DOI 10.1007/s10559-017-9923-9
Cybernetics and Systems Analysis, Vol. 53, No. 2, March, 2017

A VERSION OF THE MIRROR DESCENT METHOD TO SOLVE VARIATIONAL INEQUALITIES*

V. V. Semenov

UDC 517.988

Abstract. *Nemirovski and Yudin proposed the mirror descent algorithm at the late 1970s to solve convex optimization problems. This method is suitable to solve huge-scale optimization problems. In the paper, we describe a new version of the mirror descent method to solve variational inequalities with pseudomonotone operators. The method can be interpreted as a modification of Popov's two-step algorithm with the use of Bregman projections on the feasible set. We prove the convergence of the sequences generated by the proposed method.*

Keywords: *variational inequality, pseudomonotonicity, Bregman distance, Kullback–Leibler distance, mirror descent method, convergence.*

INTRODUCTION

Many interesting and important problems of operations research and mathematical physics can be written as variational inequalities. Solving these inequalities is a rapidly developing trend of applied nonlinear analysis [1–16]. By now, plenty of methods have been developed to solve variational inequalities, in particular, those of projection type, i.e., using metric design on a feasible set [1, 5, 7, 8]. It is well known that in problems of finding a saddle point or Nash equilibrium, for the most simple projection method to converge, strengthened monotonicity conditions should be satisfied [1]. If they are not satisfied, several approaches can be used. One of them is to regularize the original problem in order to provide a required property to it. Convergence without problem modification is ensured in iterative extragradient methods, proposed for the first time by Korpelevich [5] for inequalities with Lipschitzian operators. Later, the method with dynamic step adjustment was considered in [6]. It does not need the Lipschitz constant of the inequality operator to be known, which considerably extends the domain of potential application of the idea outlined in [5]. These methods were analyzed in many studies [7–13]. In particular, modifications of the Korpelevich algorithm with one metric projection on feasible set were proposed [8, 9, 12, 13]. In so-called subgradient–extragradient algorithms and Korpelevich algorithm, the first stages of iteration coincide, and then to obtain the next approximation, projection on some half-space that is support for the feasible set is carried out instead of projection on the feasible set. In the early 1980s, an interesting modification of the Arrow–Hurwitz algorithm of search for saddle point of convex–concave functions was proposed [14]. Some modifications of the Popov method for the solution of variational inequalities with monotone operator are analyzed in recent studies [15, 16]. The paper [17] proposes two-stage proximal algorithm to solve equilibrium programming problems, which is an adaptation of the method [14] to general Ky Fan inequalities.

In all the above-mentioned methods, Euclidean distance and projection were used, which does not allow taking into account the structure of feasible sets and solving the problems efficiently. A possible way out is more flexible choice

*The study was supported by the Ministry of Education and Science of Ukraine (project "Development of the algorithms for modeling and optimization of dynamic systems for defense, medicine, and ecology," 0116U004777).

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, *semenov.volodya@gmail.com*. Translated from Kibernetika i Sistemnyi Analiz, No. 2, March–April, 2017, pp. 83–93. Original article submitted August 3, 2016.

of the distance for projecting on feasible set. One of the first successful implementations of such strategy is the study [18], where a cyclic projection method is proposed to find a common point of convex sets. This publication created a new trend in mathematical programming and nonlinear analysis.

The mirror descent method proposed in the late 1970s by Soviet mathematicians Nemirovski and Yudin to solve convex optimization problems [19] became widely popular to solve high-dimensional problems [20–23]. In case of constrained problems, it can be interpreted as a version of the method of projection of subgradient where projection is understood in the sense of Bregman distance [21]. The mirror descent method allows taking into account the structure of admissible set of the optimization problem. For example, for simplex, it is possible to use the Kullback–Leibler distance (the Bregman distance constructed on the negative entropy) and to obtain explicitly calculated operator of projection on simplex [21]. The studies [24–28] analyze versions of this method for solution of variational inequalities and saddle problems, constructed on the basis of extragradient Korpelevich algorithm, including stochastic ones [27, 28].

The present paper analyzes a new version of the mirror descent method for the solution of variation inequalities with Lipschitzian and pseudomonotone operators constructed on the basis of the two-stage Popov algorithm [14, 15].

