The Regularized Operator Extrapolation Algorithm for Variational Inequalities

Vladimir Semenov and Oleh Kharkov

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

Abstract

In this article proposed and investigated a new algorithm for solving monotone variational inequalities in Hilbert spaces. Variational inequalities provide a universal instrument of formulating many problems of mathematical physics, machine learning, data analysis, optimal control, and operations research. The proposed iterative algorithm is a regularized (by applying the Halpern scheme) variant of the operator extrapolation method. In terms of the number of calculations required to perform an iterative step, this algorithm has an advantage over the extragradient method and the method of extrapolation from the past. For variational inequalities with monotone Lipschitz continuous operators, acting in Hilbert space, the strong convergence theorem of the method is proved.

Keywords

variational inequality, monotone operator, saddle point problem, operator extrapolation method, regularization, Halpern method, strong convergence

1. Introduction

This article continues the series of articles [1–3] devoted to the development of computationally efficient and adaptive algorithms for solving variational inequalities and equilibrium problems.

Variational inequalities provide a universal instrument of formulating many topical problems of mathematical physics, machine learning, data analysis, optimal control, and operations research [4, 5]. The development of algorithms for solving variational inequalities and related problems (equilibrium problems, game problems) is an extremely popular field of research in computational mathematics [6–35]. Some problems of non-smooth optimization can be effectively solved if they are formulated as saddle problems. This approach allows to apply algorithms for solving variational inequalities in order to get a solution of the optimization problem [11]. Recently, such a variant of building fast algorithms for convex programming problems was developed: by using a duality theory, was made a transition to some convex-concave saddle problem (Fenchel game) and then applied extragradient algorithms for solving variational inequalities [12]. Note that the increased use of generative adversarial neural networks (GANs) and other adversarial or robust learning models has led to interest among machine learning specialists in algorithms for solving saddle problems and variational inequalities [13].

The simplest method for solving variational inequalities is an analogue of the gradient descent method, which in the case of the saddle point problem is known as the gradient descent-ascent method [6]. But this method may not converge for variational inequalities with a monotone operator.

A well-known modification of the gradient descent method with projection for variational inequalities is the Korpelevich extragradient method [14–17], the iteration of which requires two calculations of the value of the operator of the problem and two metric projections onto the admissible set. Computationally cheap variants of the extragradient algorithm with one metric projection on an admissible set were proposed in the articles [18, 19]. Variants of the Korpelevich extragradient method, including adaptive ones, are proposed in the articles [20–22]. In the Popov article [23] was proposed a modification of the gradient descent-ascent method different from the extragradient algorithm for finding saddle points of convex-concave functions. The iteration of this algorithm is cheaper than the

Dynamical System Modeling and Stability Investigation (DSMSI-2023), December 19-21, 2023, Kyiv, Ukraine
EMAIL: semenov.volodya@gmail.com (V. Semenov); olehharek@gmail.com (O. Kharkov)
ORCID: 0000-0002-3280-8245 (V. Semenov); 0000-0002-9049-4846 (O. Kharkov)

iteration of the extragradient algorithm in terms of the number of operator value calculations: one instead of two. Popov's algorithm for variational inequalities became known among Machine Learning specialists as Extrapolation from the Past [13]. Important results related to this algorithm are obtained in papers [1, 2, 13, 23–25]. In particular, it's adaptive modifications are proposed in the papers [1, 2].

Further development of these ideas and attempts to reduce the complexity of iteration while preserving the nature of convergence led to the inventing of a new Forward-Reflected-Backward Algorithm for solving operator inclusions [26, 27]. The algorithm has an advantage over the Korpelevich extragradient method and the method of Extrapolation from the Past in terms of the number of calculations required for the iterative step. This scheme is known as Optimistic Gradient Descent Ascent [13] and Operator Extrapolation Algorithm [3]. For the present day, the task of developing a strongly convergent variant of the operator extrapolation algorithm for variational inequalities in Hilbert space is relevant. Strongly convergent modifications for the extragradient algorithm are proposed in [2, 7]. Recently, many results have been obtained for algorithms for solving variational problems in Banach spaces [3, 9, 28–30]. In particular, analogs of the Korpelevich, Tseng, and Popov algorithms for problems in uniformly convex Banach spaces are constructed and theoretically studied. In [3] was proposed an adaptive version of the Forward-Reflected-Backward Algorithm for monotone variational inequalities in a 2-uniformly convex and uniformly smooth Banach space.

In this article a new algorithm for solving variational inequalities in Hilbert spaces is proposed. This particular algorithm is a variant of the Operator Extrapolation Method (the Forward-Reflected-Backward Algorithm from [26]), regularized by using Halpern schemes [31, 32]. For variational inequalities with monotone Lipschitz continuous operators, acting in Hilbert space, the strong convergence theorem of the method is proved.

2. Preliminaries and problem statement

Let's consider the variational inequality:

$$\text{find } x \in C: \ \langle Ax, y-x \rangle \geq 0 \quad \forall y \in C, \tag{1}$$

where C is a nonempty subset of a Hilbert space H, A is an operator, which is acting from H in H.

