^k - \rho_k T(\tilde{u}^k) + \varepsilon^k], P_K[u^k - \rho_k T(\tilde{u}^k) + \varepsilon^k] - \tilde{u}^k \rangle \geq 0.
\]

Since \( \tilde{u}^k = P_K[u^k - \rho_k T(\tilde{u}^k) + \varepsilon^k] \) (see (3.1)) and \( d(u^k, \rho_k) = u^k - \tilde{u}^k + \varepsilon^k \) (see (3.4)), it follows from the above inequality that

\[
(d(u^k, \rho_k) - \rho_k T(\tilde{u}^k), \tilde{u}^k - \tilde{u}^k) \geq 0.
\]

Substituting (3.24) into (3.23) we obtain \( \Theta_k(z_k) \geq \gamma_k(z_k) \) and the assertion of this theorem is proved. \( \square \)

Using Lemma 3.2, (3.6) and Theorem 3.1, we get

\[
\Gamma(\tau) \geq A(\tau),
\]

where

\[
A(\tau) = \tau\|u^k - \tilde{u}^k\|^2 + \varepsilon_k \phi(u^k, \rho_k) - \tau^2\|u^k - \tilde{u}^k\|^2.
\]
The above inequality tells us how to choose a suitable \( \tau_k \). Since \( A(\tau_k) \) is a quadratic function of \( \tau_k \) and it reaches its maximum at

\[
\tau_k^* = \frac{\|u^k - \tilde{u}^k\|^2 + z_k \phi(u^k, \rho_k)}{2\|u^k - \tilde{u}^k\|^2}
\]

and

\[
A(\tau_k^*) = \frac{\tau_k^* \{\|u^k - \tilde{u}^k\|^2 + z_k \phi(u^k, \rho_k)\}}{2}.
\]

(3.27)

Then from Lemma 3.1, we get

\[
\tau_k^* \geq \frac{1 - \delta}{2} \left( \frac{\|u^k - \tilde{u}^k\|^2 + \|u^k - \tilde{u}^k\|^2}{2\|u^k - \tilde{u}^k\|^2} \right)
\]

\[
\geq \frac{1 - \delta}{4},
\]

(3.28)

and

\[
A(\tau_k^*) \geq \frac{\tau_k^* (1 - \delta)}{4} \|u^k - \tilde{u}^k\|^2
\]

\[
\geq \frac{(1 - \delta)^2}{16} \|u^k - \tilde{u}^k\|^2.
\]

(3.29)

For fast convergence, we take a relaxation factor \( \gamma \in [1, 2) \) and the step-size \( \tau_k = \gamma \tau_k^* \). Simple calculations show that

\[
A(\gamma \tau_k^*) = \gamma \tau_k^* \{\|u^k - \tilde{u}^k\|^2 + z_k \phi(u^k, \rho_k)\} - (\gamma^2 \tau_k^*) (\tau_k^* \|u^k - \tilde{u}^k\|^2)
\]

\[
= \gamma (2 - \gamma) A(\tau_k^*).
\]

(3.30)

4. Convergence of the proposed method

In this section, we prove the global convergence of the proposed method. The following theorem plays a crucial role in the convergence of the proposed method.

**Theorem 4.1.** Let \( u^* \) be a solution of problem (2.1) and let \{\( u^k \)\} be the sequence obtained from Algorithm 3.1. Then \{\( u^k \)\} is bounded and

\[
\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \frac{1}{16} \gamma (2 - \gamma) (1 - \delta)^2 \|u^k - \tilde{u}^k\|^2.
\]

(4.1)

**Proof.** Let \( u^* \) be a solution of problem (2.1). Then, from (3.25), (3.30) and (3.29), we have

\[
\|u^{k+1} - u^*\|^2 = \|u^k - u^*\|^2 - \Gamma(\gamma \tau_k^*)
\]

\[
\leq \|u^k - u^*\|^2 - \gamma (2 - \gamma) A(\tau_k^*)
\]

\[
\leq \|u^k - u^*\|^2 - \frac{1}{16} \gamma (2 - \gamma) (1 - \delta)^2 \|u^k - \tilde{u}^k\|^2.
\]

