A Descent-projection Method for Solving Monotone Structured Variational Inequalities

Min Sun, Zhenyu Liu

Abstract—In this paper, a new descent-projection method with a new search direction for monotone structured variational inequalities is proposed. The method is simple, which needs only projections and some function evaluations, so its computational load is very tiny. Under mild conditions on the problem’s data, the method is proved to converge globally. Some preliminary computational results are also reported to illustrate the efficiency of the method.

Keywords—variational inequalities, monotone function, global convergence.

I. INTRODUCTION

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

\[
(u - u^*)^T F(u^*) \geq 0, \quad \forall u \in \Omega. \tag{1}
\]

Problem VI includes nonlinear complementarity problems (when \( \Omega = \mathbb{R}^n_+ \)) and system of nonlinear equations (when \( \Omega = \mathbb{R}^n \)). VI problems find important applications in many fields, such as mathematical programming, network economics, traffic assignment, game theoretic problems, etc. In the past decades, researchers have developed many efficient numerical algorithms for solving VI problems, such as Newton-type methods, projection-type methods. For theory, numerical methods and applications, the interested reader is referred to the excellent monographs of Nagurney [1] Faccinetti and Pang [2], and the references therein.

Let \( X \subseteq \mathbb{R}^n \) and \( Y \subseteq \mathbb{R}^m \) be given nonempty closed convex sets; \( f : X \to \mathbb{R}^n \) and \( g : Y \to \mathbb{R}^m \) be given continuous monotone operators. In this paper, we focus our attention on the variational inequality problem (1) that

\[
u = \left( \begin{array}{l} x \\ y \end{array} \right), \quad F(u) = \left( \begin{array}{l} f(x) \\ g(y) \end{array} \right), \quad \Omega = \{(x, y) | x \in X, y \in Y, Ax + By = b\},
\]

where \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{m \times m} \) are given matrices; \( b \in \mathbb{R}^r \) is a given vector. The above problem is denoted by VI\((F, \Omega)\). This well-structured model captures many interesting concrete applications in diversified problems [3].

By attaching a Lagrange multiplier vector \( \lambda \in \mathbb{R}^r \) to the linear constraints \( Ax + By = b \), VI\((F, \Omega)\) can be equivalently transformed into the following compact form, denoted by VI\((Q, W)\): Find \( w^* \in W \), such that

\[
(w - w^*)^T Q(w^*) \geq 0, \quad \forall w \in W \tag{2}
\]

where

\[
w = \left( \begin{array}{l} x \\ y \\ \lambda \end{array} \right), Q(w) = \left( \begin{array}{l} f(x) - A^T \lambda \\ g(y) - B^T \lambda \\ A x + B y - b \end{array} \right),
\]

\( W = X \times Y \times \mathbb{R}^r, \) and \( Z = \mathbb{R}^r \).

For solving (2), the classical alternating direction method (ADM) proposed in [4,5] plays a fundamental role, and it has been motivating many efficient numerical algorithms[6-8,12].

As pointed in [9], an iterative method converges slowly if its generated descent direction \( d(u^k, \beta_k) \) or the step-size \( \rho(u^k, \beta_k) \) tend to zero. Recently, for co-coercive variational inequalities, Yan, Han and Sun [9] proposed a self-adaptive projection method which direction \( d(u^k, \beta_k) \) or step-size \( \rho(u^k, \beta_k) \) don't converge to zero. However, the condition for the convergence of the method, i.e., the assumption of co-coercive monotonicity, is stringent, which precludes the application of the method in reality. To overcome it, for pseudomonotone monotone variational inequalities, Zhang and Liu [10], Yu, Shao and Wang [11] proposed some projection methods which either direction \( d(u^k, \beta_k) \) or step-size \( \rho(u^k, \beta_k) \) doesn't converge to zero.

Inspired by these methods [9-11] for VI problems, a descent-projection method for solving monotone VI\((Q, W)\) is proposed in this paper, which can be viewed as an combination of the two methods proposed by Yu, Shao and Wang [11] and Sun [12]. More specifically, we have adopted both methods' technique to design our search direction.

The remainder of the paper is organized as follows. Some definitions and properties used in this paper are presented in Section 2. In Section 3, the descent-projection method is described formally and its global convergence is proved. Some preliminary computational results are given in Section 4 and some conclusions are given in Section 5.

II. PRELIMINARIES

In this section, some definitions and results from the literature are presented which are used throughout the paper.

