Convergence of Adaptive Forward-Reflected-Backward Algorithm for Solving Variational Inequalities

Sergey Denisov and Vladimir Semenov

Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska Street, Kyiv, 01161, Ukraine

Abstract

The article deals with the problem of the approximate solution of variational inequalities. A novel iterative algorithm for solving variational inequalities in a real Banach space is proposed and studied. The proposed algorithm is an adaptive variant of the forward-reflected-backward algorithm (Malitsky, Tam, 2020), where the used rule for updating the step size does not require knowledge of the Lipschitz continuous constant of the operator. In addition, the Alber generalized projection is used instead of the metric projection onto the feasible set. For variational inequalities with monotone and Lipschitz continuous operators, acting in a 2-uniformly convex and uniformly smooth Banach space, a theorem on the weak convergence of the method is proved.

Keywords [1]

Variational inequality, monotone operator, Lipschitz continuous operator, forward-reflected-backward algorithm, 2-uniformly convex Banach space, uniformly smooth Banach space, convergence

1. Introduction

Many problems of operations research and mathematical physics can be written in the form of variational inequalities [1–5]. The development and study of variational inequalities is an actively developing area of applied nonlinear analysis [4, 6–23, 25–32]. Note that often non-smooth optimization problems can be effectively solved if they are reformulated as saddle point problems and algorithms for solving variational inequalities are applied [7]. With the advent of generative adversarial networks (GANs), a steady interest in algorithms for solving variational inequalities arose among specialists in the field of machine learning [8–10].

The classical variational inequality problem (in real Hilbert space H) has the form

$$\text{find } x\in C:\ \left(Ax, y-x\right)\geq 0 \quad \forall\, y\in C,$$

where $C\subseteq H$ is convex and closed, operator $A:C\to H$ is monotone, Lipschitz continuous. The most famous method for solving variational inequalities is the Korpelevich extra-gradient algorithm [11]

$$\begin{cases} x_1\in C,\ \lambda>0, \\ y_n = P_C\left(x_n-\lambda Ax_n\right), \\ x_{n+1} = P_X\left(x_n-\lambda Ay_n\right), \end{cases}$$

where P_C is metric projection onto C. Many publications are devoted to the study of the extra-gradient algorithm and its modifications [6, 7, 12–23]. An efficient modern version of the extra-gradient method is the proximal mirror method of Nemirovski [7]. This method can be interpreted as

---

II International Scientific Symposium «Intelligent Solutions» IntSol-2021, September 28–30, 2021, Kyiv-Uzhhorod, Ukraine
EMAIL: denisov.univ@gmail.com (S. Denisov); semenov.volodya@gmail.com (V. Semenov)
ORCID: 0000-0003-4397-4617 (S. Denisov); 0000-0002-3280-8245 (V. Semenov)

a variant of the extra-gradient method with projection understood in the sense of Bregman divergence [24]. Also, an interesting method of dual extrapolation for solving variational inequalities was proposed by Yu. Nesterov [25]. Adaptive variants of the Nemirovski mirror-prox method were studied in [19–23].

In the early 1980s, L. D. Popov proposed an interesting modification of the classical Arrow-Hurwitz algorithm for finding saddle points of convex-concave functions [26]. Let X and Y are closed convex subset of spaces R^d and R^p, respectively, and $L: X \times Y \to R$ be a differentiable convex-concave function. Then, the algorithm [26] approximation of saddle points of L on $X \times Y$ can be written as

$$\begin{cases} x_0, x_1 \in X,\ y_0, y_1 \in Y,\ \lambda > 0, \\ x_n = P_X\left(x_n - \lambda \nabla_1 L\left(x_{n-1}, y_{n-1}\right)\right), \\ y_n = P_Y\left(y_n + \lambda \nabla_2 L\left(x_{n-1}, y_{n-1}\right)\right), \\ x_{n+1} = P_X\left(x_n - \lambda \nabla_1 L\left(x_n, y_n\right)\right), \\ y_{n+1} = P_Y\left(y_n + \lambda \nabla_2 L\left(x_n, y_n\right)\right), \end{cases}$$

where P_X and P_Y are metric projection onto X and Y, respectively, $\nabla_1 L$ and $\nabla_2 L$ are partial derivatives. Under some suitable assumptions, L. D. Popov proved the convergence of this algorithm.

A modification of Popov's method for solving variational inequalities with monotone operators was studied in [27]. And in the article [27], a two-stage proximal algorithm for solving the equilibrium programming problem is proposed, which is an adaptation of the method [26] to the general Ky Fan inequalities. The mentioned equilibrium problem (Ky Fan inequality) has the form

$$\text{find } x \in C: \ F\left(x, y\right) \geq 0 \quad \forall\, y \in C,$$

where C is nonempty subset of vector space H (usually Hilbert space), $F: C \times C \to R$ is function such that $F\left(x, x\right) = 0 \ \ \forall\, x \in C$ (called bifunction). And the two-stage proximal algorithm is written like this

$$\begin{cases} y_n = \operatorname{prox}_{\lambda_n F\left(y_{n-1}, \cdot\right)} x_n, \\ x_{n+1} = \operatorname{prox}_{\lambda_n F\left(y_n, \cdot\right)} x_n, \end{cases}$$

where $\lambda_n \in \left(0, +\infty\right)$, $\operatorname{prox}_\varphi$ is proximal operator for function $\varphi: C \to R$ is defined by

$$\operatorname{prox}_\varphi x = \arg\min_{y \in C} \left(\varphi\left(y\right) + \tfrac{1}{2}\left\|y - x\right\|^2\right).$$

