AMO - Advanced Modeling and Optimization, Volume 12, Number 2, 2010

A new alternating direction method for co-coercive variational inequality problems with linear equality and inequality constraints

Min Sun [1]

Department of Mathematics and Information Science, Zaozhuang University,

Shandong, 277160, China.

Abstract. In this paper, we propose a new alternating direction method for solving co-coercive variational inequality problems VI(f, S) with both linear equality and inequality constraints without the need to add any extra slack variables. We focus on the underlying function f does not have an explicit form and only its function values can be employed in the new method. Under the condition that the underlying function f is co-coercive, we prove the convergence of the new method. Preliminary numerical experiments are included to illustrate the efficiency of the new method.

Keywords. Variational inequality problems; Alternating direction method; Global convergence.

AMS(2000) Subject Classification: 90C25, 90C30, 90C90

1 Introduction

Let $S \subset R^n$ is a nonempty closed convex subset of R^n and f is a mapping from R^n into itself. The variational inequality problem, is to find a vector $x^* \in S$ such that

$$(x - x^*)^\top f(x^*) \geq 0, \quad \forall x \in S.$$

Variational inequality problem serves as very general mathematical models of numerous applications arising in economics, engineering, transportation, and so forth. It includes nonlinear complementarity problems(when $S = R^n_+$) and system of nonlinear equations(when $S = R^n$). Thus, it has been extensively investigated. There are substantial number of iterative methods including the projection method and its variant forms[4-13], the linearized Jacobi method[15], Newton-type method[15-16], etc. We refer the readers to the excellent monograph of Faccinei and Pang[1,2] and the references therein.

[1] Corresponding author, email address: sunmin_2008@yahoo.com.cn

AMO-Advanced Modeling and optimization. ISSN: 1841-4311

In this paper, we will focus our attention on the following co-coercive variational inequality problem, denoted by VI(f, S):

$$\text{Find} \quad x^* \in S \quad \text{such that} \quad (x - x^*)^\top f(x^*) \geq 0, \quad \forall x \in S, \tag{1}$$

where the feasible set S includes not only linear equality constraints but also linear inequality constraints, i.e.,

$$S = \{x \in R^n | Ax = b, Cx \leq d, x \in \mathcal{X}\}, \tag{2}$$

where $C \in R^{l \times n}$, $d \in R^l$ and $\mathcal{X}$ is a simple nonempty closed convex subset of R^n. The concerned function f is co-coercive on $\mathcal{X}$. That is, it has the following property: there exists a constant $\mu > 0$ such that

$$(x - x')^\top (f(x) - f(x')) \geq \mu \| f(x) - f(x') \|^2, \qquad \forall x, x' \in \mathcal{X}.$$

It is obvious that the co-coercivity (with modulus μ) implies the Lipschitz continuity (with constant $1/\mu$) and monotonicity (but not necessarily strongly monotonicity). This problem has several important applications in many fields, such as the capacitated transportation problem[3], the capacitated traffic assignment problem[17] and the packet routing in telecommunication with path and flow restrictions[18].

Note that the alternating direction methods in [4-7] can also be used to solve the variational inequality problem VI(f, S) by introducing a slack vector z to the linear inequality constraints to transform structure (2) to the following form:

$$S = \{(x, z) \in R^n \times R^l | Ax = b, Cx + z = d, x \in \mathcal{X}, z \geq 0\}.$$

However, this will increase the dimension of the variational inequality problem from n to $n + l$, leading to more computational complexity, especially when there are many inequality constraints in S.

Recently, Zhou, Chen and Han[11] proposed an extended alternating direction method for VI(f, S) to handle both the linear equality and inequality constraints directly. At the same time, it retains the good features of the modified alternating direction method[4,9,10]. However, the method requires an Armijo-type line search procedure to obtain a proper parameter β with a new projection needed for each trial point, and this can be very computationally expensive; Zhang and Han[6] gave an alternating direction method for co-coercive variational inequality problems with S only including linear equality constraints, which solves a series of small-scale easier problems to solving the original variational inequality problem, and is simple provided that the feasible set is simple.

In this paper, we extend Zhang and Han's method to solve the co-coercive VI(f, S), and a new alternating direction method without line search is proposed which inherits all nice properties which Zhang and Han's method has.

The paper is organized as follows. In the next section, some basic definitions and properties used in this paper are summarized. In Section 3, we formally present the new alternating direction method, and prove its global convergence. We report some preliminary computational results in Section 4 and some final conclusions are given in the last section.

2 Preliminaries

In this section, we summarize some basic properties and related definitions which will be used in the following discussions.

All matrices and vectors are real. For a vector $x \in R^n$ and a matrix $C \in R^{n\times n}$, we denote $\|x\| = \sqrt{x^\top x}$ as the Euclidean norm and $\|C\| = \sup\{\frac{\|Cx\|}{\|x\|} | \|x\| \neq 0\}$ as the induced matrix norm, where the transpose of x is denoted by $x^\top$. The projection of a point $x \in R^n$ onto the closed convex set K , denoted by $P_K[x]$, is defined as the unique solution of the problem

$$\min \|x - y\|, \quad \text{subject to} \quad y \in K.$$

The following well-known property of the projection operator plays an important role in the convergence analysis of our method.

Lemma 2.1. Let K be a nonempty closed convex subset of R^n. For any $x, y \in R^n$ and any $z \in K$, the following properties hold:

$$(x - P_K[x])^\top (z - P_K[x]) \leq 0. \tag{3}$$

$$\|P_K[x] - P_K[y]\|^2 \leq \|x - y\|^2 - \|P_K[x] - x + y - P_K[y]\|^2. \tag{4}$$

From (4), we can see that the projection operator $P_K[\cdot]$ is nonexpansive, that is,

$$\|P_K[x] - P_K[y]\| \leq \|x - y\|. \tag{5}$$

By appending a Lagrangian multiplier vector $y \in R^m$ to the linear equality constraint $Ax = b$ and another Lagrangian multiplier vector $z \in R^l$ to the linear inequality constraint $Cx \leq d$, the equivalent form of the variational inequality problem VI(f, S) can be expressed as follows, denoted by VI$(Q, \mathcal{W})$: Find a vector $w^* \in \mathcal{W}$, such that

$$(w - w^*)^\top Q(w^*) \geq 0 \quad \forall w \in \mathcal{W}, \tag{6}$$

where

$$w = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, Q(w) = \begin{pmatrix} f(x) - A^\top y + C^\top z \\ Ax - b \\ d - Cx \end{pmatrix}, \mathcal{W} = \mathcal{X} \times R^m \times \mathcal{Z},$$

where $\mathcal{Z} = R^l_+$.

