A NEW INEXACT PROXIMAL POINT ALGORITHM FOR STRUCTURED GENERAL VARIATIONAL INEQUALITIES

Qidi Sui
Department of Mathematics and Information Science
Zaozhuang University
Shandong, 277160, P.R. CHINA

Abstract: In this paper, we present a new inexact proximal point algorithm (IPPA) for structured general variational inequalities. The resulted subproblems are allowed to be solved approximately, where the accuracy constant can be relaxed into any constant in the interval (0, 1), and some correction steps are required in order to ensure convergence. The global convergence of the proposed method is proved under mild conditions.

AMS Subject Classification: 90C30, 90C33
Key Words: general variational inequalities, IPPA, global convergence

1. Introduction

Variational inequalities theory, which was introduced by Stampaccia [1], provides us with a simple, general and unified framework in which to study a wide class of problems arising in pure and applied sciences. In recent years, variational inequalities have been extended in different directions. An important extension of variational inequalities is called general variational inequalities (GVI)[2].
In this paper, we focus our attention on the following special GVI: the structured general variational inequalities, denote by GVI$(F,G,\Omega)$, which is to find $u^* \in \mathbb{R}^{m+n}$ such that
\[
F(u^*) \in \Omega \quad \text{and} \quad (u - F(u^*))^\top G(u^*) \geq 0, \quad \forall u \in \Omega,
\]
(1)
where
\[
u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad F(u) = \begin{pmatrix} f_1(x) \\ f_2(y) \end{pmatrix}, \quad G(u) = \begin{pmatrix} g_1(x) \\ g_2(y) \end{pmatrix},
\]
and
\[
\Omega = \{(x,y)|x \in \mathcal{X}, y \in \mathcal{Y}, Ax + By = b\},
\]
where $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{Y} \subseteq \mathbb{R}^m$ are given nonempty closed convex sets; $f_1, g_1$ are given continuous operators from $\mathbb{R}^n$ into itself and $f_2, g_2$ are given continuous operators from $\mathbb{R}^m$ into itself; $A \in \mathbb{R}^{r \times n}$ and $B \in \mathbb{R}^{r \times m}$ are given matrices; $b \in \mathbb{R}^r$ is a given vector. If $F(u) = u$, GVI$(F,G,\Omega)$ reduces to structured variational inequalities which have been well studied in the literature[3-5].

By attaching a Lagrange multiplier vector $\lambda \in \mathbb{R}^r$ to the linear equality constraint $Ax + By = b$, problem GVI$(F,G,\Omega)$ can be equivalently transformed into the following compact form, denoted by GVI$(M,N,W)$: Find $w^* \in \mathbb{R}^{m+n+r}$, such that
\[
M(w^*) \in W \quad \text{and} \quad (w - M(w^*))^\top N(w^*) \geq 0, \quad \forall w \in W,
\]
(2)
where
\[
w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad M(w) = \begin{pmatrix} f_1(x) \\ f_2(y) \\ \lambda \end{pmatrix}, \quad N(w) = \begin{pmatrix} g_1(x) - A^\top \lambda \\ g_2(y) - B^\top \lambda \\ Af_1(x) + Bf_2(y) - b \end{pmatrix},
\]
\[
W = \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^r.
\]

Proximal Point Algorithm (PPA)[6] is an effective numerical approach to solving classical variational inequalities, and it is natural to extend PPA method to solve general variational inequalities. This was accomplished by Li in [7] recently. However, the study of PPA method to solve structured general variational inequalities is quite little. To mention a few, Li et al. [8] proposed some PPA-based methods for linearly constrained general variational inequalities. Sun et al. [9] presented an APPA method for structured variational inequalities, in which the subproblems are solved approximately with summable accuracy. In this paper, motivated by [8-9], we propose a new PPA-based algorithm for GVI$(F,G,\Omega)$. The resulted subproblems are allowed to be solved
approximately, where the accuracy constant can be relaxed into any constant in the interval $(0, 1)$, and some correction steps are required in order to ensure convergence.

The remainder of the paper is organized as follows. In Section 2, we summarize some basic concepts about variational inequalities. In Section 3, the new IPPA method is described formally and its global convergence is also proved. Some 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. The projection of a point $x \in \mathbb{R}^n$ onto the closed convex set $K \subset \mathbb{R}^n$, denoted by $P_K[x]$, is defined as the nearest vector $y \in K$ to $x$, i.e.,

$$P_K[x] = \arg \min_{y \in K} \| x - y \|.$$

The projection mapping $P_K[\cdot]$ has the following important properties.

