AMO - Advanced Modeling and Optimization, Volume 14, Number 1, 2012

Self-adaptive prediction-correction method for constrained linear variational inequality problems

Jing Liu [1]
School of Mathematics and Statistics, Zhejiang University of Finance and Economics,
Hangzhou, 310018, China.

Abstract. This article presents a new self-adaptive prediction-correction method for solving a class of linear variational inequalities. At each iteration, it only needs to perform some orthogonal projections onto simple convex sets and some matrix-vector multiplications. The method makes use of a new descent direction to produce the new iterate and can be also viewed as a projection-based descent method. Convergence of the proposed method is proved under certain conditions. Numerical experiments are carried out to show the efficiency and robustness of our new method.

Keywords. Linear variational inequalities, Prediction-correction method, Projection-based method

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

1 Introduction

Given an $n \times n$ matrix H and a vector $c \in R^n$, the following constrained linear variational inequality problem arises frequently in traffic equilibrium and network economics problems[1-3], which is to find a vector $x^* \in S$ such that

$$(\mathrm{LVI}(H, c, S)) \quad (x - x^*)^\top (Hx^* + c) \geq 0, \quad \forall x \in S, \tag{1}$$

where

$$S = S_1 = \{x \in R^n | Ax = b, x \in X\},$$

or

$$S = S_2 = \{x \in R^n | Ax \leq b, x \in X\},$$

[1] This work was supported by the Foundation of Shandong Provincial Education Department (No.J08LI66)
***AMO-Advanced Modeling and optimization. ISSN: 1841-4311**

$A \in R^{m\times n}$, $b \in R^m$, and X is a simple nonempty closed convex subset of R^n. For example, X is the nonnegative orthant $\{x \in R^n | x \geq 0\}$, or a box $\{x \in R^n | l \leq x \leq u\}$, or a ball $\{x \in R^n | \|x\| \leq r\}$.

By introducing Lagrangian multiplier $y \in Y = R^m$ and $y \in Y = R^m_+$ to the linear constraints $Ax = b$ and $Ax \leq b$, respectively, we obtain an equivalent form of LVI(H, c, S), denoted by LVI(M, q, Ω): Find $u^* \in \Omega$, such that

$$(u - u^*)^\top (Mu^* + q) \geq 0, \forall u \in \Omega, \tag{2}$$

where

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, M = \begin{pmatrix} H & -A^\top \\ A & 0 \end{pmatrix}, q = \begin{pmatrix} c \\ -b \end{pmatrix}, \Omega = X \times Y.$$

Since the projection onto X is trivial, problem LVI(M, q, Ω) can be solved by some projection-based methods[4-12]. The alternating direction method is an attractive projection-based method for the above problem LVI(M, q, Ω); see [4,5,9-12], for example. Recently, paper [9] proposed the following alternating direction method for solving problem LVI(M, q, Ω) with $S = S_1$:

Given $(x^k, y^k) \in X \times R^m$, compute $\tilde{x}^k$ via

$$\tilde{x}^k = P_X[x^k - \frac{1}{\mu_k}(Hx^k + A^\top(Ax^k - b) - A^\top y^k + c)]. \tag{3}$$

Then, find the next iterative point by

$$x^{k+1} = P_X[x^k - \tau\rho_k B_k(x^k - \tilde{x}^k)],$$

$$y^{k+1} = y^k - \tau\rho_k(A\tilde{x}^k - b),$$

where

$$\rho_k = \frac{(x^k - \tilde{x}^k)^\top B_k(x^k - \tilde{x}^k) + \|A\tilde{x}^k - b\|^2}{\|B_k(x^k - \tilde{x}^k)\|^2 + \|A\tilde{x}^k - b\|^2},$$

and $0 < \tau < 2, \mu_k > \|H + A^\top A\|, B_k = \mu_k I - (H + A^\top A)$. The method is simple in the sense that, at each iteration, it only to perform two projections onto simple set X and some matrix-vector multiplications. Moreover, it adaptively select the parameters μ_k so as to improve its efficiency(see Step 4 of Algorithm 3.1 in [9]).