It is well known[14] that problem VI$(Q, \mathcal{W})$ is equivalent to finding zeros of

$$e(w,\beta) := \begin{pmatrix} e_1(w,\beta) \\ e_2(w,\beta) \\ e_3(w,\beta) \end{pmatrix} = \begin{pmatrix} x - P_{\mathcal{X}}[x - \beta(f(x) - A^\top y + C^\top z)] \\ \beta(Ax - b) \\ z - P_{\mathcal{Z}}[z - \beta(d - Cx)] \end{pmatrix}. \tag{7}$$

In addition, we define

$$r(w,\beta) := \begin{pmatrix} r_1(w,\beta) \\ r_2(w,\beta) \\ r_3(w,\beta) \end{pmatrix} = \begin{pmatrix} x - P_{\mathcal{X}}[x - \beta(f(x) - A^\top (y - \beta(Ax - b)) + C^\top z)] \\ \beta(Ax - b) \\ z - P_{\mathcal{Z}}[z - \beta(d - Cx)] \end{pmatrix}. \tag{8}$$

Therefore, from the second inequality of (7), solving VI$(Q, \mathcal{W})$ is equivalent to finding a zero point of $r(w, \beta)$ for any $\beta > 0$. That is

$$w \text{ is a solution of } \mathrm{VI}(Q,\mathcal{W}) \Longleftrightarrow e(w,\beta) = 0 \Longleftrightarrow r(w,\beta) = 0, \quad \forall \beta > 0.$$

Hence, both $\|e(w, \beta)\|$ and $\|r(w, \beta)\|$ can be viewed as a residual function, which measures how much w fails to be a solution point of VI$(Q, \mathcal{W})$.

We make the following standard assumptions throughout this paper:

Assumptions. • f is a co-coercive function on $\mathcal{X}$ with modulus μ.

• The solution set of problem VI$(Q, \mathcal{W})$, denoted by $\mathcal{W}^*$, is nonempty.

• $\mathcal{X}$ is a simple closed convex set. That is, the projection onto the set is simple to carry out(e.g., $\mathcal{X}$ is the nonnegative orthant R^n_+, or more generally, a box).

3 Algorithm and global convergence

In this section, we present the algorithm for solving VI$(Q, \mathcal{W})$ and show its global convergence.

Algorithm 3.1. A New Alternating Direction Method for VI$(Q, \mathcal{W})$

Step 0. Given $\varepsilon > 0$, choose $w^0 = (x^0, y^0, z^0)^\top \in \mathcal{W}$, $\delta \in (0, 2)$, $0 < \beta_L \le \beta_0 \le \beta_U < 4\mu$ and set k:=0.

Step 1. Set

$$\tilde{x}^k = P_{\mathcal{X}}[x^k - \eta_k \alpha_k (e_1(w^k, \beta_k) - \beta_k C^\top e_3(w^k, \beta_k))], \tag{9}$$

$$\tilde{y}^k = y^k - \eta_k \alpha_k [e_2(w^k, \beta_k) - \beta_k A e_1(w^k, \beta_k)], \tag{10}$$

$$\tilde{z}^k = P_{\mathcal{Z}}[z^k - \eta_k \alpha_k (e_3(w^k, \beta_k) + \beta_k C e_1(w^k, \beta_k))], \tag{11}$$

where

$$\alpha_k = \frac{1 - \beta_k/(4\mu)}{1 + \beta_k^2 \|C^\top C\|}. \tag{12}$$

$$\eta_k = \frac{\delta(1+\beta_k^2\|C^\top C\|)(\|e_1(w^k,\beta_k)\|^2+\|e_3(w^k,\beta_k)\|^2)}{(1+\beta_k^2\|C^\top C\|)(\|e_1(w^k,\beta_k)\|^2+\|e_3(w^k,\beta_k)\|^2)+\|e_2(w^k,\beta_k)-\beta_k A e_1(w^k,\beta_k)\|^2}. \tag{13}$$

Step 2. Set

$$d(\tilde{w}^k,\beta_k) = \begin{pmatrix} (I+\beta_k^2 A^\top A)r_1(\tilde{w}^k,\beta_k) - \beta_k C^\top r_3(\tilde{w}^k,\beta_k) \\ r_2(\tilde{w}^k,\beta_k) - \beta_k A r_1(\tilde{w}^k,\beta_k) \\ \beta_k C r_1(\tilde{w}^k,\beta_k) + r_3(\tilde{w}^k,\beta_k) \end{pmatrix}, \tag{14}$$

where $\tilde{w}^k = (\tilde{x}^k,\tilde{y}^k,\tilde{z}^k)^\top$.Then compute step size t_k by

$$t_k = \frac{(1-\beta_k/(4\mu))\|r_1(\tilde{w}^k,\beta_k)\|^2 + \|r_2(\tilde{w}^k,\beta_k)\|^2 + \|r_3(\tilde{w}^k,\beta_k)\|^2}{\|d(\tilde{w}^k,\beta_k)\|^2}. \tag{15}$$

Step 3. Determine the next iterate $w^{k+1} = (x^{k+1}, y^{k+1}, z^{k+1})^\top$ via

$$w^{k+1} = P_{\mathcal{W}}[\tilde{w}^k - \delta t_k d(\tilde{w}^k,\beta_k)]. \tag{16}$$

Step 4. If $\|r(\tilde{w}^k,\beta_k)\| < \epsilon$, stop; otherwise, choose $\beta_{k+1} \in [\beta_L, \beta_U]$. Set $k := k+1$ and goto Step 1.