In the following, we assume that the set \( Z \) in VI\((Q, W)\) is a general convex set in \( \mathbb{R}^r \). Let \( P_W(\cdot) \) denote the orthogonal projection mapping from \( \mathbb{R}^{n+m+r} \) onto \( W \). It is well known that VI\((Q, W)\) is equivalent to the projection equation

\[
\text{argmin}_{w \in W} \langle Q(w), w \rangle.
\]
\[ e(w, \beta) = \begin{pmatrix} e_1(w, \beta) \\ e_2(w, \beta) \\ e_3(w, \beta) \end{pmatrix} = \begin{pmatrix} x - P_X [x - \beta (f(x) - A^\top \lambda)] \\ y - P_Y [y - \beta (g(y) - B^\top \lambda)] \\ \lambda - P_Z [\lambda - \beta (Ax + By - b)] \end{pmatrix}. \]

A basic property of the projection mapping \( P_X(\cdot) \) is

\[(x - P_X(x))\top (x' - P_X(x)) \leq 0, \quad \forall x \in X, x' \in R^n. \tag{3}\]

From (3) and the Cauchy-Schwarz inequality we can see that the projection operator \( P_X(\cdot) \) is nonexpansive, namely

\[ \|P_X(x) - P_X(x')\| \leq \|x - x'\|, \quad \forall x, x' \in R^n. \]

**Lemma 2.1.** For all \( w \in R^{n+m+r} \) and \( \rho_1 > \rho_2 > 0 \), it holds that

\[ \|e(w, \rho_1)\| \leq \|e(w, \rho_2)\| \tag{4} \]

and

\[ \|e(w, \rho_1)\| \leq \|e(w, \rho_2)\|. \tag{5} \]

**Definition 2.1.** A mapping \( f : R^n \rightarrow R^n \) is said to be monotone if

\[(x - x')\top (f(x) - f(x')) \geq 0, \quad \forall x, x' \in R^n. \]

We make the following standard assumptions throughout this paper:

**Assumptions.**
- \( f \) and \( g \) are monotone mappings on \( X \) and \( Y \), respectively.
- The solution set of problem \( \text{VI}(G, W) \), denoted by \( W^* \), is nonempty.
- \( X \) and \( Y \) are simple closed convex sets. That is, the projection onto the set is simple to carry out.

**III. ALGORITHM AND CONVERGENCE**

Firstly, setting \( e_i = e_i(w, \beta), i = 1, 2, 3, F = f(x) - A^\top \lambda, G = g(y) - B^\top \lambda \) for convenience.

**Lemma 3.1.** Let \( w^* = (x^*, y^*, z^*) \in W^* \) be an arbitrary solution of \( \text{VI}(Q, W) \),

\[ g(w, \beta) = \begin{pmatrix} e_1 + \beta f(x - e_1) - \beta f(x) + \beta A^\top e_3 \\ e_2 + \beta g(y - e_2) - \beta g(y) + \beta B^\top e_3 \\ e_3 - \beta A e_1 - \beta B e_2 \end{pmatrix}, \]

then for any \( w = (x, y, z) \in R^{n+m+r} \), we have

\[ (w - w^*)\top g(w, \beta) \geq \phi(w, \beta), \]

where \( \phi(w, \beta) = \|e(w, \beta)\|^2 - \beta e_1^\top [f(x - f(x - e_1)] - \beta e_2^\top g(y - g(y - e_2)]. \]

**Proof.** Setting \( x := x - \beta f(x) - A^\top \lambda \) and \( x' := x^* \) in (3), we have

\[ \{x - \beta f(x) - A^\top \lambda]\top P_X[x - \beta f(x) - A^\top \lambda] \geq 0, \]

i.e.,

\[ e_1^\top (x - x^*) \geq \|e_1\|^2 + \beta [f(x) - A^\top \lambda]^\top [x - x^* - e_1]. \tag{6} \]

From \( w^* \) is a solution of \( \text{VI}(Q, W) \), we get

\[ \{f(x^*) - A^\top \lambda^*\}^\top P_X[x - \beta f(x) - A^\top \lambda] = x^* \geq 0, \]

that is

\[ \beta [f(x^*) - A^\top \lambda^*]^\top (x^* - x - e_1) \geq 0. \tag{7} \]

From the monotonicity of \( f \), we obtain

\[ \beta [f(x - e_1) - f(x^*)]^\top (x - e_1 - x^*) \geq 0, \]

i.e.,

\[ \beta [f(x - e_1) - f(x^*)]^\top \geq \beta e_1^\top [f(x - e_1) - f(x^*]) + \beta (x - x^*)^\top f(x^*). \tag{8} \]

Adding (6)-(8), we get

\[ \|e_1\|^2 + \beta \|x - x^*\|^2 \geq \|e_2\|^2 + \beta \|y - y^*\|^2 \geq \|e_3\|^2 + \beta \|x - x^*\|^2. \]

In a similar way, we have

\[ \|e_1\|^2 + \beta \|x - x^*\|^2 \geq \|e_2\|^2 + \beta \|y - y^*\|^2 \geq \|e_3\|^2 + \beta \|x - x^*\|^2. \]

Then, we can get the assertion of this lemma from the above inequality. This completes the proof.