In what follows, we assume that the constraint set Y in problem LVI(M, q, Ω) is a proper subset of R^m, and focus on this special class of linear variational inequality problems. Note that this special class of variational inequalities can be expressed as follows, denoted by problem LVI(H, c, W)[6]: Find a vector $w^* \in W$, such that

$$(w - w^*)^\top Q(w^*) \geq 0 \quad \forall w \in W, \tag{4}$$

where

$$w = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, Q(w) = \begin{pmatrix} Hx + c - A^\top y \\ z \\ Ax - z - b \end{pmatrix}, W = X \times Y \times R^m.$$

The purpose of this paper is to present a self-adaptive prediction-correction method for solving problem LVI(H, c, W), which inherits all nice properties which the method in [9] has.

The paper is organized as follows. In Section 2, we summarize some basic definitions and properties used in this paper, then we formally propose the new self-adaptive prediction-correction method, and the global convergence of the method is proved. In Section 3, we report some preliminary computational results of the proposed method. Section 4 gives some concluding remarks.

2 Algorithm and Convergence

We first give some basic properties and related definitions used in the subsequent sections. 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 $\|C\| = \max\{\frac{\|Cx\|}{\|x\|} | \|x\| \neq 0\}$ as the matrix 2-norm. Let K be a nonempty closed convex set in R^n, and we use $P_K[\cdot]$ to denote the orthogonal projection mapping from R^n onto K. That is,

$$P_K[x] = \text{argmin}\{\|x - y\|, y \in K\}.$$

The following well-known properties of the projection operator will be used below.

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:

$$(x - P_K[x])^\top (z - P_K[x]) \leq 0. \tag{5}$$

$$\|P_K[x] - P_K[y]\|^2 \leq \|x - y\|^2 - \|P_K[x] - x + y - P_K[y]\|^2. \tag{6}$$

From (6), we can see that the projection operator $P_K[\cdot]$ is nonexpansive, that is,

$$\|P_K[x] - P_K[y]\| \leq \|x - y\|. \tag{7}$$

It is well known[15] that problem LVI(H, c, W) is equivalent to the projection equation

$$w = P_W[w - \beta Q(w)],$$

where β is an arbitrary positive constant. Let

$$e(w,\beta) = \begin{pmatrix} e_1(w,\beta) \\ e_2(w,\beta) \\ e_3(w,\beta) \end{pmatrix} = \begin{pmatrix} x - P_X[x - \beta(Hx + c - A^\top y)] \\ y - P_Y[y - \beta z] \\ \beta(Ax - z - b) \end{pmatrix}$$

denote the residual function of the projection equation, thus problem LVI(H, c, W) is equivalent to finding a zero point of $e(w, \beta)$. In the literatures[9-12], $\|e(w, \beta)\|$ was viewed as a measure function, which measures how much w fails to be a solution point of problem LVI(H, c, W).

For any given $w \in R^{n+2m}$, the magnitude $\|e(w, \beta)\|$ is dependent on β. The following lemma plays an important role in our later convergence analysis.

Lemma 2.2. For any $w \in R^{n+2m}$ and $0 < \beta_1 < \beta_2$, we have

$$\|e(w, \beta_1)\| \leq \|e(w, \beta_2)\|, \tag{8}$$

and

$$\frac{\|e(w, \beta_1)\|}{\beta_1} \geq \frac{\|e(w, \beta_2)\|}{\beta_2}. \tag{9}$$

Proof. See Lemma 1 of [13] and (2.6) of [14]. Q.E.D.

Throughout this paper, we make the following standard assumptions.

Assumptions:

A. X and Y are simple closed convex sets. That is, the projection onto the set is simple to carry out.

B. H is a positive semi-definite matrix.

C. The solution set of problem LVI(H, c, W), denoted by W^*, is nonempty.

We are now in the position to describe our method formally.

Algorithm 2.1 A Self-adaptive Prediction-correction Method.

Step 0. Choose an arbitrary point $w^0 = (x^0, y^0, z^0) \in W$, and set a small number $\varepsilon > 0$ for the solution accuracy, $\mu_0 > 0, \tau \in (0, 2), \sigma \in (0, 1)$, and a nonnegative sequence $\{\gamma_i\}$ satisfying $\sum_{i=0}^{\infty} \gamma_i < \infty$, set k:=0.

Step 1. If

$$\|e(w^k, 1)\| \leq \varepsilon,$$

then stop; else, goto step 2.