We begin our proof with the following lemma, which is motivated by Lemma 4.1 of [6].

Lemma 3.1. Let $w^* = (x^*, y^*, z^*)^\top \in \mathcal{W}^*$, and let the function f be a co-coercive function with modulus $\mu > 0$. Then, for any β satisfying $0 \le \beta \le 4\mu$, we have

$$(x-x^*)^\top(e_1 - \beta C^\top e_3) + (y-y^*)^\top(e_2 - \beta A e_1) + (z-z^*)^\top(e_3 + \beta C e_1) \ge (1-\frac{\beta}{4\mu})\|e_1\|^2 + \|e_3\|^2, \tag{17}$$

where $e_1 = e_1(w,\beta), e_2 = e_2(w,\beta), e_3 = e_3(w,\beta)$.

Proof. Setting $x := x - \beta[f(x) - A^\top y + C^\top z]$ and $z := x^*$ in (3), we have

$$(P_{\mathcal{X}}[x-\beta(f(x)-A^\top y + C^\top z)] - x^*)^\top (x - \beta[f(x) - A^\top y + C^\top z] \\ - P_{\mathcal{X}}[x-\beta(f(x)-A^\top y + C^\top z)]) \ge 0.$$

Then from the definition of e_1, we can get

$$(x - x^* - e_1)^\top (e_1 - \beta[f(x) - A^\top y + C^\top z]) \ge 0. \tag{18}$$

Similarly, setting $x := z - \beta(d - Cx)$ and $z := z^*$ in (3), we have

$$(P_{\mathcal{Z}}[z-\beta(d-Cx)] - z^*)^\top (z - \beta(d-Cx) - P_{\mathcal{Z}}[z-\beta(d-Cx)]) \ge 0,$$

that is

$$(z - z^* - e_3)^\top (e_3 - \beta(d - Cx)) \ge 0. \tag{19}$$

Furthermore, because $w^* \in \mathcal{W}^*$ is a solution of VI$(Q,\mathcal{W})$, we have

$$(P_{\mathcal{X}}[x-\beta(f(x)-A^\top y + C^\top z)] - x^*)^\top [f(x^*) - A^\top y^* + C^\top z^*] \ge 0,$$

i.e.,

$$\beta(x-x^*-e_1)^\top[f(x^*)-A^\top y^*+C^\top z^*]\geq 0, \tag{20}$$

and

$$Ax^*-b=0. \tag{21}$$

$$(P_{\mathcal{Z}}[z-\beta(d-Cx)]-z^*)^\top(d-Cx^*)\geq 0,$$

i.e.,

$$\beta(z-z^*-e_3)^\top(d-Cx^*)\geq 0. \tag{22}$$

By adding (18) and (20) and using (21), it follows that

$$(x-x^*-e_1)^\top[e_1-\beta(f(x)-f(x^*))+\beta A^\top(y-y^*)-\beta C^\top(z-z^*)]\geq 0,$$

which means that

$$\begin{aligned}
& (x-x^*)^\top e_1+(y-y^*)^\top(e_2-\beta Ae_1)+\beta(z-z^*)^\top(Ce_1)\\
\geq\ & \|e_1\|^2+\beta(x-x^*)^\top(f(x)-f(x^*))-\beta e_1^\top(f(x)-f(x^*))+\beta(Cx-Cx^*)^\top(z-z^*)\\
\geq\ & \|e_1\|^2+\beta\mu\|f(x)-f(x^*)\|^2-(\beta\mu\|f(x)-f(x^*)\|^2+\tfrac{\beta}{4\mu}\|e_1\|^2)\\
& +\beta(Cx-Cx^*)^\top(z-z^*)\\
=\ & (1-\tfrac{\beta}{4\mu})\|e_1\|^2+\beta(Cx-Cx^*)^\top(z-z^*),
\end{aligned} \tag{23}$$

where the second inequality follows from the inequality that for any two vectors $a,b\in R^n$,

$$a^\top b\leq \varrho\|a\|^2+\|b\|^2/(4\varrho),\qquad \forall\varrho>0,$$

and f is co-coercive with modulus μ. Similarly, by adding (19) and (22), we obtain

$$(z-z^*-e_3)^\top(e_3+\beta(Cx-Cx^*))\geq 0.$$

That is

$$(z-z^*)^\top e_3-\beta(x-x^*)^\top(C^\top e_3)\geq\|e_3\|^2-\beta(Cx-Cx^*)^\top(z-z^*). \tag{24}$$

Then adding (23) and (24), we get

$$\begin{aligned}
& (x-x^*)^\top(e_1-\beta C^\top e_3)+(y-y^*)^\top(e_2-\beta Ae_1)+(z-z^*)^\top(e_3+\beta Ce_1)\\
\geq\ & (1-\frac{\beta}{4\mu})\|e_1\|^2+\|e_3\|^2.
\end{aligned}$$

This completes the proof. Q.E.D.

Lemma 3.2. Let $w^*=(x^*,y^*,z^*)^\top\in\mathcal{W}^*$, and let the function f be a co-coercive function with modulus $\mu>0$. Then, for any β satisfying $0\leq\beta\leq 4\mu$, we have

$$\begin{aligned}
& (x-x^*)^\top(r_1+\beta^2A^\top Ar_1)+(y-y^*)^\top(r_2-\beta Ar_1)+(z-z^*)^\top(\beta Cr_1+r_3)\\
\geq\ & (1-\frac{\beta}{4\mu})\|r_1\|^2+\|r_2\|^2+\|r_3\|^2,
\end{aligned}$$

where $r_1 = r_1(w,\beta), r_2 = r_2(w,\beta), r_3 = r_3(w,\beta)$.

Proof. The proof of this lemma is much similar to that of the above lemma, so is omitted. Q.E.D.