In [28, 29], the two-stage proximal mirror method was studied, which is a modification of the two-stage proximal algorithm [27] using Bregman divergence instead of the Euclidean distance. Note that recently Popov's algorithm for variational inequalities has become well known among machine learning specialists under the name “Extrapolation from the Past” [9]. Further development of this circle of ideas led to the emergence of the so-called forward-reflected-backward algorithm [30] and related methods [31, 32]. The forward-reflected-backward algorithm generates a sequence $\left(x_n\right)$ iteratively defined by

$$x_{n+1} = P_C\left(x_n - \lambda_n A x_n - \lambda_{n-1}\left(A x_n - A x_{n-1}\right)\right),$$

with $\left\{\inf_n \lambda_n, \sup_n \lambda_n\right\} \subseteq \left(0, \frac{1}{2L}\right)$, where L is the Lipschitz constant of A.

In this paper, we propose a novel algorithm for solving variational inequalities in a Banach space. Variational inequalities in Banach spaces arise and are intensively studied in mathematical physics and the theory of inverse problems [1, 2, 4]. Recently, there has been progress in the study of algorithms for problems in Banach spaces [4, 15–18]. This is due to the wide involvement of the results and constructions of the geometry of Banach spaces [33–35]. The proposed algorithm is an adaptive variant of the forward-reflected-backward algorithm [30], where the rule for updating the step size does not require knowledge of the Lipschitz constant of operator. Moreover, instead of the metric projection onto the feasible set, the Alber generalized projection is used [35]. An attractive feature of the algorithm is only one computation at the iterative step of the projection onto the feasible set. For variational inequalities with monotone Lipschitz operators acting in a 2-uniformly convex and uniformly smooth Banach space, a theorem on the weak convergence of the method is proved.

2. Preliminaries

We recall several concepts and facts of the geometry of Banach spaces that are necessary for the formulation and proof of the results [33–37].

Everywhere E denotes a real Banach space with the norm $\|\cdot\|$, E^* dual to E space, $\langle x^*, x\rangle$ is value of functional $x^* \in E^*$ on element $x \in E$. We denote norm in E^* as $\|\cdot\|_*$.

Let $S_E = \{x \in E : \|x\| = 1\}$. Banach space E is strictly convex if for all $x, y \in S_E$ and $x \neq y$ we have

$$\left\| \frac{x+y}{2} \right\| < 1.$$

The modulus of convexity of the space E is defined as follows

$$\delta_E(\varepsilon) = \inf\left\{1 - \left\|\frac{x+y}{2}\right\| : \ x, y \in B_E, \ \|x - y\| = \varepsilon\right\} \quad \forall \varepsilon \in (0, 2].$$

Banach space E is uniformly convex if $\delta_E(\varepsilon) > 0$ for all $\varepsilon \in (0, 2]$. Banach space E is called 2-uniformly convex if exists $c > 0$ that

$$\delta_E(\varepsilon) \geq c\varepsilon^2$$

for all $\varepsilon \in (0, 2]$. Obviously, a 2-uniformly convex space is uniformly convex. It is known that a uniformly convex Banach space is reflexive.

A Banach space E is called smooth if the limit

$$\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} \tag{1}$$

exists for all $x, y \in S_E$. A Banach space E is called uniformly smooth if the limit (1) exists uniformly in $x, y \in S_E$. There is a duality between the convexity and smoothness of the Banach space E and its dual E^* [33, 34]:

- E^* is strictly convex space $\Rightarrow$ E is smooth space;
- E^* is smooth space $\Rightarrow$ E is strictly convex space;
- E is uniformly convex space $\Leftrightarrow$ E^* is uniformly smooth space;
- E is uniformly smooth space $\Leftrightarrow$ E^* is uniformly convex space.

Note that if the space E is reflexive, the first two implications can be reversed. It is known that Hilbert spaces and spaces L_p ($1 < p \leq 2$) are 2-uniformly convex and uniformly smooth (spaces L_p are uniformly smooth for $p \in (1, \infty)$) [33, 34].

Multivalued operator $J : E \to 2^{E^*}$, acting as follows

$$Jx = \left\{x^* \in E^* : \ \langle x^*, x\rangle = \|x\|^2 = \|x^*\|_*^2\right\},$$

is called the normalized duality mapping. It is known that [36]:

- if the space E is smooth, then the mapping J is single valued;
- if the space E is strictly convex, then the mapping J is injective and strictly monotone;
- if the space E is reflexive, then the mapping J is surjective;
- if the space E is uniformly smooth, then the mapping J is uniformly continuous on bounded subsets of E.