**Lemma 2.1.** (see [10]) Let $K$ be a nonempty closed convex subset of $\mathbb{R}^n$. For any $x, y \in \mathbb{R}^n$ and any $z \in K$, the following properties hold:

$$(x - P_K[x])^\top (z - P_K[x]) \leq 0.$$  

$$(x - P_K[x]) - (y - P_K[y]) \leq \| x - y \|^2 - \| P_K[x] - P_K[y] \|^2.$$ 

It is well known[10] that problem GVI($M, N, W$) is equivalent to finding zeros of

$$e(w, \beta) = \begin{pmatrix} f_1(x) - P_X[f_1(x) - \beta(g_1(x) - A^T \lambda)] \\ f_2(y) - P_Y[f_2(y) - \beta(g_2(y) - B^T \lambda)] \\ \beta[Af_1(x) + Bf_2(y) - b] \end{pmatrix}.$$ 

The following definitions are used throughout this paper.

**Definition 2.1.** Let $f$ and $g$ be mappings from $\mathbb{R}^n$ into itself. Then:

(a) The mapping $f$ is said to be $g$–monotone if and only if

$$(f(s) - f(t))^\top (g(s) - g(t)) \geq 0, \quad \forall s, t \in \mathbb{R}^n;$$

(b) The mapping $f$ is said to be uniformly $g$–strongly monotone if there exists a positive constant $\mu > 0$ such that
Throughout this paper, we make the following assumptions:

**Assumptions:**

- The mapping \( f_i \) is \( g_i \)-monotone, \( i = 1, 2 \).
- The solution set of GVI\((M, N, W)\), denoted by \( W^* \), is nonempty.
- \( X \) and \( Y \) are simple closed convex sets in the sense that the projection onto them is easy to compute. (e.g. the positive orthant, boxed set, ball).

With the above assumptions, it is easily verify that

\[
(M(w) - M(\tilde{w}))^\top (N(w) - N(\tilde{w}))
= \left( f_1(x) - f_1(\tilde{x}) \right)^\top \left( (g_1(x) - g_1(\tilde{x})) - A^\top (\lambda - \tilde{\lambda}) \right)
+ \left( f_2(y) - f_2(\tilde{y}) \right)^\top \left( (g_2(y) - g_2(\tilde{y})) - B^\top (\lambda - \tilde{\lambda}) \right)
+ \lambda - \tilde{\lambda}
\geq 0,
\]

which implies that the mapping \( M \) in GVI\((M, N, W)\) is \( N \)-monotone.

### 3. IPPA Method and Global Convergence

In this section, we present the new IPPA method for solving GVI\((M, N, W)\) and show its global convergence. We first denote

\[
R(w) = \begin{pmatrix}
\beta (g_1(x) - A^\top \lambda) \\
\beta (g_2(y) - B^\top \lambda) \\
\lambda
\end{pmatrix},
\]

and

\[
P_k(w) = \begin{pmatrix}
f_1(x) + \beta [g_1(x) - g_1(x^k) - A^\top (\lambda - \lambda^k)] \\
f_2(y) + \beta [g_2(y) - g_2(y^k) - B^\top (\lambda - \lambda^k)] \\
\lambda
\end{pmatrix},
\]

\[
Q_k(w) = \begin{pmatrix}
-\beta (g_1(x) - A^\top \lambda) \\
-\beta (g_2(y) - B^\top \lambda) \\
\lambda - \lambda^k + \beta (Af_1(x) + Bf_2(y) - b)
\end{pmatrix}.
\]

According to the assumption (D), it is easy to verify that \( R \) is a continuous and nonsingular mapping, i.e., there exists a positive constant \( \rho_R \) such that

\[
\|R(w) - R(\tilde{w})\| \geq \rho_R \|w - \tilde{w}\|, \quad \forall w, \tilde{w} \in W.
\]
Now we introduce the IPPA method as follows:

**Algorithm 3.1.** The IPPA with constant accuracy for GVI($\mathcal{M}, \mathcal{N}, \mathcal{W}$)

**Step 0.** Given $\varepsilon > 0$, choose $w^0 = (x^0, y^0, \lambda^0)^\top \in \mathcal{W}$, $\beta > 0$, $\delta \in (0, 1)$, $\gamma \in (0, 2)$ and set $k:=0$.