Remark 3.1. In fact, Lemma 3.2 has proved that $-d(w,\beta)$ is a descent direction of the merit function $\frac{1}{2}||w-w^*||^2$ whenever $w \in \mathcal{W}$ is not a solution of VI$(Q,\mathcal{W})$.

In the following, we assume that the Algorithm 3.1 generates an infinite sequence $\{w^k\}$, otherwise, an approximate solution $w^k \in \mathcal{W}$ is obtained.

Lemma 3.3. Suppose that f is a co-coercive function with modulus $\mu > 0$ and $w^* = (x^*, y^*, z^*)^\top \in \mathcal{W}^*$. For given w^k and $0 \le \beta_k \le 4\mu$, let $\tilde{x}^k, \tilde{y}^k$ and $\tilde{z}^k$ be defined by (9)-(11), respectively and $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{z}^k)^\top$. Then

$$\|\tilde{w}^k - w^*\|^2 \le \|w^k - w^*\|^2 - (2-\delta)\eta_k\alpha_k^2(1+\beta_k^2\|C^\top C\|)(\|e_1\|^2 + \|e_3\|^2), \tag{25}$$

where $e_1 = e_1(w^k,\beta_k), e_2 = e_2(w^k,\beta_k), e_3 = e_3(w^k,\beta_k)$.

Proof. From (9), we have

$$\begin{aligned} &\|\tilde{x}^k - x^*\|^2 \\ \le\ & \|x^k - x^* - \eta_k\alpha_k(e_1 - \beta C^\top e_3)\|^2 - \|x^k - \eta_k\alpha_k(e_1 - \beta C^\top e_3) - \tilde{x}^k\|^2 \\ =\ & \|x^k - x^*\|^2 - 2\eta_k\alpha_k(x^k - x^*)^\top(e_1 - \beta C^\top e_3) - \|x^k - \tilde{x}^k\|^2 \\ & +2\eta_k\alpha_k(x^k - \tilde{x}^k)^\top(e_1 - \beta C^\top e_3), \end{aligned} \tag{26}$$

where the inequality follows from (4) and $x^* \in \mathcal{X}$. From(10),

$$\begin{aligned} &\|\tilde{y}^k - y^*\|^2 \\ =\ & \|y^k - \eta_k\alpha_k(e_2 - \beta_k Ae_1) - y^*\|^2 \\ =\ & \|y^k - y^*\|^2 - 2\eta_k\alpha_k(y^k - y^*)^\top(e_2 - \beta_k Ae_1) + \eta_k^2\alpha_k^2\|e_2 - \beta_k Ae_1\|^2. \end{aligned} \tag{27}$$

Similarly, from (11), we have

$$\begin{aligned} &\|\tilde{z}^k - z^*\|^2 \\ \le\ & \|z^k - z^* - \eta_k\alpha_k(e_3 + \beta_k Ce_1)\|^2 - \|z^k - \eta_k\alpha_k(e_3 + \beta_k Ce_1) - \tilde{z}^k\|^2 \\ =\ & \|z^k - z^*\|^2 - 2\eta_k\alpha_k(z^k - z^*)^\top(e_3 + \beta_k Ce_1) - \|z^k - \tilde{z}^k\|^2 \\ & +2\eta_k\alpha_k(z^k - \tilde{z}^k)^\top(e_3 + \beta_k Ce_1), \end{aligned} \tag{28}$$

where the inequality also follows from (4) and $z^* \in \mathcal{Z}$. Adding (26)-(28), it follows that

$$\begin{aligned} &\|\tilde{w}^k - w^*\|^2 \\ \le\ & \|w^k - w^*\|^2 - \|x^k - \tilde{x}^k\|^2 - \|z^k - \tilde{z}^k\|^2 \\ & +2\eta_k\alpha_k(x^k - \tilde{x}^k)^\top(e_1 - \beta_k C^\top e_3) + 2\eta_k\alpha_k(z^k - \tilde{z}^k)^\top(e_3 + \beta_k Ce_1) + \eta_k^2\alpha_k^2\|e_2 - \beta_k Ae_1\|^2 \\ & -2\eta_k\alpha_k[(1 - \frac{\beta_k}{4\mu})\|e_1\|^2 + \|e_3\|^2] \end{aligned}$$

$$\begin{aligned}
&\leq \|w^k - w^*\|^2 + \eta_k^2\alpha_k^2(\|e_1 - \beta_k C^\top e_3\|^2 + \|e_3 + \beta_k C e_1\|^2) + \eta_k^2\alpha_k^2\|e_2 - \beta_k A e_1\|^2 \\
&\quad -2\eta_k\alpha_k^2(1 + \beta_k^2\|C^\top C\|)(\|e_1\|^2 + \|e_3\|^2) \\
&= \|w^k - w^*\|^2 + \eta_k^2\alpha_k^2[(1 + \beta_k^2\|C^\top C\|)(\|e_1\|^2 + \|e_3\|^2) + \|e_2 - \beta_k A e_1\|^2] \\
&\quad -2\eta_k\alpha_k^2(1 + \beta_k^2\|C^\top C\|)(\|e_1\|^2 + \|e_3\|^2),
\end{aligned}$$

where the first inequality follows from Lemma 3.1, and the second inequality follows from Cauchy-Schwartz inequality and the definition of α_k. The assertion follows from the above inequality and the definition of η_k immediately. This completes the proof. Q.E.D.