We denote the set of solutions (1) as S.

Assume that the following conditions are met:

- $C \subseteq H$ is a convex and closed set;
- operator $A: H \to H$ is a monotone on C, which means
$$\langle Ax - Ay, x - y \rangle \geq 0 \quad \forall x, y \in C,$$
and Lipshitz operator on C (with constant $L > 0$), which means
$$\|Ax - Ay\| \leq L \|x - y\| \quad \forall x, y \in C;$$
- S is a nonempty set.

Let's consider the dual variational inequality:

$$\text{find } x \in C: \ \langle Ay, x-y \rangle \leq 0 \quad \forall y \in C. \tag{2}$$

We denote the set of solutions (2) as S^d. It is common known, that S^d is a convex and closed set [4]. Inequality (2) is called a weak or dual formulation of the variational inequality (1) (or Minty type inequality), and the solutions of the inequality (2) – weak solutions of the variational inequality (1). For the monotone operators A we always have $S \subseteq S^d$. In our particular conditions (when the operator is also continuous), we have $S^d = S$ [4].

Let K is a nonempty closed and convex subset of a Hilbert space H. We know that for each $x \in H$ there exists unique element $z \in K$ such that

$$\|z - x\| = \inf_{y \in K} \|y - x\|.$$

This element $z \in K$ denote as $P_K x$, and the corresponding operator $P_K : H \to K$ is called projection operator from H to K (metric projection) [4]. For this operator the following statements are equivalent:

$$z = P_K x \quad \Leftrightarrow \quad z \in K, \quad \langle z - x, y - z \rangle \geq 0 \quad \forall y \in K .$$

The last inequality is equivalent to the next one [4]:

$$\|y - P_K x\|^2 \leq \|y - x\|^2 - \|P_K x - x\|^2 \quad \forall y \in K .$$

The variational inequality (1) can be formulated as the problem of finding a fixed point [4]:

$$x = P_C\left(x - \lambda A x\right), \tag{3}$$

where $\lambda > 0$. Formulation (3) is useful because it leads to an iterative scheme

$$x_{n+1} = P_C\left(x_n - \lambda A x_n\right), \tag{4}$$

which is weakly convergent for inverse strongly monotone (also known as co-coercive) operators $A : H \to H$ [10]. However, in general this scheme (4) does not convergent for Lipschitz continuous monotone operators. The most famous modification of scheme (4) is the Korpelevich extragradient method [14]:

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

the iteration of which requires two calculations of the value of the operator of the problem and two metric projections onto the admissible set**.** Computationally cheap variants of the extragradient algorithm with one metric projection on an admissible set were proposed in the articles [18, 19]. Further development of these ideas and attempts to reduce the complexity of iteration while preserving the nature of convergence led to the inventing of a new Forward-Reflected-Backward Algorithm [26]

$$x_{n+1} = P_C\left(x_n - 2\lambda A x_n + \lambda A x_{n-1}\right). \tag{5}$$

This scheme is known as Optimistic Gradient Descent Ascent [13] and Operator Extrapolation Algorithm [3]. The weak convergence of algorithm (5) is proved in [26].

The task of this article is to obtain a strongly convergent variant of the Operator Extrapolation Algorithm. In order to do this, we regularize algorithm (5) using the well-known Halpern scheme [31]

$$y_{n+1} = \alpha_n y + \left(1 - \alpha_n\right) T y_n, \tag{6}$$

where $T : H \to H$ is a nonexpansive operator, $y \in H$.

If the set of fixed points $F\left(T\right) = \left\{x \in H : x = Tx\right\}$ is nonempty and $\alpha_n \in \left(0,1\right)$, $\lim_{n\to\infty} \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = +\infty$, then scheme (6) is strongly convergent: $\lim_{n\to\infty} \left\| y_n - P_{F(T)} y \right\| = 0$.

Remark 1. Halpern's iterative scheme (6) coincides with Bakushinskii's iterative regularization scheme [7] for the method of successive approximations $x_{n+1} = T x_n$ for approximation of fixed points of the operator $T : H \to H$.

Now let`s recall the well-known lemmas about recurrent numerical inequalities.

Lemma 1. Let’s consider $\left(\xi_n\right)$ is a sequence of nonnegative numbers, which satisfies the recurrence inequality

$$\xi_{n+1} \leq \left(1 - \alpha_n\right) \xi_n + \alpha_n \beta_n \quad \text{for all } n \geq 1,$$

where sequences $\left(\alpha_n\right)$ and $\left(\beta_n\right)$ have corresponding properties: $\alpha_n \in \left(0,1\right)$ and $\beta_n \leq \beta$, where $\beta \geq 0$. Then

$$\xi_n \leq e^{-\sum_{k=1}^{n-1} \alpha_k} \xi_1 + \beta .$$

Lemma 2 ([7]). Let's consider (ξ_n) is a sequence of nonnegative numbers, which satisfies the recurrence inequality

$$\xi_{n+1} \le (1-\alpha_n)\xi_n + \alpha_n\beta_n \quad \text{for all } n \ge 1,$$

where sequences (α_n) and (β_n) have corresponding properties: $\alpha_n \in (0,1)$, $\sum_{n=1}^{\infty}\alpha_n = +\infty$, and $\overline{\lim_{n\to\infty}}\beta_n \le 0$. Then $\lim_{n\to\infty}\xi_n = 0$.