Step 2.(prediction step) Set

$$\bar{x}^k = P_X[x^k - \frac{1}{\mu_k}(Hx^k + c - A^\top y^k)], \tag{10}$$

$$\bar{y}^k = P_Y[y^k - \frac{1}{\mu_k} z^k], \tag{11}$$

$$\bar{z}^k = z^k - \frac{1}{\mu_k}(Ax^k - z^k - b). \tag{12}$$

Step 3.(correction step) Set

$$d(w^k, \mu_k) = \begin{pmatrix} B_k(x^k - \bar{x}^k) + A^\top(Ax^k - z^k - b) \\ \mu_k(y^k - \bar{y}^k) + (A\bar{x}^k - z^k - b) \\ (y^k - \bar{y}^k) - (Ax^k - z^k - b) \end{pmatrix}. \tag{13}$$

where $B_k = \mu_k I - H$. Then compute step size α_k by

$$\alpha_k = \tau \frac{\|x^k - \bar{x}^k\|_{B_k}^2 + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2}{\|d(w^k, \mu_k)\|^2}. \tag{14}$$

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

$$w^{k+1} = P_W[w^k - \alpha_k d(w^k, \mu_k)]. \tag{15}$$

Step 4.(adjust μ_k)

$$\mu_{k+1} = \begin{cases} \mu_k/(1+\gamma_k), & \text{if} \quad w < \sigma, \\ (1+\gamma_k)\tau_k, & \text{if} \quad w > 1/\sigma, \\ \mu_k, & \qquad\text{otherwise,} \end{cases}$$

where

$$w = \frac{\|Ax^{k+1} - z^{k+1} - b\|}{\|z^{k+1} - z^k\|}.$$

Set $k := k+1$ and goto Step 1.

Remark 2.1. The strategy of adjusting $\{\mu_k\}$ is similar to the technique presented in [6]. From $\gamma_i \geq 0$ and $\sum_{i=0}^{\infty} \gamma_i < \infty$, it follows that

$$\Pi_{i=0}^{\infty}(1+\gamma_i) < \infty.$$

Then, the sequence $\{\mu_k\}$ is bounded, due to its updating rules in Step 4. That is to say,

$$\inf\{\mu_k\}_1^{\infty} := \mu_{\min} > 0, \quad \sup\{\mu_k\}_1^{\infty} := \mu_{\max} < \infty.$$

From now on, we assume $\mu_{\min} > \|H\|$ to ensure that the matrix $\mu_k I - H$ is positive definite.

Remark 2.2. At each iteration, the method only need to perform some projections onto simple sets and some matrix-vector multiplications, and thus its computational load is quite tiny.

To prepare for the convergence analysis of the new algorithm, we establish an important result.

Lemma 2.3. For any solution point $w^* \in W^*$, we have

$$(w^k - w^*)^\top d(w^k, \mu_k) \geq \|x^k - \bar{x}^k\|_{B_k}^2 + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2. \tag{16}$$

Proof. Setting $x = x^k - (Hx^k + c - A^\top y^k)/\mu_k$ and $z = x^*$ in (5), we have

$$(x^* - \bar{x}^k)^\top[(\mu_k I - H)x^k + A^\top y^k - c - \mu_k\bar{x}^k] \leq 0,$$

that is

$$(\bar{x}^k - x^*)^\top[B_k(x^k - \bar{x}^k) + A^\top y^k - (H\bar{x}^k + c)] \geq 0. \tag{17}$$

Similarly, we have

$$(\bar{y}^k - y^*)^\top[\mu_k(y^k - \bar{y}^k) - z^k] \geq 0. \tag{18}$$

Furthermore, because $w^* \in W^*$ is a solution of LVI(H, c, W), we can get

$$(\bar{x}^k - x^*)^\top [(Hx^* + c) - A^\top y^*] + (\bar{y}^k - y^*)^\top (Ax^* - b) \geq 0, \tag{19}$$

and

$$Ax^* - z^* - b = 0. \tag{20}$$

By adding (17)-(19), it follows that

$$(\bar{x}^k - x^*)^\top [B_k(x^k - \bar{x}^k) + A^\top (y^k - y^*)] + (\bar{y}^k - y^*)^\top [\mu_k(y^k - \bar{y}^k) - z^k + Ax^* - b] \\ \geq (\bar{x}^k - x^*)^\top H(\bar{x}^k - x^*).$$