PROBLEM STATEMENT AND AUXILIARY INFORMATION

In what follows, we will use finite-dimensional real linear space denoted by E. In this space, consider the norm $||\cdot||$ (not necessarily Euclidean). Denote the dual space by E^*. For $a \in E^*$ and $b \in E$ denote by (a, b) the value of linear function a at point b. Define the dual norm $||\cdot||_*$ on E^* in a standard way: $||a||_* = \max\{(a, b) : ||b|| = 1\}$, which ensures the Schwarz inequality $(a, b) \le ||a||_* \, ||b||$ for all $a \in E^*$, $b \in E$. The most important case is $E = E^* = \mathbb{R}^m$ with $(a, b) = \sum_{i=1}^{m} a_i b_i$, a, $b \in \mathbb{R}^m$.

Let C be a nonempty subset of space E and A be an operator acting from E into E^*. Consider the variational inequality: find

$$x \in C:\ (Ax, y - x) \ge 0 \ \ \forall\, y \in C, \tag{1}$$

whose set of solutions denote by S.

Assume that the following conditions are satisfied:

- set $C \subseteq E$ is convex and closed;
- operator $A: E \to E^*$ is pseudomonotone and Lipschitz with the constant $L > 0$ on C;
- set S is not empty.

Remark 1. Let us recall that pseudomonotonicity of operator A on set C is that for all x, $y \in C$ from $(Ax, y - x) \ge 0$ it follows that $(Ay, y - x) \ge 0$ [1]. In case $f : \mathbb{R} \to \mathbb{R}$ this means that if $f(x) = 0$ for some point $x \in \mathbb{R}$, then $f(y) \le 0$ for $y \le x$ and $f(y) \ge 0$ for $y \ge x$.

Consider so-called dual variational inequality [1]: find

$$x \in C:\ (Ay, y - x) \ge 0 \ \ \forall\, y \in C. \tag{2}$$

Denote the set of solutions of (2) by S^d. Inequality (2) is sometimes called weak or dual statement of (1), and solutions of (2) weak solutions of (1) [1]. Indeed, in case of pseudomonotonicity of A we get $S \subseteq S^d$. Under the considered conditions, $S^d = S$. In particular, the set S is convex and closed [1].

Let us introduce the structures necessary to formulate the algorithm. Let function $\varphi : E \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ satisfy the following conditions [27]:

- φ is continuous and convex on C; in particular, set $C^0 = \{x \in C : \partial\varphi(x) \ne \varnothing\}$ is not empty;
- φ is regular on C^0, i.e., the subdifferential $\partial\varphi$ on set C^0 is a continuous selector $\nabla\varphi$;

• φ is strongly convex with respect to the selected norm $||\cdot||$ with the strong convexity constant $\sigma > 0$:

$$\varphi(a) \geq \varphi(b) - (\nabla\varphi(b), a-b) + \frac{\sigma}{2}||a-b||^2 \quad \forall\, a \in C,\ b \in C^0.$$

Remark 2. Such functions are called distance generating functions [27].
Problem

$$(a, y) + \varphi(y) \to \min{}_{y\in C},\ a \in E^*,$$

has a unique solution y_a, lying in C^0, and

$$(a + \nabla\varphi(y_a), y - y_a) \geq 0 \quad \forall\, y \in C.$$

The Bregman distance respective to φ on set C is specified by the formula

$$d(a,b) = \varphi(a) - \varphi(b) - (\nabla\varphi(b), a-b) \quad \forall\, a \in C,\ b \in C^0.$$

Consider two main examples. For $\varphi(\cdot) = \frac{1}{2}||\cdot||_2^2$, where $||\cdot||_2$ is the Euclidean norm, we have $d(x, y) = \frac{1}{2}||x-y||_2^2$. For standard simplex $S_m = \left\{ x \in \mathbb{R}^m : x_i \geq 0,\ \sum_{i=1}^{m} x_i = 1 \right\}$ and negative Boltsman–Shannon entropy $\varphi(x) = \sum_{i=1}^{m} x_i \ln x_i$ (it is strongly convex with respect to ℓ_1-norm on S_m) we obtain the Kullback–Leibler distance

$$d(x, y) = \sum_{i=1}^{m} x_i \ln\frac{x_i}{y_i},\ x \in S_m,\ y \in \mathrm{ri}\,(S_m).$$

The useful three-point identity [21] takes place

$$d(a,c) = d(a,b) + d(b,c) + (\nabla\varphi(b) - \nabla\varphi(c), a-b). \tag{3}$$

From the strong convexity of φ the estimate follows

$$d(a,b) \geq \frac{\sigma}{2}||a-b||^2 \quad \forall\, a \in C,\ b \in C^0. \tag{4}$$

Assume that there exists a possibility to efficiently solve strongly convex minimization problems of the form

$$\pi_x(a) = \arg\min{}_{y\in C} \{-(a, y-x) + d(y,x)\} \quad \forall\, a \in E^*,\ x \in C^0.$$