Let E be a smooth Banach space. Consider the functional introduced by Yakov Alber [35]

$$\phi(x,y)=\|x\|^2-2\langle Jy,x\rangle+\|y\|^2 \quad \forall x,y\in E .$$

A useful identity follows from the definition of ϕ:

$$\phi(x,y)-\phi(x,z)-\phi(z,y)=2\langle Jz-Jy,x-z\rangle \quad \forall x,y,z\in E .$$

If the space E is strictly convex, then for $x,y\in E$ we have $\phi(x,y)=0 \Leftrightarrow x=y$.

Lemma 1 ([35]). Let E be a uniformly convex and uniformly smooth Banach space, (x_n) and (y_n) are bounded sequences of E elements. Then

$$\|x_n-y_n\|\to 0 \quad \Leftrightarrow \quad \|Jx_n-Jy_n\|_*\to 0 \quad \Leftrightarrow \quad \phi(x_n,y_n)\to 0.$$

Lemma 2 ([37]). Let E be a 2-uniformly convex and smooth Banach space. Then, for some number $\mu\ge 1$, the inequality holds

$$\phi(x,y)\ge\frac{1}{\mu}\|x-y\|^2 \quad \forall x,y\in E .$$

Let K be a non-empty closed and convex subset of a reflexive, strictly convex and smooth space E. It is known [35] that for each $x\in E$ there is a unique point $z\in K$ such that

$$\phi(z,x)=\inf_{y\in K}\phi(y,x).$$

This point z is denoted by $\Pi_K x$, and the corresponding operator $\Pi_K:E\to K$ is called the generalized projection of E onto K (Alber generalized projection) [35]. Note that if E is a Hilbert space, then Π_K coincides with the metric projection onto the set K.

Lemma 3 ([35]). Let K be a closed and convex subset of a reflexive, strictly convex and smooth space E, $x\in E$, $z\in K$. Then

$$z=\Pi_K x \quad \Leftrightarrow \quad \langle Jz-Jx,y-z\rangle\ge 0 \quad \forall y\in K . \tag{2}$$

Remark 1. The inequality (2) is equivalent to the following [35]:

$$\phi(y,\Pi_K x)+\phi(\Pi_K x,x)\le\phi(y,x) \quad \forall y\in K .$$

Basic information about monotone operators and variational inequalities in Banach spaces can be found in [1, 2, 4, 35, 36]. We mention only two interesting examples of monotone operators acting in a Banach space [4].

For $p\ge 2$, define the operator A by

$$Au=|u(x)|^{p-2}u(x)\int_{R^3}\frac{|u(y)|^p}{\|x-y\|_2}dy .$$

The operator A is potential and monotone, and acts from $L_p(R^3)$ to $L_q(R^3)$, where $p^{-1}+q^{-1}=1$. Note that A is the gradient of the functional

$$F(u)=\frac{1}{2p}\int_{R^3}\int_{R^3}\frac{|u(x)|^p|u(y)|^p}{\|x-y\|_2}\,dxdy.$$

Let $G\subseteq R^n$ be a bounded domain. Differential expression

$$Au=-\sum_{i=1}^{n}\tfrac{\partial}{\partial x_i}\left(a_i\left(x,\left|\tfrac{\partial u}{\partial x_i}\right|^{p-1}\right)\left|\tfrac{\partial u}{\partial x_i}\right|^{p-2}\tfrac{\partial u}{\partial x_i}\right)+a_0\left(x,|u|^{p-1}\right)|u|^{p-2}u,\ p>1,$$

where the function $a_i(x,s)$, $i=0,1,...,n$, is measurable as a function on x for every $s\in[0,+\infty)$ and continuous for almost all $x\in G$ as a function on s, $|a_i(x,s)|\le M$ for all $s\in[0,+\infty)$ and for almost all $x\in G$, specifies a monotone operator acting from Sobolev space $W_{0,p}^1(G)$ to $\left(W_{0,p}^1(G)\right)^*$.

3. Algorithm

Let E be 2-uniformly convex and uniformly smooth Banach space, C be non-empty subset of space E, A be an operator from E to E^*. Consider variational inequality:

$$\text{find } x\in C:\ \langle Ax,y-x\rangle\ge 0\quad\forall y\in C. \tag{3}$$

We denote set of solutions of (3) by S.

Assume that the following conditions are satisfied:

- set $C\subseteq E$ is convex and closed;
- operator $A:E\to E^*$ is monotone and Lipschitz -type with $L>0$ on C;
- set S is non-empty.

Remark 2. We can formulate (3) as fixed-point problem [35]:

$$x=\Pi_C J^{-1}(Jx-\lambda Ax), \tag{4}$$

where $\lambda>0$. Formulation (4) is useful because it contains an obvious algorithmic idea.

Consider dual variational inequality:

$$\text{find } x\in C:\ \langle Ay,x-y\rangle\le 0\quad\forall y\in C. \tag{5}$$

We denote set of solutions of (5) by S^d. Note that set S^d is closed and convex [2]. Inequality (5) is sometimes called weak or dual formulation of (3) (or Minty inequality) and solutions (5) are weak solutions (3). For monotone operators A we always have $S\subseteq S^d$. In our conditions $S^d=S$ [2].