Since \( \gamma \in [1, 2) \) and \( \delta \in (0, 1) \) we have

\[
\|u^{k+1} - u^*\| \leq \|u^k - u^*\| \leq \cdots \leq \|u^0 - u^*\|.
\]

Then the sequence \( u^k \) is bounded. \( \square \)

We are now in the stage to prove the convergence of the proposed method.

**Theorem 4.2.** The sequence \{\( u^k \)\} generated by the proposed method converges to a solution point of problem (2.1).
Proof. It follows from (4.1) that
\[ \sum_{k=0}^{\infty} \| u^k - \tilde{u}^k \|^2 < \infty, \]
which means that
\[ \lim_{k \to \infty} \| u^k - \tilde{u}^k \| = 0. \] (4.2)
Consequently \{\tilde{u}^k\} is also bounded. Since \( r(u, \rho) \) is a nondecreasing function of \( \rho \), it follows from \( \rho_k \geq \rho_{\text{min}} \) that
\[ \| r(\tilde{u}^k, \rho_{\text{min}}) \| \leq \| r(\tilde{u}^k, \rho_k) \| = \| \tilde{u}^k - P_K[\tilde{u}^k - \rho_k T(\tilde{u}^k)] \| \]
(using (3.1) and (3.3)) = \[ \| P_K[u^k - \rho_k T(\tilde{u}^k) + \varepsilon^k] - P_K[\tilde{u}^k - \rho_k T(\tilde{u}^k)] \| \]
(using (2.5)) \[ \leq \| u^k - \tilde{u}^k + \varepsilon^k \| \]
(using (3.2)) \[ \leq (1 + \delta) \| u^k - \tilde{u}^k \| \]
and from (4.2), we get
\[ \lim_{k \to \infty} r(\tilde{u}^k, \rho_{\text{min}}) = 0. \] (4.3)
Let \( \tilde{u} \) be a cluster point of \{\tilde{u}^k\} and the subsequence \{w^{kj}\} converges to \( \tilde{u} \). Since \( r(u, \rho) \) is a continuous function of \( u \), it follows from (4.3) that
\[ r(\tilde{u}, \rho_{\text{min}}) = \lim_{j \to \infty} r(w^{kj}, \rho_{\text{min}}) = 0. \]
According to Lemma 2.2, \( \tilde{u} \) is a solution point of problem (2.1). Note that inequality (4.1) is true for all solution point of problem (2.1), hence we have
\[ \| u^{k+1} - \tilde{u} \| \leq \| u^k - \tilde{u} \|, \quad \forall k \geq 0. \] (4.4)
Since \{w^{kj}\} \to \tilde{u} and \( u^k - \tilde{u} \to 0 \), for any given \( \varepsilon > 0 \), there is an \( l > 0 \), such that
\[ \| w^{ki} - \tilde{u} \| < \varepsilon / 2 \quad \text{and} \quad \| u^{ki} - w^{ki} \| < \varepsilon / 2. \] (4.5)
Therefore, for any \( k \geq k_l \), it follows from (4.4) and (4.5) that
\[ \| u^k - \tilde{u} \| \leq \| u^{ki} - \tilde{u} \| \leq \| u^{ki} - w^{ki} \| + \| w^{ki} - \tilde{u} \| < \varepsilon \]
and thus the sequence \{u^k\} converges to \( \tilde{u} \). \( \square \)

The detailed algorithm is as follows.