Since H is positive semi-definite,

$$(\bar{x}^k - x^*)^\top H(\bar{x}^k - x^*) \geq 0.$$

Thus

$$(\bar{x}^k - x^*)^\top [B_k(x^k - \bar{x}^k) + A^\top (y^k - y^*)] + (\bar{y}^k - y^*)^\top [\mu_k(y^k - \bar{y}^k) - z^k + Ax^* - b] \geq 0.$$

It follows that

$$(x^k - x^*)^\top B_k(x^k - \bar{x}^k) + (y^k - y^*)^\top [\mu_k(y^k - \bar{y}^k) - z^k + A\bar{x}^k - b] \\ + (y^k - \bar{y}^k)^\top (z^k - Ax^* + b) \geq \|x^k - \bar{x}^k\|_{B_k}^2 + \mu_k \|y^k - \bar{y}^k\|^2,$$

which is equivalent to the inequality

$$\begin{aligned} &(x^k - x^*)^\top B_k(x^k - \bar{x}^k) + (y^k - y^*)^\top [\mu_k(y^k - \bar{y}^k) - z^k + A\bar{x}^k - b] \\ &\qquad + (y^k - \bar{y}^k)^\top (z^k - z^*) \geq \|x^k - \bar{x}^k\|_{B_k}^2 + \mu_k \|y^k - \bar{y}^k\|^2. \end{aligned} \tag{21}$$

From (12) and (20), we obtain

$$\begin{aligned} & \mu_k^2 \|z^k - \bar{z}^k\|^2 \\ = \; & \|Ax^k - z^k - b\|^2 \\ = \; & (Ax^k - z^k - b)^\top (Ax^k - z^k - b) \\ = \; & (Ax^k - z^k - Ax^* + z^*)^\top (Ax^k - z^k - b) \\ = \; & (x^k - x^*)^\top A^\top (Ax^k - z^k - b) - (z^k - z^*)^\top (Ax^k - z^k - b). \end{aligned} \tag{22}$$

Then by adding (21) and (22), we get (16) immediately. Q.E.D.

Remark 2.3. Note that $w^k \in W^*$, when $\bar{w}^k = w^k$. Therefore, we conclude that $\|d(w^k, \mu_k)\| \neq 0$, when $w^k \notin W^*$. This fact explains that the step size (14) is well-defined.

Remark 2.4. In fact, Lemma 2.3 has proved that $-d(w^k, \mu_k)$ is a descent direction of the merit function $\frac{1}{2}||w^k - w^*||^2$ whenever $w^k \in W$ is not a solution of LVI(H, c, W).

In the following, we assume that the Algorithm 2.1 generates an infinite sequence $\{w^k\}$, otherwise, an approximate solution $w^k \in W$ is obtained.

We first investigate the technique of identifying the optimal step sizes along the descent direction $-d(w^k, \mu_k)$. To justify the strategy of choosing the step size α_k as in Step 3, we use

$$\widetilde{w}^k(\alpha) := P_W[w^k - \alpha d(w^k, \mu_k)].$$

to denote the temporary point w^k taking α as the step size along $-d(w^k, \mu_k)$. The following lemma motivates us to identify the optimal step size along this direction.

Lemma 2.4. For given w^k and $\mu_k > 0$, we have

$$\Theta_k(\alpha) := \|w^k - w^*\|^2 - \|\widetilde{w}^k(\alpha) - w^*\| \geq \Phi_k(\alpha),$$

where

$$\Phi_k(\alpha) = -\alpha^2\|d(u^k, \beta_k)\|^2 + 2\alpha(\|x^k - \bar{x}^k\|^2_{B_k} + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2).$$

Proof. Because $\widetilde{w}^k(\alpha) := P_W[w^k - \alpha d(w^k, \mu_k)]$, by setting $x = w^k - \alpha d(w^k, \mu_k)$ and $y = w^*$ in (6), we obtain

$$\|\widetilde{w}^k(\alpha) - w^*\|^2 \leq \|w^k - \alpha d(w^k, \mu_k) - w^*\|^2 - \|w^k - \alpha d(w^k, \mu_k) - \widetilde{w}^k(\alpha)\|^2,$$

and consequently

$$\Theta_k(\alpha) \geq \|w^k - \widetilde{w}^k(\alpha)\|^2 + 2\alpha(w^k - w^*)^\top d(w^k, \mu_k) - 2\alpha(w^k - \widetilde{w}^k(\alpha))^\top d(w^k, \mu_k),$$