Similar to Lemma 4.3 of [6], we have:

Lemma 3.4. Suppose that f is a co-coercive function with modulus $\mu > 0$ and $w^* = (x^*, y^*, z^*)^\top \in \mathcal{W}^*$. Let $w^k = (x^k, y^k, z^k)^\top$ be generated by Algorithm 3.1. Then

$$\begin{aligned}
\|w^{k+1} - w^*\|^2 \leq \|w^k - w^*\|^2 &- (2-\delta)\eta_k\alpha_k^2(1 + \beta^2\|C^\top C\|)(\|e_1(w^k, \beta_k)\|^2 + \|e_3(w^k, \beta_k)\|^2) \\
&- \delta(2-\delta)t_k[(1 - \frac{\beta_k}{4\mu}))\|r_1(\tilde{w}^k, \beta_k)\|^2 + \|r_2(\tilde{w}^k, \beta_k)\|^2 + \|r_3(\tilde{w}^k, \beta_k)\|^2].
\end{aligned}$$

Proof. From the definition of w^{k+1} and $w^* \in \mathcal{W}$, it follows that

$$\begin{aligned}
&\|w^{k+1} - w^*\|^2 \\
\leq\ & \|\tilde{w}^k - \delta t_k d(\tilde{w}^k, \beta_k) - w^*\|^2 \\
=\ & \|\tilde{w}^k - w^*\|^2 + \delta^2 t_k^2\|d(\tilde{w}, \beta_k)\|^2 - 2\delta t_k(\tilde{w}^k - w^*)^\top d(\tilde{w}, \beta_k) \\
\leq\ & \|\tilde{w}^k - w^*\|^2 + \delta^2 t_k^2\|d(\tilde{w}, \beta_k)\|^2 - 2\delta t_k[(1 - \frac{\beta_k}{4\mu})\|r_1(\tilde{w}^k, \beta_k)\|^2 + \|r_2(\tilde{w}^k, \beta_k)\|^2 + \|r_3(\tilde{w}^k, \beta_k)\|^2] \\
\leq\ & \|w^k - w^*\|^2 - (2-\delta)\eta_k\alpha_k^2(1 + \beta_k^2\|C^\top C\|)(\|e_1(w^k, \beta_k)\|^2 + \|e_3(w^k, \beta_k)\|^2) \\
& -\delta(2-\delta)t_k[(1 - \frac{\beta_k}{4\mu})\|r_1(\tilde{w}^k, \beta_k)\|^2 + \|r_2(\tilde{w}^k, \beta_k)\|^2 + \|r_3(\tilde{w}^k, \beta_k)\|^2],
\end{aligned}$$

where the first inequality follows from the nonexpansivity of the projection operator (5), and the second inequality follows from Lemma 3.2, and the last inequality follows from (25) and the definition of the step size t_k. This completes the proof. Q.E.D.

The following lemma shows that the step size t_k is bounded away from zero.

Lemma 3.5. For all $k \geq 0$, we have

$$t_k \geq \varsigma, \tag{29}$$

where $\varsigma > 0$ is a constant.

Proof. It follows from the definition of $r(\tilde{w}^k, \beta_k)$ and $\beta_k \leq \beta_U$ that

$$\|(I + \beta_k^2 A^\top A)r_1(\tilde{w}^k, \beta_k) - \beta_k C^\top r_3(\tilde{w}^k, \beta_k)\| \leq (1 + \beta_U^2\|A^\top A\| + \beta_U\|C^\top\|)\|r(\tilde{w}^k, \beta_k)\|,$$

$$\|r_2(\tilde{w}^k, \beta_k) - \beta_k A r_1(\tilde{w}^k, \beta_k)\| \leq (1 + \beta_U\|A\|)\|r(\tilde{w}^k, \beta_k)\|,$$

$$\|\beta_k Cr_1(\tilde{w}^k,\beta_k)+r_3(\tilde{w}^k,\beta_k)\| \le (1+\beta_U\|C\|)\|r(\tilde{w}^k,\beta_k)\|.$$

The above three inequalities and the definition of $d(\tilde{w}^k,\beta_k)$ imply

$$\|d(w^k,\mu_k)\| \le c_1\|r(\tilde{w}^k,\beta_k)\|, \quad \forall k\ge 0, \tag{30}$$

where

$$c_1 = 3+\beta_U(\|A\|+\|C\|+\|C^\top\|)+\beta_U^2\|A^\top A\|.$$

On the other hand, from $0<\beta_L\le\beta_k\le\beta_U<4\mu$, we have

$$(1-\frac{\beta_k}{4\mu})\|r_1(\tilde{w}^k,\beta_k)\|^2+\|r_2(\tilde{w}^k,\beta_k)\|^2+\|r_3(\tilde{w}^k,\beta_k)\|^2 \ge c_2\|r(\tilde{w}^k,\beta_k)\|^2, \tag{31}$$

where $c_2 = 1-\beta_U/(4\mu)>0$. Therefore, from (30) (31) and the definition of t_k, we have

$$t_k\ge\varsigma,$$

where $\varsigma = c_2/c_1^2$. This completes the proof. Q.E.D.

We are now ready to prove the global convergence of the sequence $\{w^k\}$ generated by our algorithm.

Theorem 3.1. Suppose that the function f is co-coercive with modulus μ and $\beta_L\le\beta_k\le\beta_U<4\mu$ for all $k\ge0$. The sequence $\{w^k\}$ generated by Algorithm 3.1 converges to a solution of VI(Q,W) globally.

Proof. Since $\delta\in(0,2)$ and $t_k>0$, $\eta_k>0$, it follows from Lemma 3.4 that

$$\|w^{k+1}-w^*\|^2\le\|w^k-w^*\|^2\le\cdots\le\|w^0-w^*\|^2<+\infty,$$

which means that the sequence $\{w^k\}$ is bounded. Thus, it has at least one cluster point, denoted as $w^\infty=(x^\infty,y^\infty,z^\infty)^\top$. Again from Lemma 3.4, we have

$$\begin{aligned}
&(2-\delta)\eta_k\alpha_k^2(1+\beta_k^2\|C^\top C\|)(\|e_1(w^k,\beta_k)\|^2+\|e_3(w^k,\beta_k)\|^2)+\delta(2-\delta)t_k\\
&[(1-\frac{\beta_k}{4\mu})\|r_1(\tilde{w}^k,\beta_k)\|^2+\|r_2(\tilde{w}^k,\beta_k)\|^2+\|r_3(\tilde{w}^k,\beta_k)\|^2]\\
\le\ &\|w^k-w^*\|^2-\|w^{k+1}-w^*\|^2.
\end{aligned}$$