Lemma 3 ([33]). Let`s consider (a_n) is a numerical sequence, which has a subsequence (a_{n_k}) with property $a_{n_k} < a_{n_k+1}$ for all $k \ge 1$. Then there exists such a nondecreasing sequence (m_k) of natural numbers, that $m_k \to +\infty$ and $a_{m_k} \le a_{m_k+1}$, $a_k \le a_{m_k+1}$ for all $k \ge n_1$.

3. Regularized Operator Extrapolation Algorithm

In article [26], the following Operator Extrapolation Algorithm was proposed to solve the variational inequality (1) (Forward-Reflected-Backward Algorithm)

$$x_{n+1} = P_C\left(x_n - \lambda_n Ax_n - \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right)\right) = P_C\left(x_n - \left(\lambda_n + \lambda_{n-1}\right)Ax_n + \lambda_{n-1}Ax_{n-1}\right), \qquad (7)$$

where parameters λ_n satisfy the condition $0 < \inf_n \lambda_n \le \sup_n \lambda_n < 1/2L$.

Remark 2. Modifications with the Bregman projection and the generalized Alber projection are proposed in [2, 3]. In terms of the number of calculations required to perform an iterative step, this algorithm has an advantage over the Korpelevich extragradient method

$$\begin{cases} y_n = P_C\left(x_n - \lambda_n Ax_n\right), \\ x_{n+1} = P_C\left(x_n - \lambda_n Ay_n\right), \end{cases}$$

and the method of extrapolation from the past (Popov`s method)

$$\begin{cases} y_n = P_C\left(x_n - \lambda_n Ay_{n-1}\right), \\ x_{n+1} = P_C\left(x_n - \lambda_n Ay_n\right). \end{cases}$$

It is known that for variational inequalities (1) with monotone and Lipschitz operators acting in Hilbert space, algorithm (7) weakly convergent with $O\left(\frac{1}{\varepsilon}\right)$- estimate of the efficiency in terms of the gap function [3]. Based on the well-known Halpern method of approximation of fixed points of nonexpansive operators [31, 32], we will build such a regularized version of the algorithm (7).

Algorithm 1. Regularized Operator Extrapolation Algorithm.

Initialization. We set the elements $y \in H$, $x_0, x_1 \in C$, a sequence of positive numbers (λ_n) and such a sequence (α_n), that

$$\alpha_n \in (0,1),$$

$$\lim_{n\to\infty}\alpha_n = 0, \ \sum_{n=1}^{\infty}\alpha_n = +\infty.$$

Iterations. We generate a sequence (x_n) using an iterative scheme

$$x_{n+1} = P_C\left(\alpha_n y + (1-\alpha_n)x_n - \lambda_n Ax_n - (1-\alpha_n)\lambda_{n-1}\left(Ax_n - Ax_{n-1}\right)\right).$$

For positive parameters λ_n assume that this condition is fulfilled:

$$0 < \inf_n \lambda_n \le \sup_n \lambda_n < 1/2L. \qquad (8)$$

In next sections, we will prove that the sequence (x_n), generated by Algorithm 1, strongly converges to the projection of a point y onto a set S. Therefore, to find a normal solution (a solution with the smallest norm) of the variational inequality (1), we can use the scheme

$$x_{n+1} = P_C\left(\left(1-\alpha_n\right)x_n - \lambda_n A x_n - \left(1-\alpha_n\right)\lambda_{n-1}\left(Ax_n - Ax_{n-1}\right)\right).$$

Remark 3. For a smooth saddle point problem

$$\min_{x\in C}\max_{y\in D} L(x,y)$$

Algorithm 1 has the form

$$\begin{cases} x_{n+1} = P_C\left(\alpha_n x + \left(1-\alpha_n\right)x_n - \lambda_n \nabla_1 L\left(x_n, y_n\right) - \left(1-\alpha_n\right)\lambda_{n-1}\left(\nabla_1 L\left(x_n, y_n\right) - \nabla_1 L\left(x_{n-1}, y_{n-1}\right)\right)\right), \\ y_{n+1} = P_D\left(\alpha_n y + \left(1-\alpha_n\right)y_n + \lambda_n \nabla_2 L\left(x_n, y_n\right) + \left(1-\alpha_n\right)\lambda_{n-1}\left(\nabla_2 L\left(x_n, y_n\right) - \nabla_2 L\left(x_{n-1}, y_{n-1}\right)\right)\right). \end{cases}$$

Now let's prove the strong convergence of Algorithm 1.