We assume that the following is satisfied:

- normalized duality mapping $J:E\to E^*$ sequentially weakly continuous, i.e., from $x_n\to x$ weak in E then $Jx_n\to Jx$ weak* in E^*.

Remark 3. In our situation, when the space E (and of course E^*) is reflexive, the weak* and weak convergence coincide in E^*.

Consider now a novel algorithm for solving the variational inequality (3). We will use a simple rule for updating the parameters λ_n without information about the Lipschitz constant of the operator A. The proposed algorithm is a modification of the forward-reflected-backward algorithm recently proposed in [30] for solving operator inclusions with the sum of the maximal monotone and Lipschitz continuous monotone operators acting in a Hilbert space.

Let us know the constant $\mu\ge 1$ from the Lemma 2.

Algorithm 1.

Initialization. Choose $x_0 \in E$, $x_1 \in E$, $\tau \in \left(0, \frac{1}{2\mu}\right)$ and $\lambda_0, \lambda_1 > 0$. Let $n = 1$.

1. Calculate

$$x_{n+1} = \Pi_C J^{-1}\left(Jx_n - \lambda_n Ax_n - \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right)\right).$$

2. If $x_{n-1} = x_n = x_{n+1}$, then STOP and $x_n \in S$, else go to **3**.

3. Calculate

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau \dfrac{\|x_{n+1} - x_n\|}{\|Ax_{n+1} - Ax_n\|_*}\right\}, & \text{if } Ax_{n+1} \neq Ax_n, \\ \lambda_n, & \text{otherwise.} \end{cases}$$

Let $n := n+1$ and go to **1**.

Sequence generated by rule of calculation $\left(\lambda_n\right)$ is non-increasing and lower bounded by $\min\left\{\lambda_1, \tau L^{-1}\right\}$. Then exists $\lim_{n\to\infty} \lambda_n > 0$.

The sequence $\left(x_n\right)$ generated by Algorithm 1 satisfies the inequality

$$-2\left\langle \lambda_n Ax_n + \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), y - x_{n+1}\right\rangle \leq \phi\left(y, x_n\right) - \phi\left(x_{n+1}, x_n\right) - \phi\left(y, x_{n+1}\right) \quad \forall y \in C. \tag{6}$$

Inequality (6) shows a rule of finishing the algorithm. Indeed if

$$x_{n-1} = x_n = x_{n+1}$$

then from (6) it follows then

$$\left\langle Ax_n, y - x_n\right\rangle \geq 0$$

for all $y \in C$, i.e., $x_n \in S$.

Now we go to the proof of convergence of Algorithm 1.

4. Main inequality

In this section, we state and prove the inequality on which the proof of Algorithm 1 weak convergence is based.

Lemma 4. For the sequence $\left(x_n\right)$ generated by Algorithm 1, the following inequality holds

$$\begin{aligned} &\phi\left(z, x_{n+1}\right) + 2\lambda_n \left\langle Ax_n - Ax_{n+1}, x_{n+1} - z\right\rangle + \tau\mu \frac{\lambda_n}{\lambda_{n+1}} \phi\left(x_{n+1}, x_n\right) \leq \\ &\leq \phi\left(z, x_n\right) + 2\lambda_{n-1}\left\langle Ax_{n-1} - Ax_n, x_n - z\right\rangle + \tau\mu \frac{\lambda_{n-1}}{\lambda_n}\phi\left(x_n, x_{n-1}\right) - \\ &\quad - \left(1 - \tau\mu\frac{\lambda_{n-1}}{\lambda_n} - \tau\mu\frac{\lambda_n}{\lambda_{n+1}}\right)\phi\left(x_{n+1}, x_n\right), \end{aligned}$$

where $z \in S$.

Proof. Let $z \in S$. We have

$$\phi\left(z, x_{n+1}\right) \leq \phi\left(z, x_n\right) - \phi\left(x_{n+1}, x_n\right) + 2\left\langle \lambda_n Ax_n + \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), z - x_{n+1}\right\rangle. \tag{7}$$

From monotonicity of operator A we have

$$\left\langle \lambda_n Ax_n + \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), z - x_{n+1}\right\rangle = \lambda_n \left\langle Ax_n - Ax_{n+1}, z - x_{n+1}\right\rangle +$$

$$+\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, z - x_{n+1}\right\rangle + \underbrace{\lambda_n \left\langle Ax_{n+1}, z - x_{n+1}\right\rangle}_{\le 0} \le$$

$$\le \lambda_n \left\langle Ax_n - Ax_{n+1}, z - x_{n+1}\right\rangle + \lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, z - x_n\right\rangle +$$

$$+\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\right\rangle. \quad (8)$$

Applying (8) to (7) we obtain

$$\phi\left(z, x_{n+1}\right) \le \phi\left(z, x_n\right) - \phi\left(x_{n+1}, x_n\right) + 2\lambda_n \left\langle Ax_n - Ax_{n+1}, z - x_{n+1}\right\rangle +$$

$$+2\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, z - x_n\right\rangle + 2\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\right\rangle. \quad (9)$$