Point $\pi_x(a)$ in the Euclidean case coincides with the Euclidean metric projection

$$P_C(x+a) = \arg\min{}_{y\in C} ||\, y - (x+a)||_2.$$

For the case of simplex S_m and Kullback–Leibler distance, we get [21]

$$\pi_x(a) = \left(\frac{x_1 e^{a_1}}{\sum_{j=1}^{m} x_j e^{a_j}}, \frac{x_2 e^{a_2}}{\sum_{j=1}^{m} x_j e^{a_j}}, \ldots, \frac{x_m e^{a_m}}{\sum_{j=1}^{m} x_j e^{a_j}} \right),\ a \in \mathbb{R}^m,\ x \in \mathrm{ri}\,(S_m).$$

Remark 3. In [22, 23], for $\pi_x(a)$ the notation $\mathrm{Mirr}_x(a)$ is accepted. The operator $\pi_x : E^* \to C^0$ is called mirror-prox [27].

VARIANT OF THE MIRROR DESCENT METHOD

Let us describe a variant of the mirror descent method for problem (1).

Algorithm 1

Beginning with $x_1 \in C^0$ and $y_1 \in C$, generate the sequence of elements x_n, y_n with the help of iterative scheme

$$x_{n+1} = \pi_{x_n}(-\lambda A y_n), \quad y_{n+1} = \pi_{x_{n+1}}(-\lambda A y_n),$$

where $\lambda > 0$.

Below we will formulate the rule to choose the regularization parameter λ.

Remark 4. If $\varphi(\cdot) = \frac{1}{2}||\cdot||_2^2$, then Algorithm 1 becomes [14, 15, 17]:

$$\begin{cases} x_{n+1} = P_C(x_n - \lambda A y_n), \\ y_{n+1} = P_C(x_{n+1} - \lambda A y_n). \end{cases}$$

If for some $n \in \mathbb{N}$ the equalities

$$x_{n+1} = x_n = y_n \tag{5}$$

hold, then the inclusion $y_n \in S$ takes place as well as the stationarity condition $x_k = y_k = y_n$ for $k \geq n$. Indeed, the equality $x_{n+1} = \pi_{x_n}(-\lambda A y_n)$ means

$$(Ay_n, y - x_{n+1}) + \frac{(\nabla\varphi(x_{n+1}) - \nabla\varphi(x_n), y - x_{n+1})}{\lambda} \geq 0 \ \ \forall\, y \in C.$$

From (5) it follows that $(Ay_n, y - y_n) \geq 0 \ \forall\, y \in C$, i.e., $y_n \in S$.

Taking this into account, we can make the practical variant of Algorithm 1 as follows.

Algorithm 2

Step 0. Specify $x_1 \in C^0$, $y_1 \in C$, $\lambda > 0$ and $\varepsilon > 0$.

Step 1. For x_n and y_n calculate

$$x_{n+1} = \pi_{x_n}(-\lambda A y_n) = \arg\min_{y \in C} \{\lambda(Ay_n, y - x_n) + d(y, x_n)\}.$$

Step 2. If $\max\{||x_{n+1} - x_n||, ||x_n - y_n||\} \leq \varepsilon$, then STOP; otherwise calculate

$$y_{n+1} = \pi_{x_{n+1}}(-\lambda A y_n) = \arg\min_{y \in C} \{\lambda(Ay_n, y - x_{n+1}) + d(y, x_{n+1})\}.$$

Step 3. Put $n = n+1$ and go to Step 1.

Remark 5. We can also use the condition $\max\{d(x_{n+1}, x_n), d(y_n, x_n)\} \leq \varepsilon$.

In what follows, we assume that for all numbers $n \in \mathbb{N}$ condition (5) is not satisfied and pass to substantiating the convergence of Algorithm 1.

PROOF OF THE CONVERGENCE OF THE METHOD

To prove the convergence of the method, we will use the following lemma.

LEMMA 1. Let (a_n) and (b_n) be sequences of nonnegative numbers satisfying the inequality $a_{n+1} \leq a_n - b_n$ for all $n \in \mathbb{N}$. Then there exists a limit $\lim_{n\to\infty} a_n$ and $\sum_{n=1}^{\infty} b_n < +\infty$.

Let us obtain an important estimate that describes the behavior of the Bregman distance between the point generated by the algorithm x_n and arbitrary element from the set S.