**Step 1.** If $\|e(w^k, \beta)\| \leq \varepsilon$, then stop. Otherwise, goto Step 2.

**Step 2.** Generate temporal iterate $\tilde{w}^k$ such that
\[
P_k(\tilde{w}^k) \in \mathcal{W}, \quad [w - (P_k(\tilde{w}^k) - \xi^k)]^\top Q_k(\tilde{w}^k) \geq 0, \quad \forall w \in \mathcal{W},
\]
with the inexactness criterion
\[
\|\xi^k\| \leq \delta \|R(w^k) - R(\tilde{w}^k)\|.
\]

**Step 3.** Compute the next iterate $w^{k+1}$ via
\[
R(w^{k+1}) = R(w^k) - \alpha_k d(w^k, \tilde{w}^k, \xi^k)
\]
where
\[
d(w^k, \tilde{w}^k, \xi^k) = [R(w^k) - R(\tilde{w}^k)] + \xi^k,
\]
\[
\alpha_k = \gamma \alpha_k^*
\]
and
\[
\alpha_k^* = \frac{[R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi^k)}{\|d(w^k, \tilde{w}^k, \xi^k)\|^2}.
\]

**Step 4.** With $k = k + 1$, go to Step 1.

**Remark 3.1.** According to (6), the GVI (10) is characterized by the projection equation:
\[
P_k(w^{k+1}) - \xi^k = P_Y[P_k(w^{k+1}) - \xi^k - Q_k(w^{k+1})],
\]
which is equivalent to the equalities
\[
\begin{align*}
f_1(\tilde{x}^k) + \beta[g_1(\tilde{x}^k) - g_1(x^k) - A^\top(\tilde{\lambda}^k - \lambda^k)] - \xi^k_{\tilde{x}} &= P_X[f_1(\tilde{x}^k) - \beta(g_1(x^k) - A^\top \lambda^k) - \xi^k_{\tilde{x}}], \\
f_2(\tilde{y}^k) + \beta[g_2(\tilde{y}^k) - g_2(y^k) - B^\top(\tilde{\lambda}^k - \lambda^k)] - \xi^k_{\tilde{y}} &= P_Y[f_2(\tilde{y}^k) - \beta(g_2(y^k) - B^\top \lambda^k) - \xi^k_{\tilde{y}}],
\end{align*}
\]
\[
\tilde{\lambda}^k = \lambda^k - \beta[Af_1(\tilde{x}^k) + Bf_2(\tilde{y}^k) - b],
\]
with $d(w^k, \tilde{w}^k, \xi^k)$.
Remark 3.2. The correction step (12) of Algorithm 3.1 consists of the following computation:

\[
\begin{align*}
\beta g_1(x^{k+1}) &= \beta g_1(x^k) - \beta A^\top (\lambda^k - \lambda^{k+1}) \\
&\quad - \alpha_k[\beta(g_1(x^k) - g_1(\tilde{x}^k) - A^\top (\lambda^k - \tilde{\lambda}^k)) + \xi^k], \\
\beta g_2(y^{k+1}) &= \beta g_2(y^k) - \beta B^\top (\lambda^k - \lambda^{k+1}) \\
&\quad - \alpha_k[\beta(g_2(y^k) - g_2(\tilde{y}^k) - B^\top (\lambda^k - \tilde{\lambda}^k)) + \xi^k].
\end{align*}
\]

(17)

\[
\lambda^{k+1} = \lambda^k - \alpha_k(\lambda^k - \tilde{\lambda}^k).
\]

Lemma 3.1. The GVI(10) is uniformly strongly monotone GVI under the assumption that \( f_i \) is \( g_i \)-monotone, \( i = 1, 2 \).

\[ f_1(x^*) \in \mathcal{X}, \] 

\[ f_2(x^*) \in \mathcal{Y}, \] 

\[ f_1(x^*) - f_2(x^*) \geq 0. \]

Proof. See Lemma 3.1 of [9].

Lemma 3.2. Let \( \{\tilde{w}^k\} \) be the sequence generated by Algorithm 3.1. Then, for any \( w^* \in \mathcal{W}^* \), we have

\[
[R(w^k) - R(w^*)]^\top d(w^k, \tilde{w}^k, \xi^k) \geq [R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi^k). \quad (18)
\]

Proof. From the definition of \( R \) (see (6)) and (16), we have

\[
R_x(w^k) - R_x(\tilde{w}^k) = \beta[(g_1(x^k) - g_1(\tilde{x}^k)) - A^\top (\lambda^k - \tilde{\lambda}^k)] - f_1(\tilde{x}^k) - \xi^k - P_x[f_1(\tilde{x}^k) - \beta(g_1(x^k) - A^\top \lambda^k) - \xi^k]. \quad (19)
\]