From rule of calculation λ_n we have upper estimation for $2\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\right\rangle$ in (9). We have

$$2\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\right\rangle \le$$

$$\le 2\lambda_{n-1}\left\|Ax_n - Ax_{n-1}\right\|_* \left\|x_n - x_{n+1}\right\| \le 2\tau \frac{\lambda_{n-1}}{\lambda_n}\left\|x_n - x_{n-1}\right\|\left\|x_{n+1} - x_n\right\| \le$$

$$\le \tau \frac{\lambda_{n-1}}{\lambda_n}\left\|x_n - x_{n-1}\right\|^2 + \tau \frac{\lambda_{n-1}}{\lambda_n}\left\|x_n - x_{n+1}\right\|^2 \le$$

$$\le \tau\mu \frac{\lambda_{n-1}}{\lambda_n}\phi\left(x_n, x_{n-1}\right) + \tau\mu \frac{\lambda_{n-1}}{\lambda_n}\phi\left(x_{n+1}, x_n\right).$$

We obtain

$$\phi\left(z, x_{n+1}\right) + 2\lambda_n \left\langle Ax_n - Ax_{n+1}, x_{n+1} - z\right\rangle + \tau\mu \frac{\lambda_n}{\lambda_{n+1}}\phi\left(x_{n+1}, x_n\right) \le$$

$$\le \phi\left(z, x_n\right) + 2\lambda_{n-1}\left\langle Ax_{n-1} - Ax_n, x_n - z\right\rangle + \tau\mu \frac{\lambda_{n-1}}{\lambda_n}\phi\left(x_n, x_{n-1}\right) -$$

$$-\left(1 - \tau\mu \frac{\lambda_{n-1}}{\lambda_n} - \tau\mu \frac{\lambda_n}{\lambda_{n+1}}\right)\phi\left(x_{n+1}, x_n\right).$$

The proof is complete. ■

Remark 4. We can change rule of updating for step 3 of Algorithm 1 to the following:

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau \dfrac{\sqrt{\mu\, \phi\left(x_{n+1}, x_n\right)}}{\left\|Ax_{n+1} - Ax_n\right\|_*}\right\}, & \text{if} \quad Ax_{n+1} \ne Ax_n, \\ \lambda_n, & \text{otherwise.} \end{cases} \quad (10)$$

Lemma 4 holds also for variant of Algorithm 1 with the rule (10).

5. Convergence

Let us formulate the main result.

Theorem 1. Let C be a non-empty convex and closed subset of 2-uniformly convex and uniformly smooth Banach space E, $A: E \to E^*$ is monotone Lipschitz continuous operator, $S \neq \varnothing$. Assume that normalized duality mapping J is sequentially weakly continuous. Then sequence generated by Algorithm 1 (x_n) converge weakly to $z \in S$.

Proof. Let $z' \in S$. Assume

$$a_n = \phi(z', x_n) + 2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z'\rangle + \tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_n, x_{n-1}),$$

$$b_n = \left(1 - \tau\mu\frac{\lambda_{n-1}}{\lambda_n} - \tau\mu\frac{\lambda_n}{\lambda_{n+1}}\right)\phi(x_{n+1}, x_n).$$

Inequality from Lemma 4 takes form

$$a_{n+1} \leq a_n - b_n.$$

Since there exists $\lim\limits_{n\to\infty}\lambda_n > 0$, then

$$1 - \tau\mu\frac{\lambda_{n-1}}{\lambda_n} - \tau\mu\frac{\lambda_n}{\lambda_{n+1}} \to 1 - 2\tau\mu \in (0,1), \quad n \to \infty.$$

Show that $a_n \geq 0$ for all large $n \in N$. We have

$$a_n = \phi(z', x_n) + 2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z'\rangle + \tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_n, x_{n-1}) \geq$$

$$\geq \frac{1}{\mu}\|x_n - z'\|^2 - 2\lambda_{n-1}\|Ax_{n-1} - Ax_n\|_*\|x_n - z'\| + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_{n-1} - x_n\|^2 \geq$$

$$\geq \frac{1}{\mu}\|x_n - z'\|^2 - 2\tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|\|x_n - z'\| + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_{n-1} - x_n\|^2 \geq \left(\frac{1}{\mu} - \tau\frac{\lambda_{n-1}}{\lambda_n}\right)\|x_n - z'\|^2.$$

Since there exists such $n_0 \in N$ that

$$\frac{1}{\mu} - \tau\frac{\lambda_{n-1}}{\lambda_n} > 0 \text{ for all } n \geq n_0,$$

then $a_n \geq 0$ starting from n_0.