**Remark 3.1.** \( g(w, \beta) \) in Lemma 3.1 is different from the search direction \( d(w, \beta) \) in [12], because \( e_3(w, \beta) \) in both directions is different.

Now, we give the new descent-projection method.

**Algorithm 3.1.**

**Step 0:** Given \( \varepsilon > 0 \). Choose \( w^0 \in W \) and \( \gamma \in (0, 2) \).

**Step 1:** Compute \( e(w^k, \beta_k) \). If \( \|e(w^k, \beta_k)\| < \varepsilon \), then stop; otherwise, go to step 2.
Lemma 3.1. The step-size defined by (15) has a positive value.

\[ \beta_k \left( \| f(x^k) - f(x^k - e_1(w^k, \beta_k)) \| + \| g(y^k) - g(y^k - e_2(w^k, \beta_k)) \| \right) \leq L_1 \| e(w^k, \beta_k) \|. \]  

(14)

Step 3: Calculate

\[ \rho_k = \frac{(1 - L_1) \| e(w^k, \beta_k) \|^2}{\| g(w^k, \beta_k) + \tau_k Q(w^k) \|^2}, \]  

(15)

where \( \tau_k = \frac{\| e(w^k, \beta_k) \|}{\| Q(w^k) \|} \) is defined in Lemma 3.1.

Step 4: Compute the new iterate

\[ w^{k+1} = P_{W}[w^k - \gamma \rho_k d(w^k, \beta_k)], \]

where

\[ d(w^k, \beta_k) = g(w^k, \beta_k) + \tau_k Q(w^k). \]  

(16)

Step 5: If

\[ \beta_k \left( \| f(x^k) - f(x^k - e_1(w^k, \beta_k)) \| + \| g(y^k) - g(y^k - e_2(w^k, \beta_k)) \| \right) \leq L_2 \| e(w^k, \beta_k) \|. \]  

(17)

then \( \beta = \beta_k / \eta; \) else \( \beta = \beta_k. \) Set \( k := k + 1, \) and go to Step 1.

Lemma 3.2. If \( w \in W \) is not a solution of \( VI(Q, W) \), then there exist \( L \in (0, 1) \) and \( \beta > 0 \), such that for all \( \beta \in (0, \beta) \),

\[ \beta \left( \| f(x) - f(x - e_1(w, \beta)) \| + \| g(y) - g(y - e_2(w, \beta)) \| \right) \leq L \| e(w, \beta) \|. \]

(18)

Lemma 3.3. For any \( w^* \in W^* \), we have

\[ (w^k - w^*)^\top d(w^k, \beta_k) \geq (1 - L_1) \| e(w^k, \beta_k) \|^2. \]  

(19)

Proof. It follows from (16) that

\[
\begin{align*}
(w^k - w^*)^\top d(w^k, \beta_k) & = (w^k - w^*)^\top g(w^k, \beta_k) + \tau_k (w^k - w^*)^\top Q(w^k) \\
& \geq (w^k - w^*)^\top g(w^k, \beta_k) + \tau_k (w^k - w^*)^\top Q(w^*) \\
& \geq (1 - L_1) \| e(w^k, \beta_k) \|^2,
\end{align*}
\]

where the first inequality follows from the monotonicity of \( Q(\cdot) \), and the second inequality follows from \( w^* \in W^* \), and the last inequality follows from Lemma 3.1 and Lemma 3.2.

The proof is completed.

Lemma 3.3 indicates that \( -d(w^k, \beta_k) \) is a descent direction of the unknown distance function \( \| w^k - w^* \|^2 / 2 \).