Summarizing both sides of the above inequality, we get

$$\begin{aligned}
&\sum_{k=0}^{\infty}\{(2-\delta)\eta_k\alpha_k^2(1+\beta_k^2\|C^\top C\|)(\|e_1(w^k,\beta_k)\|^2+\|e_3(w^k,\beta_k)\|^2)+\delta(2-\delta)t_k\\
&[(1-\frac{\beta_k}{4\mu})\|r_1(\tilde{w}^k,\beta_k)\|^2+\|r_2(\tilde{w}^k,\beta_k)\|^2+\|r_3(\tilde{w}^k,\beta_k)\|^2]\}\\
\le\ &\sum_{k=0}^{\infty}\{\|w^k-w^*\|^2-\|w^{k+1}-w^*\|^2\}\\
\le\ &\|w^0-w^*\|^2\\
<\ &+\infty,
\end{aligned}$$

which together with (29) and $\beta_k \le \beta_U < 4\mu$ imply that

$$\lim_{k\to\infty} \eta_k\alpha_k^2(\|e_1(w^k,\beta_k)\|^2 + \|e_3(w^k,\beta_k)\|^2) = 0. \tag{32}$$

$$\lim_{k\to\infty} \|r_1(\tilde{w}^k,\beta_k)\| = \lim_{k\to\infty} \|r_2(\tilde{w}^k,\beta_k)\| = \lim_{k\to\infty} \|r_3(\tilde{w}^k,\beta_k)\| = 0. \tag{33}$$

On the other hand, from the boundedness of $\{w^k,\}$ and $\{\beta_k\}$, it is true that the dominator of η_k is bounded. That is, there is a constant $M > 0$, such that

$$(1+\beta_k^2\|C^\top C\|)(\|e_1(w^k,\beta_k)\|^2 + \|e_3(w^k,\beta_k)\|^2) + \|e_2(w^k,\beta_k) - \beta_k A e_1(w^k,\beta_k)\|^2 < M, \quad \forall k \ge 0.$$

We therefore have that

$$\begin{aligned}\eta_k\alpha_k^2(\|e_1(w^k,\beta_k)\|^2 + \|e_3(w^k,\beta_k)\|^2) &\ge \frac{\delta\alpha_k^2(\|e_1(w^k,\beta_k)\|^2 + \|e_3(w^k,\beta_k)\|^2)^2}{M}\\ &\ge \frac{\delta[1-\beta_U/(4\mu)]^2(\|e_1(w^k,\beta_k)\|^2 + \|e_3(w^k,\beta_k)\|^2)^2}{(1+\beta_U^2\|C^\top C\|)^2 M},\end{aligned}$$

which together with (32) and $\beta_U < 4\mu$ imply that

$$\lim_{k\to\infty} \|e_1(w^k,\beta_k)\| = \lim_{k\to\infty} \|e_3(w^k,\beta_k)\| = 0. \tag{34}$$

From (9), we have

$$\begin{aligned}&\|x^k - \tilde{x}^k\|\\ =\ &\|x^k - P_{\mathcal{X}}[x^k - \eta_k\alpha_k(e_1(w^k,\beta_k) - \beta_k C^\top e_3(w^k,\beta_k))]\|\\ \le\ &\eta_k\alpha_k\|e_1(w^k,\beta_k) - \beta_k C^\top e_3(w^k,\beta_k)\|\\ \le\ &\delta(1-\frac{\beta_L}{4\mu})(\|e_1(w^k,\beta_k)\| + \beta_U\|C^\top\|\|e_3(w^k,\beta_k)\|),\end{aligned}$$

where the first inequality follows from the nonexpansivity of the projection operator and $x^k \in \mathcal{X}$. From the above inequality and (34), we have

$$\lim_{k\to\infty} \|x^k - \tilde{x}^k\| = 0. \tag{35}$$

Similarly, we have

$$\begin{aligned}&\|y^k - \tilde{y}^k\|\\ =\ &\|y^k - y^k + \eta_k\alpha_k[e_2(w^k,\beta_k) - \beta_k A e_1(w^k,\beta_k)]\|\\ =\ &\eta_k\alpha_k\|\beta_k(Ax^k - b) - \beta_k A e_1(w^k,\beta_k)\|\\ \le\ &\delta(1-\frac{\beta_L}{4\mu})(\beta_k\|A\|\|x^k - \tilde{x}^k\| + \beta_k\|A\tilde{x}^k - b\| + \beta_U\|A\|\|e_1(w^k,\beta_k)\|)\\ \le\ &\delta(1-\frac{\beta_L}{4\mu})(\beta_U\|A\|\|x^k - \tilde{x}^k\| + \|r_2(\tilde{w}^k,\beta_k)\| + \beta_U\|A\|\|e_1(w^k,\beta_k)\|).\end{aligned}$$

Therefore, from the above inequality and (33)-(35), we have

$$\lim_{k\to\infty} \|y^k - \tilde{y}^k\| = 0. \tag{36}$$

Similarly, from (5) (11) and $z^k \in \mathcal{Z}$, we obtain

$$\begin{aligned} & \|z^k - \bar{z}^k\| \\ = \;& \|z^k - P_{\mathcal{Z}}[z^k - \eta_k\alpha_k(e_3(w^k,\beta_k) + \beta_k C e_1(w^k,\beta_k))]\| \\ \le \;& \delta(1 - \frac{\beta_L}{4\mu})(\beta_U\|C\|\|e_1(w^k,\beta_k)\| + \|e_3(w^k,\beta_k)\|). \end{aligned}$$

From the above inequality and (34), we get

$$\lim_{k\to\infty} \|z^k - \tilde{z}^k\| = 0. \tag{37}$$

From (35)-(37) and the boundedness of the sequence $\{w^k\}$, we can get that w^∞ is also a cluster point of $\{\tilde{w}^k\}$ .Therefore, there exists a subsequence $\{\tilde{w}^{k_j}\} = (\tilde{x}^{k_j}, \tilde{y}^{k_j}, \tilde{z}^{k_j})^\top$ converging to it. Without loss of generality we can assume that

$$\lim_{k\to\infty} \beta_{k_j} = \beta_*.$$

Taking limit along such a sequence in (33), we have

$$\|r_1(w^\infty,\beta_*)\| = \|r_2(w^\infty,\beta_*)\| = \|r_3(w^\infty,\beta_*)\| = 0.$$

So w^∞ is a solution of VI$(Q,\mathcal{W})$.