Since w^* is a solution, it follows from Lemma 2.3 that

$$\begin{aligned}
& \Theta_k(\alpha) \\
\geq\ & \|w^k - \widetilde{w}^k(\alpha)\|^2 + 2\alpha(\|x^k - \bar{x}^k\|^2_{B_k} + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2) \\
& -2\alpha(w^k - \widetilde{w}^k(\alpha))^\top d(w^k, \mu_k) \\
=\ & \|w^k - \widetilde{w}^k(\alpha) - \alpha d(w^k, \mu_k)\|^2 - \alpha^2\|d(w^k, \mu_k)\|^2 \\
& +2\alpha(\|x^k - \bar{x}^k\|^2_{B_k} + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2) \\
\geq\ & -\alpha^2\|d(w^k, \mu_k)\|^2 + 2\alpha(\|x^k - \bar{x}^k\|^2_{B_k} + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2) \\
=\ & \Phi_k(\alpha)
\end{aligned}$$

We get the assertion of this Lemma. Q.E.D.

Clearly, $\Theta_k(\alpha)$ means the progress made by the new iterate $w^{k+1}(\alpha)$ at the kth iteration. Therefore, in order to accelerate the convergence, it is reasonable to choose

$$\alpha_k = (\|x^k - \bar{x}^k\|^2_{B_k} + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2)/\|d(w^k, \mu_k)\|^2,$$

i.e., the optimal value of α maximizing the quadratic function $\Phi_k(\alpha)$ which provides a lower bound function of $\Theta_k(\alpha)$. Based on numerical experiences, we prefer to attach a relax factor $\tau \in (0,2)$ to α_k, and simple calculation show that

$$\Phi_k(\tau\alpha_k) = \tau(2-\tau)\Phi_k(\alpha_k) = \tau(2-\tau)\alpha_k(\|x^k - \bar{x}^k\|^2_{B_k} + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2). \qquad (23)$$

Remark 2.5. It follows from (23) that

$$\|w^{k+1} - w^*\|^2 \le \|w^k - w^*\|^2 - \tau(2-\tau)\alpha_k(\|x^k - \bar{x}^k\|^2_{B_k} + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2). \qquad (24)$$

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

Lemma 2.5. For all $k \ge 0$, we have

$$\alpha_k \ge \varsigma, \qquad (25)$$

where $\varsigma > 0$ is a constant.

Proof. It follows from (8) and (10)-(12) that

$$\begin{aligned} & \|B_k(x^k - \bar{x}^k) + A^\top(Ax^k - z^k - b)\| \\ = \; & \|(\mu_k I - H)(x^k - \bar{x}^k) + \mu_k A^\top(z^k - \bar{z}^k)\| \\ \le \; & (\mu_{\max} + \|H\|)\|x^k - \bar{x}^k\| + \mu_{\max}\|A\|\|z^k - \bar{z}^k\| \\ \le \; & (\mu_{\max} + \|H\| + \mu_{\max}\|A\|)\|w^k - \bar{w}^k\|, \end{aligned}$$

and

$$\begin{aligned} & \|\mu_k(y^k - \bar{y}^k) + (A\bar{x}^k - z^k - b)\| \\ = \; & \|\mu_k(y^k - \bar{y}^k) + (A\bar{x}^k - Ax^k) + (Ax^k - z^k - b)\| \\ \le \; & \mu_{\max}\|y^k - \bar{y}^k\| + \|A\|\|x^k - \bar{x}^k\| + \mu_{\max}\|z^k - \bar{z}^k\| \\ \le \; & (2\mu_{\max} + \|A\|)\|w^k - \bar{w}^k\|, \end{aligned}$$

and

$$\begin{aligned} & \|(y^k - \bar{y}^k) - (Ax^k - z^k - b)\| \\ \le \; & (1 + \mu_{\max})\|w^k - \bar{w}^k\|, \end{aligned}$$