4. Main inequalities

First, we will prove two auxiliary inequalities that will allow us to use Lemmas 1 and 2 to prove the convergence of Algorithm 1

Lemma 4. For the sequence (x_n), generated by algorithm 1, the next inequality holds

$$\begin{aligned} &\left\|x_{n+1}-z\right\|^2 + 2\lambda_n\left\langle Ax_n - Ax_{n+1}, x_{n+1}-z\right\rangle + \frac{1}{2}\left\|x_{n+1}-x_n\right\|^2 \le \\ &\le \left(1-\alpha_n\right)\left(\left\|x_n - z\right\|^2 + 2\lambda_{n-1}\left\langle Ax_{n-1} - Ax_n, x_n - z\right\rangle + \frac{1}{2}\left\|x_n - x_{n-1}\right\|^2\right) + \\ &+\alpha_n\left\|y-z\right\|^2 - \alpha_n\left\|y - x_{n+1}\right\|^2 - \left(\frac{1}{2} - \alpha_n - \left(1-\alpha_n\right)\lambda_{n-1}L\right)\left\|x_{n+1}-x_n\right\|^2 - \\ &-\left(1-\alpha_n\right)\left(\frac{1}{2} - \lambda_{n-1}L\right)\left\|x_n - x_{n-1}\right\|^2, \end{aligned} \tag{9}$$

where $z \in S$.

Proof. Let $z \in S$. Then

$$\left\langle x_{n+1} - \alpha_n y - \left(1-\alpha_n\right)x_n + \lambda_n A x_n + \left(1-\alpha_n\right)\lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), z - x_{n+1}\right\rangle \ge 0. \tag{10}$$

The monotonicity of the operator A and inclusion $z \in S$ gives us

$$\begin{aligned} &\left\langle \lambda_n A x_n + \left(1-\alpha_n\right)\lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), z - x_{n+1}\right\rangle = \\ &= \lambda_n\left\langle Ax_n - Ax_{n+1}, z - x_{n+1}\right\rangle + \left(1-\alpha_n\right)\lambda_{n-1}\left\langle\left(Ax_n - Ax_{n-1}\right), z - x_{n+1}\right\rangle + \underbrace{\lambda_n\left\langle Ax_{n+1}, z - x_{n+1}\right\rangle}_{\le 0} \le \\ &\le \lambda_n\left\langle Ax_n - Ax_{n+1}, z - x_{n+1}\right\rangle + \\ &+\left(1-\alpha_n\right)\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, z - x_n\right\rangle + \left(1-\alpha_n\right)\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\right\rangle. \end{aligned} \tag{11}$$

By using (11) in (10), we obtain

$$\begin{aligned} &0 \le 2\left\langle x_{n+1} - \alpha_n y - \left(1-\alpha_n\right)x_n, z - x_{n+1}\right\rangle + 2\lambda_n\left\langle Ax_n - Ax_{n+1}, z - x_{n+1}\right\rangle + \\ &+2\left(1-\alpha_n\right)\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, z - x_n\right\rangle + 2\left(1-\alpha_n\right)\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\right\rangle. \end{aligned} \tag{12}$$

Now let's estimate from above the application $2\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\right\rangle$ in (12). We obtain

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

$$\leq 2\lambda_{n-1}L\|x_n - x_{n-1}\|\cdot\|x_{n+1} - x_n\| \leq \lambda_{n-1}L\|x_n - x_{n-1}\|^2 + \lambda_{n-1}L\|x_n - x_{n+1}\|^2 .$$

Then we transform application $2\langle x_{n+1} - \alpha_n y - (1-\alpha_n)x_n, z - x_{n+1}\rangle$ in (12). We obtain

$$2\langle x_{n+1} - \alpha_n y - (1-\alpha_n)x_n, z - x_{n+1}\rangle =$$

$$= \|\alpha_n y + (1-\alpha_n)x_n - z\|^2 - \|x_{n+1} - z\|^2 - \|x_{n+1} - \alpha_n y - (1-\alpha_n)x_n\|^2 . \tag{13}$$

In order to transform the difference $\|\alpha_n y + (1-\alpha_n)x_n - z\|^2 - \|\alpha_n y + (1-\alpha_n)x_n - x_{n+1}\|^2$ in (13) let`s use the following identity

$$\|\alpha u + (1-\alpha)v - w\|^2 = \|v - w - \alpha(v-u)\|^2 = \|v-w\|^2 - 2\alpha\langle v-w, v-u\rangle + \alpha^2\|v-u\|^2 =$$

$$= \|v-w\|^2 - \alpha\|v-u\|^2 - \alpha\|v-w\|^2 + \alpha\|u-w\|^2 + \alpha^2\|v-u\|^2 ,$$

where $u, v, w \in H$, $\alpha > 0$. Then

$$\|\alpha_n y + (1-\alpha_n)x_n - z\|^2 - \|\alpha_n y + (1-\alpha_n)x_n - x_{n+1}\|^2 =$$

$$= (1-\alpha_n)\|x_n - z\|^2 - (1-\alpha_n)\|x_n - x_{n+1}\|^2 + \alpha_n\|y-z\|^2 - \alpha_n\|y - x_{n+1}\|^2 .$$