So, we came into conclusion that there exists a limit

$$\lim_{n\to\infty}\left(\phi(z', x_n) + 2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z'\rangle + \tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_n, x_{n-1})\right)$$

and

$$\sum_{n=1}^{\infty}\left(1 - \tau\mu\frac{\lambda_{n-1}}{\lambda_n} - \tau\mu\frac{\lambda_n}{\lambda_{n+1}}\right)\phi(x_{n+1}, x_n) < +\infty.$$

Hence, we obtain that the sequence (x_n) is bounded and

$$\lim_{n\to\infty}\phi(x_{n+1}, x_n) = \lim_{n\to\infty}\|x_{n+1} - x_n\| = 0.$$

Since

$$\lim_{n\to\infty}\left(2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z'\rangle + \tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_n, x_{n-1})\right) = 0,$$

then sequences $(\phi(z', x_n))$ converge for all $z' \in S$.

Show that all weak cluster points of sequence (x_n) are in the set S. Consider subsequence (x_{n_k}) which converges weakly to $z \in E$. Easy to see that $z \in C$. Show that $z \in S$. We have

$$\left\langle Jx_{n+1} - Jx_n + \lambda_n Ax_n + \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), y - x_{n+1}\right\rangle \geq 0 \quad \forall y \in C.$$

Hence using monotonicity of operator A we have an inequality

$$\langle Ay, y - x_n\rangle + \langle Ax_n, x_n - x_{n+1}\rangle \geq \langle Ax_n, y - x_{n+1}\rangle \geq$$

$$\geq \frac{1}{\lambda_n}\langle Jx_n - Jx_{n+1}, y - x_{n+1}\rangle - \frac{\lambda_{n-1}}{\lambda_n}\langle Ax_n - Ax_{n-1}, y - x_{n+1}\rangle \qquad \forall y \in C.$$

From $\lim_{n\to\infty}\|x_n - x_{n-1}\| = 0$ and Lipschitz property of operator A it follows

$$\lim_{n\to\infty}\|Ax_n - Ax_{n-1}\|_* = 0.$$

From uniform continuity of normalized duality mapping J on bounded sets we get

$$\lim_{n\to\infty}\|Jx_n - Jx_{n+1}\|_* = 0.$$

Hence,

$$\varliminf_{n\to\infty}\langle Ay, y - x_n\rangle \geq 0 \qquad \forall y \in C.$$

From other side

$$\langle Ay, y - z\rangle = \lim_{k\to\infty}\left\langle Ay, y - x_{n_k}\right\rangle \geq \varliminf_{n\to\infty}\left\langle Ay, y - x_n\right\rangle \geq 0 \qquad \forall y \in C.$$

Then it follows that $z \in S$.

Show that sequence (x_n) converges weakly to z. Arguing by contradiction. Let exists the subsequence (x_{m_k}) such that $x_{m_k} \to z'$ weakly and $z \neq z'$. Easy to see that $z' \in S$. We have

$$2\langle Jx_n, z - z'\rangle = \phi(z', x_n) - \phi(z, x_n) + \|z\|^2 - \|z'\|^2.$$

From that we see the existence of limit $\lim_{n\to\infty}\langle Jx_n, z - z'\rangle$. From sequentially weak continuity of normalized duality mapping J we get

$$\langle Jz, z - z'\rangle = \lim_{k\to\infty}\left\langle Jx_{n_k}, z - z'\right\rangle = \lim_{k\to\infty}\left\langle Jx_{m_k}, z - z'\right\rangle = \langle Jz', z - z'\rangle,$$

i.e., $\langle Jz - Jz', z - z'\rangle = 0$. Then it follows that $z = z'$. The proof is complete. ■

The weak convergence of the variant of the algorithm with a constant parameter $\lambda > 0$ is similarly proved.

Algorithm 2.

Initialization. Choose $x_0 \in E$, $x_1 \in E$, $\lambda \in \left(0, \frac{1}{2\mu L}\right)$. Let $n = 1$.

1. Calculate

$$x_{n+1} = \Pi_C J^{-1}\left(Jx_n - 2\lambda Ax_n + \lambda Ax_{n-1}\right).$$

2. If $x_{n-1} = x_n = x_{n+1}$, then STOP and $x_n \in S$, else let $n := n+1$ and go to **1**.

Remark 5. A special case of Algorithm 2 is the optimistic gradient descent ascent (OGDA) algorithm, popular among machine learning specialists [8, 9].

Lemma 5. For the sequence (x_n) generated by Algorithm 2, the following inequality holds

$$\phi(z, x_{n+1}) + 2\lambda \langle Ax_n - Ax_{n+1}, x_{n+1} - z \rangle + \lambda\mu L\phi(x_{n+1}, x_n) \le$$
$$\le \phi(z, x_n) + 2\lambda \langle Ax_{n-1} - Ax_n, x_n - z \rangle + \lambda\mu L\phi(x_n, x_{n-1}) - (1 - 2\lambda\mu L)\phi(x_{n+1}, x_n),$$

where $z \in S$.

Theorem 2. Let C be a nonempty convex and closed subset of 2-uniformly convex and uniformly smooth Banach space E, operator $A : E \to E^*$ is monotone and Lipschitz continuous with constant $L > 0$. Let $S \ne \varnothing$ and normalized duality mapping J is sequentially weakly continuous. Then sequence generated by Algorithm 2 (x_n) converge weakly to $z \in S$.