This and the definition of $d(w^k, \mu_k)$ imply

$$\|d(w^k, \mu_k)\| \le c_1\|w^k - \bar{w}^k\|, \quad \forall k \ge 0, \qquad (26)$$

where

$$c_1 = 1 + \|A\| + \|H\| + (3 + \|A\|)\mu_{\max}.$$

From $\mu_k \geq \mu_{\min} > \|H\|$, it is true that

$$\begin{aligned}\|x^k - \bar{x}^k\|_{B_k}^2 &\geq (x^k - \bar{x}^k)^\top(\mu_{\min} I - H)(x^k - \bar{x}^k) \\ &\geq \lambda_{\min}\|x^k - \bar{x}^k\|^2,\end{aligned}$$

where $\lambda_{\min}$ is the minimum eigenvalue of the positive definite matrix $\mu_{\min} I - H$. Then, it follows that

$$\|x^k - \bar{x}^k\|_{B_k}^2 + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2 \geq c_2\|w^k - \bar{w}^k\|^2, \tag{27}$$

where $c_2 = \min\{\lambda_{\min}, \mu_{\min}, \mu_{\min}^2\}$. Therefore, from (26)(27) and the definition of α_k, we have

$$\alpha_k \geq \varsigma,$$

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

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

Theorem 2.1. The sequence $\{w^k\}$ generated by Algorithm 2.1 converges to a solution of LVI(H, c, W) globally.

Proof. Since $\tau \in (0, 2)$ and $\alpha_k > 0$, it follows from (24) that

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

which means that the sequence $\{w^k\}$ is bounded. Thus, it has at least one cluster point, denoted as $w^\infty = (x^\infty, y^\infty, z^\infty)^\top$, and $\{w^{k_j}\}$ be the corresponding subsequence converging to w^∞. Again from (24), we have

$$\tau(2 - \tau)\alpha_k(\|x^k - \bar{x}^k\|_{B_k}^2 + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2) \leq \|w^k - w^*\|^2 - \|w^{k+1} - w^*\|^2.$$

Summarizing both sides of the above inequality, we get

$$\begin{aligned} & \sum_{k=0}^{\infty}\{\tau(2 - \tau)\alpha_k(\|x^k - \bar{x}^k\|_{B_k}^2 + \mu_k\|y^k - \bar{y}^k\|^2 + \mu_k^2\|z^k - \bar{z}^k\|^2)\} \\ \leq\ & \sum_{k=0}^{\infty}\{\|w^k - w^*\|^2 - \|w^{k+1} - w^*\|^2\} \\ \leq\ & \|w^0 - w^*\|^2 \\ <\ & +\infty, \end{aligned}$$

which together with (25) means that

$$\lim_{k\to\infty}\|x^k - \bar{x}^k\|_{B_k}^2 = \lim_{k\to\infty}\mu_k\|y^k - \bar{y}^k\|^2 = \lim_{k\to\infty}\mu_k^2\|z^k - \bar{z}^k\|^2 = 0.$$

Thus, from $\mu_k \geq \mu_{\min} > 0$, we have

$$\lim_{k\to\infty} \|x^k - \bar{x}^k\|_{B_k}^2 = \lim_{k\to\infty} \|y^k - \bar{y}^k\|^2 = \lim_{k\to\infty} \|z^k - \bar{z}^k\|^2 = 0. \tag{28}$$

On the other hand, from (8) (10) and $\mu_{\max} \geq \mu_k \geq \mu_{\min} > \|H\|$, it is true that

$$\begin{aligned} & \|x^k - \bar{x}^k\|_{B_k}^2 \\ \geq\ & (x^k - \bar{x}^k)^\top (\mu_{\min} I - H)(x^k - \bar{x}^k) \\ \geq\ & \lambda_{\min} \|x^k - \bar{x}^k\|^2 \\ =\ & \lambda_{\min} \|e_1(w^k, 1/\mu_k)\|^2 \\ \geq\ & \lambda_{\min} \|e_1(w^k, 1/\mu_{\max})\|^2, \end{aligned}$$

where $\lambda_{\min}$ is the minimum eigenvalue of the positive definite matrix $\mu_{\min} I - H$. In a similar way, we can prove that

$$\|y^k - \bar{y}^k\| = \|e_2(w^k, 1/\mu_k)\| \geq \|e_2(w^k, 1/\mu_{\max})\|,$$

and

$$\|z^k - \bar{z}^k\| = \|e_3(w^k, 1/\mu_k)\| \geq \|e_3(w^k, 1/\mu_{\max})\|,$$

Therefore, from the above three inequalities and (28), we have

$$\lim_{k\to\infty} \|e(w^k, 1/\mu_{\max})\| = 0. \tag{29}$$

Taking the limit in (29) along the subsequence $\{w^{k_j}\}$ and using the continuity of the operator $\|e(\cdot, 1/\mu_{\max})\|$, we have

$$\|e(w^\infty, 1/\mu_{\max})\| = 0.$$

So w^∞ is a solution of LVI(H, c, 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

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

On the other hand, since $\{\|w^k - w^\infty\|\}$ is monotonically non-increasing (since (24) and that w^∞ is a solution of LVI(H, c, W)), we have $\|w^k - w^\infty\| \leq \|w^{k_0} - w^\infty\|$ for all $k \geq k_0$, and it follows that

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

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 LVI(H, c, W). This completes the proof. Q.E.D.