From Lemma 3.1 and Lemma 3.2, it is easy to deduce that there exist a constant \( \zeta > 0 \), such that

\[ g(w^k, \beta_k) \leq \zeta \| e(w^k, \beta_k) \|. \]  

(20)

Lemma 3.4. The step-size defined by (15) has a positive lower bound from below. That is, there exist a constant \( \nu > 0 \), such that

\[ \rho_k \geq \nu (1 - L_1), \forall k \geq 0. \]  

(21)

Proof. Note that \( \tau_k = \frac{\| e(w^k, \beta_k) \|}{\| Q(w^k) \|} \), it follows that

\[
\begin{align*}
\| g(w^k, \beta_k) + \tau_k Q(w^k) \| & = \| g(w^k, \beta_k) + \tau_k Q(w^k) \| \\
& = \| g(w^k, \beta_k) \|^2 + 2 \| e(w^k, \beta_k) \| \| g(w^k, \beta_k) \| Q(w^k) \\
& + \| e(w^k, \beta_k) \|^2 \\
& \leq \| g(w^k, \beta_k) \|^2 + 2 \| e(w^k, \beta_k) \| \| g(w^k, \beta_k) \| \\
& + \| e(w^k, \beta_k) \|^2 \\
& \leq \frac{1}{\nu} \| e(w^k, \beta_k) \|^2,
\end{align*}
\]

where the last inequality follows from (19) and \( \nu = 1/(1+\zeta)^2 \). Hence, it holds that

\[ \rho_k = \frac{(1 - L_1) \| e(w^k, \beta_k) \|^2}{\| g(w^k, \beta_k) + \tau_k Q(w^k) \|^2} \geq \nu (1 - L_1). \]

This completes the proof.

Theorem 3.1. The sequence \( \{ w^k \} \) generated by Algorithm 3.1 is bounded. More specifically, we have

\[ \| w^{k+1} - w^* \|^2 \leq \| w^k - w^* \|^2 - \gamma (2 - \gamma) (1 - L_1) \rho_k \| e(w^k, \beta_k) \|^2. \]  

(22)

Proof. From the nonexpansivity of the projection operator and \( w^* \in W^* \), we have

\[ \| w^{k+1} - w^* \|^2 \leq \| w^k - w^* \|^2 - 2 \gamma \rho_k \| e(w^k, \beta_k) \|^2. \]  

(20)

From (20) and (21), we have

\[ \gamma (2 - \gamma) (1 - L_1) \rho_k \| e(w^k, \beta_k) \|^2. \]

where the second inequality follows from (18). This completes the proof.

Theorem 3.2. Suppose that the Assumptions hold. Then

\[ (1) \lim_{k \to \infty} \| e(w^k, \beta_k) \| / \| \beta_k \| = 0. \]

(2) The whole sequence \( \{ w^k \} \) converges to a solution of \( VI(Q, W) \).

Proof. Let \( w^* \) be a solution of \( VI(Q, W) \). From (20) and (21), we have

\[ \gamma (2 - \gamma) (1 - L_1) \| e(w^k, \beta_k) \|^2. \]

which means that

\[ \lim_{k \to \infty} \| e(w^k, \beta_k) \| = 0. \]  

(23)

(1) Suppose that there is an infinite index set \( K_0 \), such that

\[ \| e(w^k, \beta_k) \| / \| \beta_k \| \geq \varepsilon > 0. \]  

\[ \forall k \in K_0. \]

From (22), we have,

\[ \lim_{k \to \infty, k \in K_0} \| \beta_k \| = 0. \]
From the bound of \(\{w^k\}\), we have that \(\{f(x^k) - A^T\lambda^k\}, \{g(y^k) - B^T\lambda^k\}\) are also bounded. From the non-expansivity of the projection operator, we have

\[
\|x^k - PX[x^k - \beta_k(f(x^k) - A^T\lambda^k)]/\mu\| \\
\leq \beta_k \|f(x^k) - A^T\lambda^k\|/\mu \\
\leq \beta_k \|y^k - PV[y^k - \beta_k(g(y^k) - B^T\lambda^k)]/\mu\| \\
\leq \beta_k \|y^k - B^T\lambda^k\|/\mu \\
\to 0.
\]

By the choice of \(\beta_k\) we know that \(\beta_k \to 0\) is not satisfied for \(m_k - 1\). That is,

\[
\|f(x^k) - f(PX[x^k - \beta_kF^k]/\mu])\| \\
+ \|g(y^k) - g(PY[y^k - \beta_kG^k]/\mu])\| \\
\geq L_1 \|e(w^k, \beta_k/\mu]\|/\beta_k \\
\geq \mu L_1 \|e(w^k, \beta_k]\|.
\]

This contradicts with (23). Thus the assertion of (1) holds.