Now we have this inequality

$$0 \leq (1-\alpha_n)\|x_n - z\|^2 - (1-\alpha_n)\|x_n - x_{n+1}\|^2 + \alpha_n\|y-z\|^2 - \alpha_n\|y - x_{n+1}\|^2 - \|x_{n+1} - z\|^2 +$$
$$+2\lambda_n\langle Ax_n - Ax_{n+1}, z - x_{n+1}\rangle + 2(1-\alpha_n)\lambda_{n-1}\langle Ax_n - Ax_{n-1}, z - x_n\rangle +$$
$$+(1-\alpha_n)\lambda_{n-1}L\|x_n - x_{n-1}\|^2 + (1-\alpha_n)\lambda_{n-1}L\|x_n - x_{n+1}\|^2 . \tag{14}$$

We rearrange the terms in (14) and finally get

$$\|x_{n+1} - z\|^2 + 2\lambda_n\langle Ax_n - Ax_{n+1}, x_{n+1} - z\rangle + \frac{1}{2}\|x_{n+1} - x_n\|^2 \leq$$
$$\leq (1-\alpha_n)\left(\|x_n - z\|^2 + 2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z\rangle + \frac{1}{2}\|x_n - x_{n-1}\|^2\right) +$$
$$+\alpha_n\|y-z\|^2 - \alpha_n\|y - x_{n+1}\|^2 - \left(\frac{1}{2} - \alpha_n - (1-\alpha_n)\lambda_{n-1}L\right)\|x_{n+1} - x_n\|^2 -$$
$$-(1-\alpha_n)\left(\frac{1}{2} - \lambda_{n-1}L\right)\|x_n - x_{n-1}\|^2 ,$$

which had to be proved. ■

Lemma 5. For the sequence (x_n), generated by Algorithm 1, the inequality holds

$$\|x_{n+1} - z\|^2 + 2\lambda_n\langle Ax_n - Ax_{n+1}, x_{n+1} - z\rangle + \frac{1}{2}\|x_{n+1} - x_n\|^2 \leq$$
$$\leq (1-\alpha_n)\left(\|x_n - z\|^2 + 2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z\rangle + \frac{1}{2}\|x_n - x_{n-1}\|^2\right) + 2\alpha_n\langle y - z, x_{n+1} - z\rangle -$$
$$-\left(\frac{1}{2} - \alpha_n - (1-\alpha_n)\lambda_{n-1}L\right)\|x_{n+1} - x_n\|^2 - (1-\alpha_n)\left(\frac{1}{2} - \lambda_{n-1}L\right)\|x_n - x_{n-1}\|^2 , \tag{15}$$

where $z \in S$.

Proof. Let's apply an elementary inequality $\|a+b\|^2 \leq \|a\|^2 + 2\langle b, a+b\rangle$. We obtain

$$\|y - z\|^2 = \|y - x_{n+1} + x_{n+1} - z\|^2 \leq \|y - x_{n+1}\|^2 + 2\langle y - z, x_{n+1} - z\rangle . \tag{16}$$

By using (16) in (9), we get (15), which had to be proved. ■

5. Strong convergence

Now let`s prove that the sequence is bounded (x_n).

Lemma 6. Let the condition (8) be fulfilled. Then the sequence (x_n), generated by Algorithm 1, is bounded.

Proof. Since $0 < \inf_n \lambda_n \le \sup_n \lambda_n < \frac{1}{2L}$ and $\lim_{n\to\infty} \alpha_n = 0$, there exists such number $n_0 \ge 1$, that

$$\frac{1}{2} - \alpha_n - (1-\alpha_n)\lambda_{n-1}L = \frac{1}{2} - \lambda_{n-1}L - \alpha_n(1-\lambda_{n-1}L) > 0 \quad \text{and} \quad (1-\alpha_n)\left(\frac{1}{2} - \lambda_{n-1}L\right) > 0 \cdot \tag{17}$$

From (9) and (17) we obtain, that for $n \ge n_0$ the next inequality holds

$$\xi_{n+1} \le (1-\alpha_n)\xi_n + \alpha_n \|y - z\|^2, \tag{18}$$

where $\xi_n = \|x_n - z\|^2 + 2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z\rangle + \frac{1}{2}\|x_n - x_{n-1}\|^2$, $z \in S$.

Let's get the lower bound of ξ_n. We obtain

$$\xi_n = \|x_n - z\|^2 + 2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z\rangle + \frac{1}{2}\|x_n - x_{n-1}\|^2 \ge$$
$$\ge \|x_n - z\|^2 - 2\lambda_{n-1}\|Ax_{n-1} - Ax_n\|\|x_n - z\| + \frac{1}{2}\|x_{n-1} - x_n\|^2 \ge$$
$$\ge \|x_n - z\|^2 - 2\lambda_{n-1}L\|x_{n-1} - x_n\|\|x_n - z\| + \frac{1}{2}\|x_{n-1} - x_n\|^2 \ge$$
$$\ge (1-\lambda_{n-1}L)\|x_n - z\|^2 + \left(\frac{1}{2} - \lambda_{n-1}L\right)\|x_n - x_{n-1}\|^2 \ge 0 \cdot \tag{19}$$

From inequalities (18), (19) and Lemma 1 follows the boundedness of the sequences (ξ_n) and (x_n), which had to be proved. Let's formulate the main result.