**Algorithm 4.1.** Step 0. Let \( \rho_0 > 0, \delta := 0.95 < 1, \gamma \in [1, 2), \varepsilon > 0, k = 0 \) and \( u^0 \in K \).
Step 1. If \( \| r(u^k, \rho_k) \| \leq \varepsilon \), then stop. Otherwise, go to Step 2.
Step 2,
\[ \tilde{u}^k = P_K[u^k - \rho_k T(u^k)], \quad \varepsilon^k = \rho_k(T(\tilde{u}^k) - T(u^k)), \]
\[ r = \frac{\| \varepsilon^k \|}{\| u^k - \tilde{u}^k \|}. \]
While \((r > \delta)\)
\[
\rho_k = \frac{0.8}{r} \rho_k, \quad \tilde{u}^k = P_K[u^k - \rho_k T(u^k)],
\]
\[
\varepsilon^k = \rho_k (T(\tilde{u}^k) - T(u^k)), \quad r = \frac{\|\varepsilon^k\|}{\|u^k - \tilde{u}^k\|}.
\]
end While

Step 3. Set
\[
d(u^k, \rho_k) := u^k - \tilde{u}^k + \varepsilon^k,
\]
\[
\phi(u^k, \rho_k) := \langle u^k - \tilde{u}^k, d(u^k, \rho_k) \rangle, \quad \alpha_k = \frac{\phi(u^k, \rho_k)}{\|d(u^k, \rho_k)\|^2},
\]
\[
\tilde{u}^k = P_K[u^k - \alpha_k \rho_k T(\tilde{u}^k)],
\]
\[
\tau_k = \frac{\|u^k - \tilde{u}^k\|^2 + \alpha_k \phi(u^k, \rho_k)}{2\|u^k - \tilde{u}^k\|^2}, \quad \tau_k = \gamma \tau_k^*,
\]
\[
u^{k+1} = P_K[u^k - \tau_k(u^k - \tilde{u}^k)].
\]

Step 4. \(\rho_{k+1} = \begin{cases} \rho_k * 0.7 & \text{if } r \leq 0.5; \\ \rho_k & \text{otherwise}. \end{cases}\)

Step 5. \(k := k + 1; \) go to Step 1.

5. Relationship to some existing methods

Some methods can be viewed as special case of our method for example:

- **Method of He and Liao [8]:** If \(\tau = 1\), Algorithm 3.1 reduces to the method of He and Liao [8].
- **Method of Solodov and Svaiter [20]:** One can easily show that the method of Solodov and Svaiter [20] is a special case of Algorithm 3.1 (when \(\tau_k = \tau = 1\)).
- For given \(u^k\) and \(\rho_k > 0\), denote \(F_k(u) := (u - u^k) + \rho_k T(u)\). It is known that Solodov and Svaiter’s method [20, p. 385, Algorithm 2] consists of the following steps:

**Algorithm 5.1.** Step 1. Find a \(y^k\) which is an approximate solution of
\[
u \in K, \quad (u' - u)^T F_k(u) \geq 0, \quad \forall u' \in K,
\]
such that
\[
\{y^k - P_K[y^k - F_k(y^k)]\}^T F_k(y^k) - \frac{1}{2} \|y^k - P_K[y^k - F_k(y^k)]\|^2 \leq \frac{\delta}{2} \|y^k - u^k\|^2.
\]

Step 2. Set
\[
u^{k+1} = P_K[u^k - \rho_k T(y^k)].
\]

Note that the term \(y^k\) in Algorithm 5.1 plays the same role as the term \(\tilde{u}^k\) in our method. Now let us observe the differences between Algorithm 5.1 and our framework.

First we compare the error restrictions of the two methods. Since \(y^k \in K\), it follows from (2.2) that
\[
\{y^k - P_K[y^k - F_k(y^k)]\}^T F_k(y^k) \geq \|y^k - P_K[y^k - F_k(y^k)]\|^2.
\]
In order to satisfy Condition (5.2), one needs at least
\[ \{y^k - P_K[y^k - F_k(y^k)]\}^T F_k(y^k) \leq \delta \|y^k - u^k\|^2. \tag{5.4} \]

Notice that \( \tilde{u}^k \) generated from our method (see (3.1)) can be written as
\[ \tilde{u}^k = P_K[\tilde{u}^k - F_k(\tilde{u}^k) + \varepsilon^k] \tag{5.5} \]
and it requires at most \( \|\varepsilon^k\| \leq \delta \|\tilde{u}^k - u^k\| \).