3 Preliminary Computational Results

In this section, we illustrate the necessity and efficiency of our methods. To this end, we also code the algorithm proposed by Wang and Luo[6].

Example 1. The example used here is a modification of the test minimax problem of Wang and Liao[6],

$$\min_{x\in X}\{\max_{y\in Y}\{\frac{1}{2}x^\top Hx + c^\top x - y^\top Ax + b^\top y\}\}.$$

In particular, $A \in R^{n\times n}$ is a randomly generated matrix as suggested in [6],

$$A = B^\top B + C + D,$$

where every entry of the $n \times n$ matrix B and the $n \times n$ skew-symmetric matrix C is uniformly generated from $(-5, 5)$, every diagonal entry of the $n \times n$ diagonal D is uniformly generated from $(0, 0.3)$, and every entry of the vectors b and c is randomly generated from $(-1, 1)$. Besides, we take

$$X = \{x | x \in R^n_+\}, \quad Y = \{y \in R^n_+ | \|y\| \le 1\},$$

and

$$H = \begin{pmatrix} 1 & 2 & \cdots & \cdots & 2 \\ 0 & 1 & 2 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 2 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

For any $y \in R^n$, $P_Y[y]$ by component is defined by

$$(P_Y[y])_i = \begin{cases} (y_+)_i, & \text{if } \|y_+\| \le 1, \\ (y_+)_i / \|y_+\|, & \text{otherwise}, \end{cases}$$

where

$$(y_+)_i = \max\{0, y_i\}.$$

Other parameters used in the Algorithm 2.1 are set as $\tau = 1.95$, $\sigma = 0.1, \gamma_i = 1$. Actually, we limit the adjustment of μ_k up to $k_{\max} = 100$ times. It is hoped and anticipated that, after at most $k_{\max}$ adequate adjustments, the parameter μ_k could be close enough to a proper value. For the method in [6], we take $\gamma = 1$, $\beta_0 = 2\|H + A^\top A\|$, and update the parameters β_k as μ_k in Algorithm 2.1. To solve the linear subVI problem in Wang and Luo's method[6], we utilize the quadratic-program solver 'quadprog.m' from the Matlab optimization toolbox. In our tests, we take

$$\varepsilon = 1 \times 10^{-5}, \quad y^0 = z^0 = (0, \cdots, 0)^\top,$$

and use the stopping criterion, $||w^k - \bar{w}^k|| \le \varepsilon$ for both methods. All programs are coded in Matlab. 'N' denotes the dimension of the tested problem, and 'IN' denotes the number of iterations and 'CPU' denotes the CPU time in seconds. We also code the proposed algorithm with a fixed parameter μ throughout the entire algorithm without any change, denoted in the tables as "Method (F)", and "Method(V)" means Algorithm 2.1.

Table 1: Numerical results for different μ_0

v		5	10	15	20	25	30	35	40
Wang and	IN	59	56	58	54	55	54	54	51
Liao's method	CPU	0.92	0.90	0.91	0.83	0.93	0.86	0.81	0.79
Proposed	IN	4	4	3	2	2	3	2	2
Algorithm 2.1	CPU	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 2: Numerical results with randomly generated $x^0 \in (0, 1)$

N		50	100	200	300	400	500	1000
	IN	130	132	112	108	99	84	56
Method (F)	CPU	0.02	0.08	0.46	1.89	3.79	6.18	37.63
	IN	5	4	4	4	3	3	2
Method (V)	CPU	0.01	0.10	0.19	0.68	1.67	3.47	40.62

Table 3: Numerical results with $x^0 = (1, \cdots, 1)^\top$

N		50	100	200	300	400	500	1000
	IN	68	189	195	232	176	148	130
Method (F)	CPU	0.01	0.10	0.70	3.27	5.46	8.56	48.78
	IN	6	5	5	4	4	4	4
Method (V)	CPU	0.01	0.11	0.20	0.75	1.67	3.37	40.90