Theorem 1. Let C is a nonempty convex closed subset of Hilbert space H, $A: H \to H$ is a monotone and Lipschitz continuous operator on the set C, $S \ne \varnothing$, $y \in H$, condition (8) is fulfilled. Then the sequence (x_n), generated by Algorithm 1, strongly converges to $z = P_S y$.

Proof. Let $z = P_S y$. Lemma 6 implies the existence of such a number $M \ge 0$, that

$$|\langle y - z, x_{n+1} - z\rangle| \le M \quad \text{for all} \quad n \ge 1.$$

Then from Lemma 5 the next inequality follows

$$\xi_{n+1} - \xi_n + \alpha_n\xi_n + \left(\frac{1}{2} - \alpha_n - (1-\alpha_n)\lambda_{n-1}L\right)\|x_{n+1} - x_n\|^2 +$$
$$+ (1-\alpha_n)\left(\frac{1}{2} - \lambda_{n-1}L\right)\|x_n - x_{n-1}\|^2 \le 2\alpha_n M, \tag{20}$$

where $\xi_n = \|x_n - z\|^2 + 2\lambda_{n-1}\langle Ax_{n-1} - Ax_n, x_n - z\rangle + \frac{1}{2}\|x_n - x_{n-1}\|^2$.

Consider a numerical sequence (ξ_n). Then two options are possible:

1. there exists a number $\bar{n} \ge 1$ that $\xi_{n+1} \le \xi_n$ for all $n \ge \bar{n}$;
2. there exists an increasing sequence of numbers (n_k) that $\xi_{n_k+1} > \xi_{n_k}$ for all $k \ge 1$.

First let`s consider the option 1. In this case there exists $\lim_{n\to\infty} \xi_n$. Since $\xi_{n+1} - \xi_n \to 0$, $\alpha_n \to 0$ and inequalities (20) are fulfilled, then for $n \to \infty$ we obtain

$$\|x_{n+1}-x_n\|\to 0. \qquad (21)$$

Let`s show that all weak partial limits of the sequence $(x_n)$ belong to $S$. Consider a subsequence $(x_{n_k})$, which weakly converges to some point $w\in H$. It is obvious, that $w\in C$. Let`s show that $w\in S$. We have

$$\left\langle x_{n_k+1}-\alpha_{n_k}y-(1-\alpha_{n_k})x_{n_k}+\lambda_{n_k}Ax_{n_k}+(1-\alpha_{n_k})\lambda_{n_k-1}\left(Ax_{n_k}-Ax_{n_k-1}\right),y-x_{n_k+1}\right\rangle\geq 0 \quad \forall y\in C.$$

By using the monotonicity of the operator A, derive an estimate:

$$\left\langle Ay,y-x_{n_k}\right\rangle+\left\langle Ax_{n_k},x_{n_k}-x_{n_k+1}\right\rangle\geq\left\langle Ax_{n_k},y-x_{n_k+1}\right\rangle\geq$$

$$\geq\frac{1}{\lambda_{n_k}}\left\langle\alpha_{n_k}\left(y-x_{n_k}\right)+x_{n_k}-x_{n_k+1},y-x_{n_k+1}\right\rangle-(1-\alpha_{n_k})\frac{\lambda_{n_k-1}}{\lambda_{n_k}}\left\langle Ax_{n_k}-Ax_{n_k-1},y-x_{n_k+1}\right\rangle \quad \forall y\in C.$$

From $\lim_{n\to\infty}\alpha_n=0$, constraint of the sequence (x_n), (21) and Lipshitz property of operator A, we obtain

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

On the other hand

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

Thus, $w\in S$.

Let`s prove that

$$\varlimsup_{n\to\infty}\langle y-z,x_{n+1}-z\rangle\leq 0. \qquad (22)$$

Consider the following subsequence (x_{n_k}), that

$$\lim_{k\to\infty}\left\langle y-z,x_{n_k}-z\right\rangle=\varlimsup_{n\to\infty}\langle y-z,x_{n+1}-z\rangle.$$

Let`s consider that $x_{n_k}\to w\in S$ weakly. Then we obtain

$$\lim_{k\to\infty}\left\langle y-z,x_{n_k}-z\right\rangle=\langle y-z,w-z\rangle=\langle y-P_Sy,w-P_Sy\rangle\leq 0,$$

which is a proof for (22).

Now from (22), inequality

$$\xi_{n+1}\leq(1-\alpha_n)\xi_n+2\alpha_n\langle y-z,x_{n+1}-z\rangle,$$

which holds for sufficiently large n, and Lemma 2 we conclude that

$$\xi_n=\|x_n-z\|^2+2\lambda_{n-1}\langle Ax_{n-1}-Ax_n,x_n-z\rangle+\frac{1}{2}\|x_n-x_{n-1}\|^2\to 0.$$

From (19) we obtain $\lim_{n\to\infty}\|x_n-z\|=0$.