In the following we prove that the sequence $\{w^k\}$ has exactly one cluster point. Assume that $\hat{w}$ is another cluster point of $\{w^k\}$. Then we have

$$\delta := \|w^\infty - \hat{w}\| > 0.$$

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

$$\|w^{k_0} - w^\infty\| \le \frac{\delta}{2}.$$

On the other hand, since $\{\|w^k - w^\infty\|\}$ is monotonically non-increasing, we have $\|w^k - w^\infty\| \le \|w^{k_0} - w^\infty\|$ for all $k \ge k_0$, and it follows that

$$\|w^k - \hat{w}\| \ge \|w^\infty - \hat{w}\| - \|w^k - w^\infty\| \ge \frac{\delta}{2}, \forall k \ge k_0,$$

which contradicts the fact that $\hat{w}$ is a cluster point of $\{w^k\}$. This contradiction assures that the sequence $\{w^k\}$ converges to its unique cluster point w^∞, which is a solution of VI$(Q,\mathcal{W})$. This completes the proof. Q.E.D.

4 Preliminary Computational Results

In this section, we illustrate the necessity and efficiency of our method. To this end, we also code the algorithm proposed by Zhang and Han[6] and the algorithm proposed by Zhou, Chen and Han[11].

Example 4.1. The example used here is a modification of the test problem in paper[6], which constraint set S and the mapping f are taken, respectively, as

$$S = \{x \in R^5_+ | \sum_{i=1}^{5} x_i \leq 10\},$$

and

$$f(x) = Mx + \rho C(x) + q,$$

where M is an $R^{5\times 5}$ asymmetric positive matrix and $C_i(x) = \arctan(x_i - 2), i = 1, 2, \cdots, 5$. The parameter ρ is used to vary the degree of asymmetry and nonlinearity. The data of example are illustrate as follows:

$$M = \begin{pmatrix} 0.726 & -0.949 & 0.266 & -1.193 & -0.504 \\ 1.645 & 0.678 & 0.333 & -0.217 & -1.443 \\ -1.016 & -0.225 & 0.769 & 0.943 & 1.007 \\ 1.063 & 0.587 & -1.144 & 0.550 & -0.548 \\ -0.256 & 1.453 & -1.073 & 0.509 & 1.026 \end{pmatrix}$$

and

$$q = (5.308,\ 0.008\ , -0.938,\ 1.024,\ -1.312)^\top.$$

Thus, in this experiment, $A = 0, b = 0, C = (1, 1, 1, 1, 1), d = 10$. For Algorithm 3.1, we take $\beta_k = 0.06$, $\delta = 1.35$. For the method in [6], denoted by Zhang and Han's method, we add slack variables z to convert the inequality constraints to equality constraints and take $\beta_k = 0.05$, $\delta = 1.35$ when $\rho = 10$ and $\beta_k = 0.06$, $\delta = 1.35$ when $\rho = 20$. For the method in [11], denoted by Zhou, Chen and Han's method, we take $\tau = 1.98, \mu = 0.4, v = 0.6, \delta = 0.6$. In this experiment, we take the stopping criterion $\varepsilon = 10^{-6}$, $z^0 = 0$ as the initial point. All programs are coded in Matlab 7.1. 'IN' denotes the number of iterations and 'CPU' denotes the CPU time in seconds.

The results in the Table 1 and Table 2 indicate that the performance of the Algorithm 3.1 is better than that of Zhang and Han's method and Zhou, Chen and Han's method in terms of number of iteration.

Example 4.2. To show the advantage of the new alternating direction method for large scale problems, we implement Algorithm 3.1 to a set of spatial price equilibrium problem, which is a modification

Table 1: Numerical results for $\rho = 10$.

Starting point	Method	IN	CPU
	Zhang and Han's method	35	0.01
	Zhou, Chen and Han's method	31	0.01
(0 2.5 2.5 2.5 2.5)	Algorithm 3.1	9	0.01
	Zhang and Han's method	38	0.01
	Zhou, Chen and Han's method	49	0.01
(25 0 0 0 0)	Algorithm 3.1	17	0.01
	Zhang and Han's method	46	0.01
	Zhou, Chen and Han's method	21	0.01
(10 0 0 0 0)	Algorithm 3.1	12	0.01
	Zhang and Han's method	35	0.01
	Zhou, Chen and Han's method	54	0.01
(10 0 10 0 10)	Algorithm 3.1	9	0.01

of the problem in [5] by adding some inequality constraints, as follows:

$$\min \sum_{i=1}^{m}\sum_{j=1}^{n}(c_{ij}x_{ij} + \frac{1}{2}h_{ij}x_{ij}^2).$$

$$\text{s.t.} \sum_{j=1}^{n} x_{ij} = s_i, i = 1, 2, \cdots, m,$$

$$\sum_{i=1}^{m} x_{ij} = d_j, j = 1, 2, \cdots, n,$$

$$x_{i1} \leq 0.1 s_i, i = 1, 2, \cdots, m,$$

$$x_{ij} \geq 0, i = 1, 2, \cdots, m, j = 1, 2, \cdots, n,$$

where s_i is the supply amount on the ith supply market, $i = 1, \cdots, m$ and d_j the demand amount on the jth demand market, $j = 1, \cdots, n$. $c_{ij} \in (1, 100)$, $h_{ij} \in (0.005, 0.01)$, s_j and d_j are generated randomly in $(0, 100)$ for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$. We take $\beta_k = 0.4, \delta = 1.65$ for Algorithm 3.1, and $\beta_k = 0.2, \delta = 1.6$ for Zhang and Han's method. The calculations were started with $w^0 = 0$ and stopped when

$$\|r_1(\tilde{w}^k, \beta_k)\| + \|r_2(\tilde{w}^k, \beta_k\| + \|r_3(\tilde{w}^k, \beta_k)\| \leq \varepsilon,$$

for Algorithm 3.1. For Zhang and Han's method, the stop criterion is

$$\|r_1(u^k, \beta_k)\| + \|r_2(u^k, \beta_k)\| \leq \varepsilon,$$

Table 2: Numerical results for $\rho = 20$.