LEMMA 2. For sequences (x_n) and (y_n) generated by the algorithm, the inequality

$$d(z, x_{n+1}) \le d(z, x_n) - \left(1 - (1+\sqrt{2})\frac{\lambda L}{\sigma}\right) d(y_n, x_n)$$

$$- \left(1 - \sqrt{2}\frac{\lambda L}{\sigma}\right) d(x_{n+1}, y_n) + \frac{\lambda L}{\sigma} d(x_n, y_{n-1}), \tag{6}$$

holds, where $z \in S$.

Proof. Applying identity (3) twice yields

$$d(z, x_{n+1}) = d(z, x_n) - d(x_{n+1}, x_n) + (\nabla\varphi(x_{n+1}) - \nabla\varphi(x_n), x_{n+1} - z)$$

$$= d(z, x_n) - d(x_{n+1}, y_n) - d(y_n, x_n) - (\nabla\varphi(y_n) - \nabla\varphi(x_n), x_{n+1} - y_n)$$

$$+ (\nabla\varphi(x_{n+1}) - \nabla\varphi(x_n), x_{n+1} - z). \tag{7}$$

From the definition of points x_{n+1} and y_n the inequalities follow

$$\lambda(Ay_n, z - x_{n+1}) + (\nabla\varphi(x_{n+1}) - \nabla\varphi(x_n), z - x_{n+1}) \ge 0, \tag{8}$$

$$\lambda(Ay_{n-1}, x_{n+1} - y_n) + (\nabla\varphi(y_n) - \nabla\varphi(x_n), x_{n+1} - y_n) \ge 0. \tag{9}$$

Using (8) and (9) for the estimate of scalar products in (7), we obtain

$$d(z, x_{n+1}) \le d(z, x_n) - d(x_{n+1}, y_n) - d(y_n, x_n)$$

$$+ \lambda\{(Ay_{n-1}, x_{n+1} - y_n) + (Ay_n, z - x_{n+1})\}$$

$$= d(z, x_n) - d(x_{n+1}, y_n) - d(y_n, x_n)$$

$$+ \lambda\{(Ay_{n-1} - Ay_n, x_{n+1} - y_n) + (Ay_n, z - y_n)\}. \tag{10}$$

From the pseudomonotonicity of A it follows that $(Ay_n, z - y_n) \le 0$. Using this estimate in (10), we obtain

$$d(z, x_{n+1}) \le d(z, x_n) - d(x_{n+1}, y_n) - d(y_n, x_n) + \lambda(Ay_{n-1} - Ay_n, x_{n+1} - y_n). \tag{11}$$

Let us now estimate the term $\lambda(Ay_{n-1} - Ay_n, x_{n+1} - y_n)$. We have

$$\lambda(Ay_{n-1} - Ay_n, x_{n+1} - y_n) \le \lambda\|Ay_{n-1} - Ay_n\|_* \|x_{n+1} - y_n\|$$

$$\le \lambda L\|y_{n-1} - y_n\| \|x_{n+1} - y_n\| \le \lambda L\left\{\frac{1}{2\sqrt{2}}\|y_{n-1} - y_n\|^2 + \frac{1}{\sqrt{2}}\|x_{n+1} - y_n\|^2\right\}$$

$$\le \frac{\lambda L}{2\sqrt{2}}\{\sqrt{2}\|y_{n-1} - x_n\|^2 + (2+\sqrt{2})\|x_n - y_n\|^2\} + \frac{\lambda L}{\sqrt{2}}\|x_{n+1} - y_n\|^2$$

$$= \frac{\lambda L}{2}\|y_{n-1} - x_n\|^2 + \lambda L\frac{1+\sqrt{2}}{2}\|x_n - y_n\|^2 + \frac{\lambda L}{\sqrt{2}}\|x_{n+1} - y_n\|^2. \tag{12}$$

In this relation, we have used the elementary inequalities

$$ab \le \frac{\varepsilon^2}{2}a^2 + \frac{1}{2\varepsilon^2}b^2,\ (a+b)^2 \le \sqrt{2}a^2 + (2+\sqrt{2})b^2.$$

Estimating the norms in (12) by means of inequality (4), we obtain

$$\lambda(Ay_{n-1}-Ay_n, x_{n+1}-y_n) \le \frac{\lambda L}{\sigma} d(x_n, y_{n-1})$$

$$+\frac{\lambda L}{\sigma}(1+\sqrt{2})d(y_n, x_n)+\frac{\lambda L}{\sigma}\sqrt{2}d(x_{n+1}, y_n). \tag{13}$$