By (2) and \( P_x[f_1(\tilde{x}^k) - \beta(g_1(x^k) - A^\top \lambda^k) - \xi^k] \in \mathcal{X} \), we have

\[
\{P_x[f_1(\tilde{x}^k) - \beta(g_1(x^k) - A^\top \lambda^k) - \xi^k] - f_1(x^*)\}^\top \beta(g_1(x^*) - A^\top \lambda^*) \geq 0. \quad (20)
\]

Because \( f_1(x^*) \in \mathcal{X} \), it follows from (3) that

\[
\{P_x[f_1(\tilde{x}^k) - \beta(g_1(x^k) - A^\top \lambda^k) - \xi^k] - f_1(x^*)\}^\top [(f_1(\tilde{x}^k) - \beta(g_1(x^k) - A^\top \lambda^k) - \xi^k) - f_1(x^*)] \geq 0.
\]

(21)

Adding (20) and (21), and using (19), we obtain

\[
[R_x(\tilde{w}^k) - R_x(w^k) + f_1(\tilde{x}^k) - f_1(x^*) - \xi^k]^\top \\
\{R_x(w^k) - R_x(\tilde{w}^k) - \beta[g_1(x^k) - g_1(\tilde{x}^k) - A^\top (\lambda^k - \lambda^*)]\} \geq 0.
\]

Therefore, from (6), we have

\[
[R_x(\tilde{w}^k) - R_x(w^k) + f_1(\tilde{x}^k) - f_1(x^*) - \xi^k]^\top \\
\{\beta[g_1(x^*) - g_1(\tilde{x}^k) - A^\top (\lambda^* - \tilde{\lambda}^k)]\} \geq 0. \quad (22)
\]
Similarly, we have

\[
[R_y(\tilde{\omega}^k) - R_y(w^k) + f_2(\tilde{y}^k) - f_2(y^*) - \xi_y^k]^\top
\{\beta[g_2(y^*) - g_2(\tilde{y}^k) - B^\top(\lambda^* - \tilde{\lambda}^k)]\} \geq 0.
\]  

(23)

Combining (22) and (23), and using the third equality of (16), we have

\[
\begin{pmatrix}
\beta[g_1(x^*) - g_1(\tilde{x}^k) - A^\top(\lambda^* - \tilde{\lambda}^k)] \\
\beta[g_2(y^*) - g_2(\tilde{y}^k) - B^\top(\lambda^* - \tilde{\lambda}^k)] \\
\lambda^k - \tilde{\lambda}^k - \beta[Af_1(\tilde{x}^k) + Bf_2(\tilde{y}^k) - b]
\end{pmatrix} \geq 0.
\]  

(24)

Then using Assumption (A) and the definition of \(R\), we have

\[
[R(\tilde{\omega}^k) - R(w^*)]^\top [R(w^k) - R(\tilde{\omega}^k) + \xi^k] \geq 0.
\]  

(25)

Assertion (18) follows from (13) and (25) immediately.

**Remark 3.3.** Based on Lemma 3.2 and (11), we have

\[
[R(w^k) - R(w^*)]^\top d(w^k, \tilde{\omega}^k, \xi^k)
\]

\[
\geq [R(w^k) - R(\tilde{\omega}^k)]^\top [R(w^k) - R(\tilde{\omega}^k) + \xi^k]
\]

\[
\geq \|R(w^k) - R(\tilde{\omega}^k)\|^2 - \|R(w^k) - R(\tilde{\omega}^k)\| \cdot \|\xi_k\|
\]

\[
\geq (1 - \delta)\|R(w^k) - R(\tilde{\omega}^k)\|^2.
\]

This inequality and \(\delta \in (0, 1)\) imply that \(-d(w^k, \tilde{\omega}^k, \xi^k)\) is a descent direction of the unknown function \(\|R(w^k) - R(w^*)\|^2/2\) at the point \(w^k\).

Now, we explain the reason for determining the stepsize \(\alpha_k^*\) in the way of (15) for Algorithm 3.1. We use the quantity

\[
\|R(w^k) - R(w^*)\|^2 - \|R(w^{k+1}) - R(w^*)\|^2
\]

to measure the progress made by the new iterate. To identify an appropriate stepsize along the direction \(-d(w^k, \tilde{\omega}^k, \xi^k)\), we denote by \(w^{k+1}(\alpha)\) the new iterate generated by solving the following general scheme with undetermined stepsize \(\alpha > 0\):

\[
R(w^{k+1}(\alpha)) = R(w^k) - \alpha d(w^k, \tilde{\omega}^k, \xi^k).
\]  

(26)