6. Conclusions

In this paper, we have proposed and studied a new algorithm for solving variational inequalities in a Banach space. The proposed algorithm is an adaptive variant of the forward-reflected-backward algorithm [30], where the rule for updating the step size does not require knowledge of the Lipschitz continuous operator constant. Moreover, instead of the metric projection onto the admissible set, the Alber generalized projection is used [35]. An attractive feature of the algorithm is only one computation at the iterative step of the generalized projection onto the feasible set. For variational inequalities with monotone Lipschitz continuous operators acting in a 2-uniformly convex and uniformly smooth Banach space, a theorem on the weak convergence of the method is proved.

Based on the technique [38], similar results can most likely be obtained for problems with pseudo-monotone, Lipschitz continuous, and sequentially weakly continuous operators acting in a uniformly convex and uniformly smooth Banach space. Also, in a future article we will present a proof of the convergence of a modification of the algorithm using the Bregman projection.

Note that a problem of significant interest in nonlinear analysis applications is to find $x \in (A_1 + A_2)^{-1} 0$, where $A_1 : E \to 2^{E^*}$ is a maximal monotone operator and $A_2 : E \to E^*$ is a monotone and Lipschitz operator. Based on the results of this work and [30] for solving this problem, we can construct the following adaptive splitting method

$$x_{n+1} = (J + \lambda_n A_1)^{-1} \left(Jx_n - \lambda_n A_2 x_n - \lambda_{n-1} (A_2 x_n - A_2 x_{n-1}) \right),$$

$$\lambda_{n+1} = \begin{cases} \min\left\{ \lambda_n, \tau \dfrac{\|x_{n+1} - x_n\|}{\|A_2 x_{n+1} - A_2 x_n\|_*} \right\}, & \text{if } A_2 x_{n+1} \ne A_2 x_n, \\ \lambda_n, & \text{otherwise,} \end{cases}$$

where $\tau \in \left(0, \frac{1}{2\mu}\right)$. The proof of its convergence will be presented in another work soon.

Acknowledgements

This work was supported by Ministry of Education and Science of Ukraine (project "Mathematical modeling and optimization of dynamical systems for defense, medicine and ecology", 0119U100337).

References

[1] J.-L. Lions, Some Methods of Solving Non-Linear Boundary Value Problems, Dunod-Gauthier-Villars, Paris, 1969.

[2] D. Kinderlehrer, G. Stampacchia, An Introduction to Variational Inequalities and Their Applications, Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[3] A. Nagurney, Network economics: A variational inequality approach, Kluwer Academic Publishers, Dordrecht, 1999.

[4] Y. Alber, I. Ryazantseva, Nonlinear Ill Posed Problems of Monotone Type, Springer, Dordrecht, 2006.

[5] S. I. Lyashko, D. A. Klyushin, D. A. Nomirovsky, V. V. Semenov, Identification of age-structured contamination sources in ground water, in: R. Boucekkline, N. Hritonenko, Y. Yatsenko, Y. (Eds.), Optimal Control of Age-Structured Populations in Economy, Demography, and the Environment, Routledge, London–New York, 2013, pp. 277–292.

[6] F. Facchinei, J. S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Springer Series in Operations Research, vol. I, Springer, New York, 2003.

[7] A. Nemirovski, Prox-method with rate of convergence O(1/T) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems, SIAM J. Optim. 15 (2004) 229–251. doi:10.1137/S1052623403425629.

[8] C. Daskalakis, A. Ilyas, V. Syrgkanis, H. Zeng, Training GANs with optimism, arXiv preprint arXiv:1711.00141. (2018).

[9] G. Gidel, H. Berard, P. Vincent, S. Lacoste-Julien, A Variational Inequality Perspective on Generative Adversarial Networks, arXiv preprint arXiv:1802.10551. (2018).

[10] M. Liu, W. Zhang, Y. Mroueh, X. Cui, J. Ross, T. Yang, P. Das, A decentralized parallel algorithm for training generative adversarial nets, arXiv preprint arXiv:1910.12999. (2020).

[11] G. M. Korpelevich, An extragradient method for finding saddle points and for other problems, Matecon. 12 (1976) 747–756.

[12] Y. Censor, A. Gibali, S. Reich, The subgradient extragradient method for solving variational inequalities in Hilbert space, Journal of Optimization Theory and Applications 148 (2011) 318–335 (2011). doi:10.1007/s10957-010-9757-3.

[13] V. V. Semenov, Modified Extragradient Method with Bregman Divergence for Variational Inequalities, Journal of Automation and Information Sciences 50(8) (2018) 26–37. doi:10.1615/JAutomatInfScien.v50.i8.30.

[14] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, SIAM Journal on Control and Optimization 38 (2000) 431–446.

[15] Y. Shehu, Convergence Results of Forward-Backward Algorithms for Sum of Monotone Operators in Banach Spaces, Results Math. 74 (2019). doi:10.1007/s00025-019-1061-4.