Starting point	Method	IN	CPU
	Zhang and Han's method	48	0.01
	Zhou, Chen and Han's method	45	0.01
(0 2.5 2.5 2.5 2.5)	Algorithm 3.1	6	0.01
	Zhang and Han's method	51	0.01
	Zhou, Chen and Han's method	98	0.01
(25 0 0 0 0)	Algorithm 3.1	10	0.01
	Zhang and Han's method	51	0.01
	Zhou, Chen and Han's method	20	0.01
(10 0 0 0 0)	Algorithm 3.1	7	0.01
	Zhang and Han's method	50	0.01
	Zhou, Chen and Han's method	104	0.02
(10 0 10 0 10)	Algorithm 3.1	7	0.01

where $r_1(u^k, \beta_k)$ and $r_2(u^k, \beta_k)$ is defined in [6]. The computational results are given in Table 3 for some m and n. The numerical results given in Table 3 show that Algorithm 3.1 outperforms Zhang and Han's method, and it is attractive in practice.

5 Conclusions

In this paper, we observe a new descent direction at each iteration, and present a new alternating direction method for co-coercive VI(f, S). Total computational cost of the method is very tiny provided that the projection is easy to implement. Thus, the new method is applicable in practice. Under some mild conditions, we proved the global convergence of the method. Some preliminary computational results illustrated the efficiency of the algorithm.

References

[1] F.Facchinei, J.S.Pang, Finite-demensional variational inequalities and complementarity problems. Vol. I, Springer Series in Operations Research, Springer, New York, USA, 2003.

[2] F.Facchinei, J.S.Pang, Finite-demensional variational inequalities and complementarity problems. Vol. II, Springer Series in Operations Research, Springer, New York, USA, 2003.

Table 3: Numerical results for different scale and precisions.

m	n	Algorithm		$\varepsilon = 0.1$	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$
5	10	Zhang and Han's method	IN	312	745	1111	2332
			CPU	0.05	0.10	0.16	0.31
		Algorithm 3.1	IN	249	306	756	843
			CPU	0.05	0.07	0.15	0.17
10	15	Zhang and Han's method	IN	333	959	1572	2564
			CPU	0.12	0.34	0.52	0.87
		Algorithm 3.1	IN	297	637	1066	1881
			CPU	0.15	0.28	0.50	0.88
20	25	Zhang and Han's method	IN	355	945	2350	3558
			CPU	1.30	3.50	9.20	13.20
		Algorithm 3.1	IN	342	857	1589	3016
			CPU	0.64	1.60	3.01	5.78
30	40	Zhang and Han's method	IN	438	1444	1788	3194
			CPU	9.83	32.88	40.28	68.64
		Algorithm 3.1	IN	371	1125	1319	3368
			CPU	3.10	8.75	11.59	28.61

[3] S.Dafernos, Traffic equilibrium and variational inequalities. *Transportation Science*, **14** (1980), 42–54.

[4] Han D.R., A modified alternating direction method for variational inequality problems. *Applied Mathematics and Optimization*, **45** (2002), 63–74.

[5] Han D.R., A proximal decomposition algorithm for variational inequality problems. *Journal of Computation and Applied Mathematics*, **161** (2003), 231–244.

[6] Zhang W.X., Han D.R., A new alternating direction method for co-coercive variational inequality problems. *Computers and Mathematics with Applications.* **57(7)**, (2009) ,1168-1178

[7] Ye C.H.,Yuan X.M., A descent method for structured monotone variational inequalities. *Optimization Methods and Software*, **22** (2007), 329–338.

[8] He B.S., Hai Y., Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities. *Operations Research Letters*, **23** (1998), 151–161.

[9] Han, D.R., Hong, K.Lo, Wang, Z.W., A simple self-adaptive alternating direction method for linear variational inequality problems. *Computers and Mathematics with Applications*, **53** (2007), 1595–1604.

[10] Han, D.R., Hong, K.Lo, New alternating direction method for a class of nonlinear variational inequality problems. *Journal of Optimization Theory and Applications*, **112(3)** (2002), 549–560.

[11] Zhou Z., Chen, A., Han D.R., An extended alternating direction method for variational inequality problems with linear equality and inequality constraints. *Appllied Mathematics and Computation*, **184** (2007), 769–782.

[12] E.M.Gafni, D.P.Bertsekas, Two-matric projection methods for constrained optimization. *SIAM Journal on Control and Optimization*, **22** (1984), 936–964.

[13] Jiang Z.K., Abdellah B., A projection-based prediction-correction method for structured monotone variational inequalities. *Applied Mathematics and Computation*, **202** (2008), 747–759.

[14] Eaves B.C., On the basic theorem of complementarity. *Mathematical Programming*, **1** (1971), 68–75.

[15] Harker P.T., Pang J.S, Finite-dimensional variational inequality and complementarity problems: a survey of theory, algorithms and applications. *Mathematical Programming*, **48** (1990), 161–220.

[16] Marcotte P., Dussault J.P., A note on a globally convergent Newton method for solving variational inequalities. *Operations Research Letters*, **6** (1987), 35–42.

[17] Larsson T., Patriksson M., Equilibrium characterizations of soltions to side constrained asymmetric traffic assignment models. *Le Matematiche*, **49** (1994), 249–280.

[18] Leblanc L., Chifflet J., Mahey P., Packet routing in telecommunication networks with path and flow restrictions. *Networks*, **11(2)** (1999), 188–197.