It is worthy to discuss the relation between \( \varepsilon^k \) and \( e^k := y^k - P_K[y^k - F_k(y^k)] \) in formula (5.4). Note that \( y^k = \tilde{u}^k \). Hence according to (5.5) we have
\[ e^k = P_K[\tilde{u}^k - F_k(\tilde{u}^k) + \varepsilon^k] - P_K[y^k - F_k(y^k)]. \]

Since the projection is nonexpansive, we have \( \|e^k\| \leq \|\varepsilon^k\| \). Therefore, compared with Algorithm 5.1, the method proposed has a much relaxed error restriction.

Next we compare the step lengths employed in the correction step. In Algorithm 5.1, the step length is \( \kappa_1 = 1 \), which is different from (3.6) in our framework. In [21], the authors have proposed a range of step length which is similar to (3.6). (The \( v^k \) in [21], when \( \kappa_2 \) is taken to be zero, is the same as \( d^k \) defined by (3.4).) However, since \( v^k \) is defined after the correction step [26, p. 386] such step length cannot be applied to Algorithm 5.1.

6. Preliminary computational results

In this section, we set two examples and applied the proposed algorithm.

6.1. Numerical experiments I

In order to verify the theoretical assertions, we consider the following problems:
\[ u \geq 0, \quad T(u) \geq 0, \quad u^T T(u) = 0, \tag{6.1} \]
where
\[ T(u) = D(u) + Mu + q, \tag{6.2} \]
\( D(u) \) and \( Mu + q \) are the nonlinear part and the linear part of \( T(u) \), respectively.

We form the linear part in the test problems similarly as in Harker and Pang [6]. 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 \((-500, 500)\). In \( D(u) \), the nonlinear part of \( T(u) \), the components are chosen to be \( D_j(u) = d_j \times \arctan(u_j) \), where \( d_j \) is a random variable in \((0, 1)\). A similar type of problems was tested in [15,22].

In all tests we take \( \delta = 0.95 \) and \( \gamma = 1.98 \). All iterations start with \( u^0 = (1, \ldots, 1)^T \) and \( \rho_0 = 1 \), and stopped whenever \( \|r(u^k)\|_\infty \leq 10^{-7} \). All codes are written in Matlab and run on a P4-2.00G notebook computer. The iteration numbers and the computational time for Algorithms 4.1, 5.1 and the method in [8] with different dimensions are given in the Table 1.

6.2. Traffic equilibrium problems

In this subsection, we apply the proposed method to the traffic equilibrium problems and present corresponding numerical results.

Consider a network \([N, L]\) of nodes \( N \) and directed links \( L \), which consists of a finite sequence of connecting links with a certain orientation. Let \( a, b, \) etc., denote the links, and let \( p, q, \) etc., denote the paths. We let \( \omega \) denote an origin/destination \((O/D)\) pair of nodes of the network and \( P_\omega \) denotes the set of all paths connecting \( O/D \) pair \( \omega \).
Table 1
Numerical results for problem (6.1)

<table>
<thead>
<tr>
<th>Dimension of the problem</th>
<th>Algorithm 5.1</th>
<th>The method in [8]</th>
<th>Algorithm 4.1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. It.</td>
<td>CPU (s)</td>
<td>No. It.</td>
</tr>
<tr>
<td>$n = 100$</td>
<td>501</td>
<td>0.11</td>
<td>261</td>
</tr>
<tr>
<td>$n = 200$</td>
<td>757</td>
<td>1.42</td>
<td>402</td>
</tr>
<tr>
<td>$n = 300$</td>
<td>848</td>
<td>3.64</td>
<td>442</td>
</tr>
<tr>
<td>$n = 500$</td>
<td>913</td>
<td>9.24</td>
<td>496</td>
</tr>
<tr>
<td>$n = 700$</td>
<td>885</td>
<td>29.91</td>
<td>479</td>
</tr>
</tbody>
</table>

Fig. 1. An illustrative example of given directed network and the O/D pairs.