[16] Y. Shehu, Single projection algorithm for variational inequalities in Banach spaces with application to contact problem, Acta Math. Sci. 40 (2020) 1045–1063. doi:10.1007/s10473-020-0412-2.

[17] J. Yang, P. Cholamjiak, P. Sunthrayuth, Modified Tseng's splitting algorithms for the sum of two monotone operators in Banach spaces, AIMS Mathematics 6(5) (2021) 4873–4900. doi:10.3934/math.2021286.

[18] P. Cholamjiak, Y. Shehu, Inertial forward-backward splitting method in Banach spaces with application to compressed sensing, Appl. Math. 64 (2019) 409–435.

[19] F. Bach, K. Y. Levy, A Universal Algorithm for Variational Inequalities Adaptive to Smoothness and Noise, arXiv preprint arXiv:1902.01637. (2019).

[20] K. Antonakopoulos, V. Belmega, P. Mertikopoulos, An adaptive mirror-prox method for variational inequalities with singular operators, in: Advances in Neural Information Processing Systems 32 (NeurIPS), Curran Associates, Inc., 2019, 8455–8465.

[21] F. S. Stonyakin, On the adaptive proximal method for a class of variational inequalities and related problems, Trudy Inst. Mat. i Mekh. UrO RAN. 25 (2019) 185–197. doi:10.21538/0134-4889-2019-25-2-185-197.
[22] F. S. Stonyakin, E. A. Vorontsova, M. S. Alkousa, New Version of Mirror Prox for Variational Inequalities with Adaptation to Inexactness, in: Jacimovic M., Khachay M., Malkova V., Posypkin M. (Eds.), Optimization and Applications, OPTIMA 2019, volume 1145 of Communications in Computer and Information Science, Springer, Cham, 2020, 427–442. doi:10.1007/978-3-030-38603-0_31.
[23] S. V. Denisov, V. V. Semenov, P. I. Stetsyuk, Bregman Extragradient Method with Monotone Rule of Step Adjustment, Cybernetics and Systems Analysis. 55 (2019) 377–383. doi:10.1007/s10559-019-00144-5.
[24] L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, USSR Computational Mathematics and Mathematical Physics 7(3) (1967) 200–217.
[25] Yu. Nesterov, Dual extrapolation and its applications to solving variational inequalities and related problems, Mathematical Programming 109 (2007), 319–344.
[26] L. D. Popov, A modification of the Arrow-Hurwicz method for search of saddle points, Mathematical notes of the Academy of Sciences of the USSR. 28 (1980) 845–848.
[27] L. Chabak, V. Semenov, Y. Vedel, A New Non-Euclidean Proximal Method for Equilibrium Problems, In: O. Chertov et al. (Eds.), Recent Developments in Data Science and Intelligent Analysis of Information, volume 836 of Advances in Intelligent Systems and Computing, Springer, Cham, 2019, pp. 50–58. doi:10.1007/978-3-319-97885-7_6.
[28] D. A. Nomirovskii, B. V. Rublyov, V. V. Semenov, Convergence of Two-Stage Method with Bregman Divergence for Solving Variational Inequalities, Cybernetics and Systems Analysis. 55 (2019) 359–368. doi:10.1007/s10559-019-00142-7.
[29] A. Gibali, D. V. Thong, A new low-cost double projection method for solving variational inequalities, Optim. Eng. 21 (2020) 1613–1634. doi:10.1007/s11081-020-09490-2.
[30] Y. Malitsky, M. K. Tam, A Forward-Backward Splitting Method for Monotone Inclusions Without Cocoercivity, SIAM Journal on Optimization 30 (2020) 1451–1472. doi:10.1137/18M1207260.
[31] E. R. Csetnek, Y. Malitsky, M. K. Tam, Shadow Douglas-Rachford Splitting for Monotone Inclusions, Appl Math Optim. 80 (2019) 665–678. doi:10.1007/s00245-019-09597-8.
[32] V. Cevher, B. C. Vu, A Reflected Forward-Backward Splitting Method for Monotone Inclusions Involving Lipschitzian Operators, Set-Valued and Variational Analysis 29 (2021) 163–174. doi:10.1007/s11228-020-00542-4.
[33] J. Diestel, Geometry of Banach Spaces, Springer-Verlag, Berlin–Heidelberg, 1975.
[34] B. Beauzamy, Introduction to Banach Spaces and Their Geometry, North-Holland, Amsterdam, 1985.
[35] Y.I. Alber, Metric and generalized projection operators in Banach spaces: properties and applications. in: Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, vol. 178, Dekker, New York, 1996, pp. 15–50.
[36] M. M. Vainberg, Variational Method and Method of Monotone Operators in the Theory of Nonlinear Equations, Wiley, New York, 1974.
[37] K. Aoyama, F. Kohsaka, Strongly relatively nonexpansive sequences generated by firmly nonexpansive-like mappings, Fixed Point Theory Appl. 95 (2014). doi:10.1186/1687-1812-2014-95.
[38] R.I. Bot, E.R. Csetnek, P.T. Vuong, The forward-backward-forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in Hilbert spaces, European Journal of Operational Research 287(1) (2020) 49–60.