We conduct the numerical study with dimensions varying from 10 to 1000, and with different initial points. Table 1 reports the results with the initial points as x^0 generated randomly between $(0, 1)$, $n = 10$, and $\mu_0 = v\|H + A^\top A\|$ with different v; Table 2 reports the results with the initial points as x^0 generated randomly between $(0, 1)$; Table 3 reports the results with the initial point as $x^0 = (1, \cdots, 1)$;

Table 4: Numerical results with randomly generated $x^0 \in (0, 10)$

N		50	100	200	300	400	500	1000
	IN	200	330	336	396	472	490	296
Method (F)	CPU	0.05	0.16	1.10	5.42	11.90	19.34	73.39
	IN	8	7	7	7	7	6	6
Method (V)	CPU	0.01	0.10	0.18	0.70	1.79	3.53	41.02

and Table 4 reports the results with the initial point x^0 generated randomly between $(0, 10)$. In Table 2-4, we set $\mu_0 = 2\|H + A^\top A\|$ for the both methods: Method(F) and Method(V).

The results in the Table 1 indicate that the performance of the Algorithm 2.1 is much better than that of Wang and Liao's method. The reason for this result may be that Algorithm 2.1 only needs to perform some projections onto simple sets and some matrix-vector multiplications in each iteration, while Wang and Liao's method need to solve a subvariational inequality problem, which is difficult to solve efficiently and exactly in each iteration. Also from this table, we can observe that the number of iteration both methods varies slightly with different initial parameters.

The results summarized in Tables 2-4 show that adjusting the parameter μ significantly improves both the CPU time and the iteration number compared to the case with fixed μ. The self-adaptive strategy make the method more robust than with a fixed μ, and thus it is important to adapt μ dynamically according to different problems.

Example 2. To show the advantage of the new prediction-correction method for large scale problems, we implement Algorithm 2.2 to a set of spatial price equilibrium problems. The details of these problems can be found in [5], as follows:

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

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

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

$$x_{ij} \geq 0,$$

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 $j = 1, \cdots, n$, and the other parameters $\tau = 1.98, \mu_0 = 21\|H\|$, $\alpha_1 = 0.5, \alpha_2 = 2, v_1 = 0.1/n, v_2 = 0.9/n, \mu_{\min} = 5\|H\|, \mu_{\max} = 50\|H\|$. The calculations were started with $w^0 = 0$ and stopped when

$$\max\{\frac{\|e_1(u^k)\|}{\|c\|}, \frac{\|e_2(u^k)\|}{\|b\|}\} \le \varepsilon,$$

for some prescribed $\varepsilon > 0$, where $b = (s, d)^\top$. The computational results are given in Table 5 for some m and n. The numerical results given in Table 5 show that Algorithm 2.2 is relatively efficient, and it is

Table 5: Number of iterations for different scale and precisions

m	n	m×n	$\varepsilon = 0.1$	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$
5	10	50	12	33	126	239
10	10	100	13	36	132	273
10	15	150	17	99	178	264
20	25	500	21	62	183	409
30	40	1200	18	82	286	522

attractive from a computational point of view.

4 Conclusions

Based on the alternating direction method, we observe a new descent direction at each iteration, and present a new self-adaptive prediction-correction method for LVI(H, c, S). The new method uses only the function values and the total computational cost is very small. Thus, the new method is applicable in practice. Under mild conditions, we proved the global convergence of the method. Some preliminary computational results illustrated the efficiency of the algorithm.

References

[1] M.C.Ferris, J.S.Pang, Engineering and economic applications of complementarity problems, *SIAM Review*, **39** (1997), 669–713.

[2] D.P.Bertsekas, E.M.Gafni, Projection method for variational inequalities with applications to the traffic assignment problem. *Mathematical Programming Study*, **17** (1982), 139–159.

[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] Wang S.L., Liao L.Z.: Decomposition method with a variable parameters for a class of monotone variational inequality problems. *Journal of Optimization Theory and Application*, **109** (2001), 415–429.

[7] 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.

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

[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, Solving variational inequality problems with linear constraints by a proximal decomposition algorithm. *Journal of Global Optimization*, **28** (2004), 97–113.

[11] 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.

[12] 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.

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

[14] Peng J.M., Yuan Y.X., Unconstrained methods for generalized complementarity problems. *Journal of Computational Mathematics*, **15** (1997), 253–264.

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