Note that the path-arc incidence matrix and the path-O/D pair incidence matrix, denoted by $A$ and $B$, respectively, are determined by the given network and O/D pairs. To see how to convert a traffic equilibrium problem into a variational inequality, we take into account a simple example depicted in Fig. 1.

For the given example in Fig. 1, the path-arc incidence matrix $A$ and the path-O/D pair incidence matrix $B$ have the following forms:

$$
A = \begin{pmatrix}
0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1
\end{pmatrix}, \quad B = \begin{pmatrix}
1 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 1
\end{pmatrix}.
$$

Let $u_p$ represent the traffic flow on path $p$ and $f_a$ denote the link load on link $a$, then the arc-flow vector $f$ is given by

$$
f = A^T u.
$$

Let $d_{\omega}$ denote the traffic amount between O/D pair $\omega$, which must satisfy

$$
d_{\omega} = \sum_{p \in P_\omega} u_p.
$$

Thus, the O/D pair-traffic amount vector $d$ is given by

$$
d = B^T u.
$$

Let $t(f) = \{t_a, a \in L\}$ be the vector of link travel costs, which is a function of the link flow. A user travelling on path $p$ incurs a (path) travel cost $\theta_p$. For given link travel cost vector $t$, the path travel cost vector $\theta$ is given by

$$
\theta = At(f) \quad \text{and thus} \quad \theta(u) = At(A^T u).
$$
We apply the proposed method to the example taken from [16] (Example 7.5 in [16]), which consisted of 25 nodes, 37 links and 6 O/D pairs. The network is depicted in Fig. 2. For this example, there are in total 55 paths for the six given O/D pairs and hence the dimension of the variable $u$ is 55. Therefore, the path-arc incidence matrix $A$ is a $55 \times 37$
The $O/D$ pairs and the parameters in (6.4) of the example

<table>
<thead>
<tr>
<th>$(O, D)$ pair $\omega$</th>
<th>$(1,20)$</th>
<th>$(1,25)$</th>
<th>$(2,20)$</th>
<th>$(3,25)$</th>
<th>$(1,24)$</th>
<th>$(11,25)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{\omega}$</td>
<td>1</td>
<td>6</td>
<td>10</td>
<td>5</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>$q_{\omega}$</td>
<td>1000</td>
<td>800</td>
<td>2000</td>
<td>6000</td>
<td>8000</td>
<td>7000</td>
</tr>
<tr>
<td>$</td>
<td>P_{\omega}</td>
<td>$</td>
<td>10</td>
<td>15</td>
<td>9</td>
<td>6</td>
</tr>
</tbody>
</table>

The disutility function is given by

$$\lambda_{\omega}(d) = -m_{\omega}d_{\omega} + q_{\omega}$$

(6.4)

and the coefficients $m_{\omega}$ and $q_{\omega}$ in the disutility function of different $O/D$ pairs for this example are given in Table 3.

In all test implementations we take $u^0 = (1, \ldots, 1)^T$ as starting point, $\rho_0 = 1$ and the stop criterion is

$$\frac{\| \min \{ u, T(u) \} \|_{\infty}}{\| \min \{ u^0, T(u^0) \} \|_{\infty}} \leq \varepsilon. \quad (6.5)$$

The numbers of iteration and the CPU time for different $\varepsilon$ are reported in Table 4.

Tables 1 and 4 show that the proposed method is more efficient, $\tau_k$ and $\tau_k$ play important role to reduce the iterative numbers due the fact that they are dependent on the previous points, and thus more precise.

References