(2) We divide our proof into two cases: \(\limsup_{k \to \infty} \beta_k > 0\) and \(\lim_{k \to \infty} \beta_k = 0\).

(i) There is \(e_0 > 0\) and an infinite set \(K_1\), such that \(\beta_k \geq e_0\). If \(k \in K_1\). From (4), we have \(\|e(w^k, \beta_k]\| \geq \|e(w^k, e_0]\|\), if \(k \in K_1\). Combining (22), we get

\[
\|e(w^k, e_0]\| \\
\to 0.
\]

Since \(\{w^k\}\) is bounded, it has a cluster point \(\bar{w} \in W\) such that \(\|e(\bar{w}, e_0]\| = 0\). That is, \(\bar{w}\) is a solution of VI\((Q, W)\). From Theorem 3.1, we have

\[
\|w^{k+1} - \bar{w}\| \leq \|w^k - \bar{w}\|.
\]

Thus the whole sequence \(\{w^k\}\) converges to \(\bar{w}\), a solution of VI\((Q, W)\).

(ii) For sufficiently large \(k\), from (5), we have

\[
\|e(w^k, \beta_k]\|/\beta_k \\
\geq \|e(w^k, 1]\|.
\]

From (1) of this theorem and the above inequality, we get

\[
\|e(w^k, 1]\| \\
\to 0.
\]

The remainder is similar to the proof in (i). This completes the proof.

IV. PRELIMINARY COMPUTATIONAL RESULTS

In this section, we implemented Algorithm 3.1 in MATLAB and tested it on a PC.

**Problem 4.1** We consider the following optimization problem:

\[
\min \frac{1}{2}\|f(x)\|^2 + \frac{1}{2}\|g(y)\|^2 \\
\text{s.t.} \quad x - y = 0, \\
x \geq 0, \quad y \geq 0.
\]

The KKT condition of the above problem is the following variational inequality: Find \(w^* = (x^*, y^*, z^*)^T \in W\) such that

\[
\begin{cases} 
(x - x^*)^T(f(x^*) - z^*) \geq 0, \\
(y - y^*)^T(g(y^*) + z^*) \geq 0, \\
x^* - y^* = 0,
\end{cases}
\]

where

\[
W = R^n_x \times R^n_y \times R^n_z.
\]

The problem (24) is the special case of problem VI\((Q, W)\) for \(A = I_n\), \(B = -I_n\), and \(b\) is a zero vector of order \(n\).

In the following numerical experiments, we set \(L_1 = 0.4\), \(L_2 = 0.6\), \(\gamma = 1.98\), \(\mu = 0.7\), \(w^0 = 0\).

**Problem 4.2** In (24), we take \(A = \text{rand}(n)\) and \(B = -\text{rand}(n)\).

**TABLE I**

<table>
<thead>
<tr>
<th>Problem size</th>
<th>n</th>
<th>No. of iter.</th>
<th>CPU time(s)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>78</td>
<td>0.0301</td>
<td>2.3929 \times 10^{-7}</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>82</td>
<td>0.0701</td>
<td>1.6979 \times 10^{-7}</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>91</td>
<td>0.1102</td>
<td>2.3856 \times 10^{-7}</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>88</td>
<td>2.8441</td>
<td>4.7576 \times 10^{-7}</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>89</td>
<td>7.0301</td>
<td>5.7576 \times 10^{-7}</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE II**

<table>
<thead>
<tr>
<th>Problem size</th>
<th>n</th>
<th>No. of iter.</th>
<th>CPU time(s)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>75</td>
<td>0.0100</td>
<td>9.0246 \times 10^{-7}</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>87</td>
<td>0.0300</td>
<td>9.5531 \times 10^{-7}</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>228</td>
<td>0.2303</td>
<td>9.5993 \times 10^{-7}</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>933</td>
<td>28.3408</td>
<td>9.8335 \times 10^{-7}</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 and Table 2 show the computational results. There, we report the number of iterations it takes to convergence, the run time in seconds, and the error \(\|e(w^k, \beta_k]\|\). Numerical results show that the suggested method is effective for the problems considered here.

V. CONCLUSION

In this paper, we observe a new descent direction at each iteration, and present a descent-projection method for monotone structured variational inequalities. Under some mild conditions, we proved the global convergence of the two new methods. Some preliminary computational results illustrated the efficiency of the proposed method.
ACKNOWLEDGMENT

This work was supported by the Shangdong Outstanding Young Scientists Award Project under grant no. BS2010SF031.

REFERENCES