Then, the following theorem provides the reason behind the choice (15).
Theorem 3.1. Let \( w^* \in \mathcal{W}^* \), \( w^{k+1}(\alpha) \) is given by the general scheme (26) and \( \tilde{w}^k \) be generated by (10). Let

\[
\theta^k(\alpha) := \| R(w^k) - R(w^*) \|^2 - \| R(w^{k+1}) - R(w^*) \|^2
\]

and

\[
q^k(\alpha) := 2\alpha [R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi^k) - \alpha^2\|d(w^k, \tilde{w}^k, \xi^k)\|^2.
\]

Then, we have

\[
\theta^k(\alpha) \geq q^k(\alpha).
\]

Proof. Using (12) and (18), we get

\[
\theta^k(\alpha) = 2\alpha [R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi^k) - \alpha^2\|d(w^k, \tilde{w}^k, \xi^k)\|^2 \\
\geq 2\alpha [R(w^k) - R(w^*)]^\top d(w^k, \tilde{w}^k, \xi^k) - \alpha^2\|d(w^k, \tilde{w}^k, \xi^k)\|^2 \\
= q^k(\alpha).
\]

The proof is complete.

Theorem 3.1 suggests that \( q^k(\alpha) \) is a lower bound of the progress made by the iterate \( w^{k+1}(\alpha) \). Despite that it is impossible to identify the optimal value of \( \alpha \) to maximize \( \theta^k(\alpha) \), we are interested in finding the suboptimal value of \( \alpha \) that maximizes the lower bound \( q^k(\alpha) \). Note that \( q^k(\alpha) \) is a quadratic function and it reaches its maximum at

\[
\alpha^*_k = \frac{[R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi^k)}{\|d(w^k, \tilde{w}^k, \xi^k)\|^2},
\]

which justifies the choice in (15).

Theorem 3.2. The sequence \( \{\alpha^*_k\} \) generated by Algorithm 3.1 is bounded away from \( 1/2 \).

Proof. Notice that

\[
\alpha^*_k = \frac{[R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi^k)}{\|d(w^k, \tilde{w}^k, \xi^k)\|^2} \\
= \frac{\|R(w^k) - R(\tilde{w}^k)\|^2 + [R(w^k) - R(\tilde{w}^k)]^\top \xi^k}{\|d(w^k, \tilde{w}^k, \xi^k)\|^2}.
\]

Recall that \( \|\xi^k\| \leq \|R(w^k) - R(\tilde{w}^k)\| \). We have

\[
\|R(w^k) - R(\tilde{w}^k)\|^2 + [R(w^k) - R(\tilde{w}^k)]^\top \xi^k
\]
\[
\begin{align*}
\geq \frac{1}{2} \| R(w^k) - R(\tilde{w}^k) \|^2 + [R(w^k) - R(\tilde{w}^k)]^\top \xi_k + \frac{1}{2} \| \xi_k \|^2 \\
= \frac{1}{2} \| R(w^k) - R(\tilde{w}^k) + \xi_k \|^2.
\end{align*}
\]

Then, it follows from the above inequalities that \( \alpha_k^* \geq 1/2 \). This complete the proof.

The parameter \( \gamma \) in (14) plays the role of a relaxation or scaling parameter. We can easily prove that \( \gamma \in (0, 2) \) can ensure convergence.

**Theorem 3.3.** Let \( \{w^k\} \) be the sequence generated by Algorithm 3.1. Then, for any \( w^* \in W^* \), we have

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

**Proof.** From (27) and (29), we have

\[
\| R(w^k) - R(w^*) \|^2 - \| R(w^{k+1}) - R(w^*) \|^2 \geq q^k(\alpha_k).
\]

Using (14)-(15) and (28), we have

\[
q^k(\alpha_k) = 2\alpha_k [R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi_k) - \alpha_k^2 \| d(w^k, \tilde{w}^k, \xi_k) \|^2 \\
= 2\gamma \alpha_k^* [R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi_k) - \gamma^2 \alpha_k^* [R(w^k) - R(\tilde{w}^k)]^\top d(w^k, \tilde{w}^k, \xi_k) \\
\geq \frac{\gamma(2-\gamma)(1-\delta)}{2} \| R(w^k) - R(\tilde{w}^k) \|^2.
\]

The assertion (31) follows from the above inequalities, and this complete the proof.

**Theorem 3.4.** The sequence \( \{w^k\} \) generated by Algorithm 3.1 converges to a point in \( W^* \).

**Proof.** The proof is similar to that of Theorem 2.3 of [8], thus is omitted.

### 4. Conclusions

In this paper, we presented a new IPPA method for structured general variational inequalities. Each iteration of the method consists of solving a proximally regularized subproblem, and this subproblem is allowed to be solved approximately subject to some inexactness criterion. Global convergence of the new method is proved under mild assumptions.
References