Now let`s consider option 2. In this case there exists a nondecreasing sequence of numbers (m_k) with the following properties (Lemma 3):

1. $m_k\to+\infty$;
2. $\xi_{m_k+1}\geq\xi_{m_k}$ for all $k\geq n_1$;
3. $\xi_{m_k+1}\geq\xi_k$ for all $k\geq n_1$.

From the inequality of Lemma 5, (19) and second property we get

$$\left(\frac{1}{2}-\alpha_{m_k}-(1-\alpha_{m_k})\lambda_{m_k-1}L\right)\|x_{m_k+1}-x_{m_k}\|^2+(1-\alpha_{m_k})\left(\frac{1}{2}-\lambda_{m_k-1}L\right)\|x_{m_k}-x_{m_k-1}\|^2\leq 2\alpha_{m_k}M.$$

This leads us to

$$\lim_{k\to\infty}\left\|x_{m_k+1}-x_{m_k}\right\|=\lim_{k\to\infty}\left\|x_{m_k}-x_{m_k-1}\right\|=0 .$$

By similar reasoning, we prove that the partial limits of the sequence are weak $\left(x_{m_k}\right)$ belong to S. From identity

$$\left\langle y-z, x_{m_k+1}-z\right\rangle=\left\langle y-z, x_{m_k}-z\right\rangle+\left\langle y-z, x_{m_k+1}-x_{m_k}\right\rangle$$

we obtain

$$\overline{\lim_{k\to\infty}}\left\langle y-z, x_{m_k+1}-z\right\rangle=\overline{\lim_{k\to\infty}}\left\langle y-z, x_{m_k}-z\right\rangle .$$

As in the previous part, we obtain the inequality

$$\overline{\lim_{k\to\infty}}\left\langle y-z, x_{m_k+1}-z\right\rangle\le 0 .$$

Then we get

$$\xi_{m_k+1}\le\left(1-\alpha_{m_k}\right)\xi_{m_k}+2\alpha_{m_k}\left\langle y-z, x_{m_k+1}-z\right\rangle\le\left(1-\alpha_{m_k}\right)\xi_{m_k+1}+2\alpha_{m_k}\left\langle y-z, x_{m_k+1}-z\right\rangle .$$

With respect to the third property, we obtain

$$\xi_k\le\xi_{m_k+1}\le 2\left\langle y-z, x_{m_k+1}-z\right\rangle .$$

As a result, we get

$$\overline{\lim_{k\to\infty}}\,\xi_k\le 2\,\overline{\lim_{k\to\infty}}\left\langle y-z, x_{m_k+1}-z\right\rangle\le 0 .$$

Thus, $\lim_{n\to\infty}\xi_n=0$ and, consequently, $\lim_{n\to\infty}\left\|x_n-z\right\|=0$, which had to be proved. ■

6. Conclusions

In this article proposed and investigated a new algorithm for solving variational inequalities in Hilbert spaces. The proposed iterative algorithm is a regularized (by applying the Halpern scheme [32, 33]) variant of the Operator Extrapolation Method (Forward-Reflected-Backward Algorithm from [26]). For variational inequalities with monotone Lipschitz continuous operators, acting in Hilbert space, the strong convergence theorem of the method is proved.

An important issue is the study of the asymptotic behavior of Algorithm 1 in the situation $C=H$:

$$x_{n+1}=\alpha_n y+\left(1-\alpha_n\right)x_n-\lambda_n Ax_n-\left(1-\alpha_n\right)\lambda_{n-1}\left(Ax_n-Ax_{n-1}\right).$$

To be more precise, this issue is about the behavior of the norm $\left\|Ax_n\right\|$. In our opinion, the estimation should be $\left\|Ax_n\right\|=\mathrm{O}\left(1/n\right)$. Note that in [34] was obtained an estimate for the extragradient method such as $\left\|Ax_n\right\|=\mathrm{O}\left(1/\sqrt{n}\right)$, and in [35] $\left\|Ax_n\right\|=\mathrm{O}\left(1/n\right)$ for the extragradient method with Halpern regularization

$$\begin{cases} y_n=x_n+\dfrac{1}{n+2}\left(x_0-x_n\right)-\dfrac{1}{8L}Ax_n, \\ x_{n+1}=x_n+\dfrac{1}{n+2}\left(x_0-x_n\right)-\dfrac{1}{8L}Ay_n. \end{cases}$$

The parameters λ_n of Algorithm 1 satisfy the condition $0<\inf_n\lambda_n\le\sup_n\lambda_n<1/2L$. This means that the information about the Lipschitz constants of the operator A was used a priori. Algorithm 1 and the scheme from articles [1–3] allow you to build such an algorithm with adaptive value selection λ_n, that which does not require knowledge of Lipschitz constants of operators and linear search type procedures.

Algorithm 2. Adaptive regularized operator extrapolation algorithm.

Initialization. Set $y \in H$, elements $x_0, x_1 \in C$, numbers

$$\tau \in \left(0, \tfrac{1}{2}\right), \ \lambda_1, \lambda_0 > 0,$$

and such a sequence (α_n), which have properties $\alpha_n \in (0,1)$,

$$\lim_{n\to\infty} \alpha_n = 0, \ \sum_{n=1}^{\infty} \alpha_n = +\infty.$$

Iterations. We generate a sequence (x_n) by using an iterative scheme

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

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

In addition, based on the results of work [3], it is possible to obtain an analogue of Algorithm 1 with a generalized Alber projection for solving variational inequalities in uniformly convex and uniformly smooth Banach spaces.

7. Acknowledgements

This work was supported by the Ministry of Education and Science of Ukraine (project "Computational algorithms and optimization for artificial intelligence, medicine and defense", 0122U002026).

8. References

[1] Y. I. Vedel, S. V. Denisov, V. V. Semenov, An Adaptive Algorithm for the Variational Inequality Over the Set of Solutions of the Equilibrium Problem, Cybernetics and Systems Analysis 57 (2021) 91–100. doi:10.1007/s10559-021-00332-2.

[2] V. V. Semenov, S. V. Denisov, A. V. Kravets, Adaptive Two-Stage Bregman Method for Variational Inequalities, Cybernetics and Systems Analysis 57 (2021) 959–967. doi:10.1007/s10559-021-00421-2.

[3] Y. Vedel, V. Semenov, S. Denisov, A Novel Algorithm with Self-adaptive Technique for Solving Variational Inequalities in Banach Spaces, in: N. N. Olenev, Y. G. Evtushenko, M. Jaćimović, M. Khachay, V. Malkova (Eds.), Advances in Optimization and Applications, volume 1514 of Communications in Computer and Information Science, Springer, Cham, 2021, pp. 50–64. doi:10.1007/978-3-030-92711-0_4.

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

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

[6] S. P. Uryas'ev, Adaptive algorithms of stochastic optimization and game theory, Nauka, Moscow, 1990.

[7] A. B. Bakushinskii, A. V. Goncharskii, Iterative Methods for Solving Ill-Posed Problems, Nauka, Moscow, 1989.

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

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

[10] H. H. Bauschke, P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, Berlin–Heidelberg–New York, 2017.

[11] 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.

[12] J.-K. Wang, J. Abernethy, K. Y. Levy, No-Regret Dynamics in the Fenchel Game: A Unified Framework for Algorithmic Convex Optimization, arXiv preprint arXiv:2111.11309. (2021).

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

[14] G. M. Korpelevich, An extragradient method for finding saddle points and for other problems, Matecon. 12 (1976) 747–756.
[15] E. N. Khobotov, Modification of the extra-gradient method for solving variational inequalities and certain optimization problems, USSR Comput. Math. Phys. 27 (1987) 120–127. doi:10.1016/0041-5553(87)90058-9.
[16] N. Nadezhkina, W. Takahashi, Weak convergence theorem by an extragradient method for nonexpansive mappings and monotone mappings, Journal of Optimization Theory and Applications 128 (2006) 191–201.
[17] S. V. Denisov, D. A. Nomirovskii, B. V. Rublyov, V. V. Semenov, Convergence of Extragradient Algorithm with Monotone Step Size Strategy for Variational Inequalities and Operator Equations, Journal of Automation and Information Sciences 51 (2019) 12–24. doi:10.1615/JAutomatInfScien.v51.i6.20.
[18] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM Journal* on Control and *Optimization* 38 (2000) 431–446.
[19] 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.
[20] F. Bach, K. Y. Levy, A Universal Algorithm for Variational Inequalities Adaptive to Smoothness and Noise, arXiv preprint arXiv:1902.01637. (2019).
[21] 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.
[22] 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.
[23] 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.
[24] 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.
[25] E. Gorbunov, A. Taylor, G. Gidel, Last-Iterate Convergence of Optimistic Gradient Method for Monotone Variational Inequalities, arXiv preprint arXiv:2205.08446. (2022).
[26] 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.
[27] 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.
[28] H. Iiduka, W. Takahashi, Weak convergence of a projection algorithm for variational inequalities in a Banach space, Journal of Mathematical Analysis and Applications 339 (2008) 668–679. https://doi.org/10.1016/j.jmaa.2007.07.019.
[29] 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.
[30] 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.
[31] B. Halpern, Fixed points of nonexpanding maps, Bull. Amer. Math. Soc. 73 (1967) 957–961.
[32] H. K. Xu, Viscosity approximation methods for nonexpansive mappings, Journal of Mathematical Analysis and Applications 298 (2004) 279–291. doi:10.1016/j.jmaa.2004.04.059.
[33] P.-E. Mainge, Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization, Set-Valued Analysis 16 (2008) 899–912.
[34] E. Gorbunov, N. Loizou, G. Gidel, Extragradient Method: $O(1/K)$ Last-Iterate Convergence for Monotone Variational Inequalities and Connections With Cocoercivity, arXiv preprint arXiv: 2110.04261. (2021). doi:10.48550/arXiv.2110.04261.
[35] T. Yoon, E. K. Ryu, Accelerated algorithms for smooth convex-concave minimax problems with $O(1/k^2)$ rate on squared gradient norm, Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research 139 (2021) 12098–12109.