Applying (13) in (11) yields

$$d(z, x_{n+1}) \le d(z, x_n)-d(x_{n+1}, y_n)-d(y_n, x_n)$$

$$+\frac{\lambda L}{\sigma} d(x_n, y_{n-1})+\frac{\lambda L}{\sigma}(1+\sqrt{2})d(y_n, x_n)+\frac{\lambda L}{\sigma}\sqrt{2}d(x_{n+1}, y_n)$$

$$\le d(z, x_n)-\left(1-\frac{\lambda L}{\sigma}\sqrt{2}\right)d(x_{n+1}, y_n)-\left(1-\frac{\lambda L}{\sigma}(1+\sqrt{2})\right)d(y_n, x_n)+\frac{\lambda L}{\sigma} d(x_n, y_{n-1}),$$

as was to be shown. ■

Let us formulate the main result of the study.

THEOREM 1. Let the set $C \subseteq E$ be convex and closed, operator $A: E \to E^*$ be pseudomonotone and Lipschitz with constant $L>0$, $S \ne \varnothing$ and $\lambda \in \left(0, (\sqrt{2}-1)\frac{\sigma}{L}\right)$. Then the sequences (x_n) and (y_n) generated by Algorithm 1 converge to some point $\bar{z} \in S$.

Proof. Let $z \in S$. Put

$$a_n = d(z, x_n)+\frac{\lambda L}{\sigma} d(x_n, y_{n-1}),$$

$$b_n = \left(1-\frac{\lambda L}{\sigma}(1+\sqrt{2})\right)(d(y_n, x_n)+d(x_{n+1}, y_n)).$$

Inequality (6) becomes $a_{n+1} \le a_n - b_n$. Then from Lemma 1 we can conclude that there exists the limit

$$\lim_{n\to\infty}\left(d(z, x_n)+\frac{\lambda L}{\sigma} d(x_n, y_{n-1})\right),$$

$$\sum_{n=1}^{\infty}\left(1-\frac{\lambda L}{\sigma}(1+\sqrt{2})\right)(d(y_n, x_n)+d(x_{n+1}, y_n))< +\infty.$$

From here we obtain

$$\lim_{n\to\infty} d(y_n, x_n) = \lim_{n\to\infty} d(x_{n+1}, y_n) = 0 \tag{14}$$

and the convergence of the numerical sequence $(d(z, x_n))$ for all $z \in S$. From (14) it follows that

$$\lim_{n\to\infty} \| y_n - x_n \| = \lim_{n\to\infty} \| x_{n+1} - y_n \| = 0, \tag{15}$$

and also

$$\lim_{n\to\infty} \| x_{n+1} - x_n \| = 0. \tag{16}$$

The inequality $d(z, x_n) \ge \frac{\sigma}{2}\| z - x_n \|^2$ and (16) yield the boundedness of sequences (x_n) and (y_n).

Consider the subsequence (x_{n_k}) converging to some point $\bar{z} \in C$. Then from (15) it follows that $y_{n_k} \to \bar{z}$ and $x_{n_k+1} \to \bar{z}$. Let us show that $\bar{z} \in S$. We have

$$(Ay_{n_k}, y - x_{n_k+1}) + \frac{1}{\lambda}(\nabla\varphi(x_{n_k+1}) - \nabla\varphi(x_{n_k}), y - x_{n_k+1}) \ge 0 \quad \forall\, y \in C. \tag{17}$$

Passing to the limit in (17) and taking into account (15) and (16) we obtain $(A\bar{z}, y - \bar{z}) \ge 0 \ \forall\, y \in C$, i.e., $\bar{z} \in C$.

Let us show that $x_n \to \bar{z}$ (then from $||x_n - y_n|| \to 0$ it follows that $y_n \to \bar{z}$). It is generally known that there exists the limit

$$\lim_{n\to\infty} d(\bar{z}, x_n) = \lim_{n\to\infty} (\varphi(\bar{z}) - \varphi(x_n) - (\nabla\varphi(x_n), \bar{z} - x_n)).$$

Since $\lim\limits_{n\to\infty} d(\bar{z}, x_{n_k}) = 0$, we get $\lim\limits_{n\to\infty} d(\bar{z}, x_n) = 0$. From here $||x_n - \bar{z}|| \to 0$. ■

Remark 6. We can specify asymptotics (14) and (15) to the following relations:

$$\underline{\lim}_{n\to\infty}\, nd(y_n, x_n) = \underline{\lim}_{n\to\infty}\, nd(x_{n+1}, y_n) = 0, \tag{18}$$

$$\underline{\lim}_{n\to\infty}\, \sqrt{n}||\, y_n - x_n|| = \underline{\lim}_{n\to\infty}\, \sqrt{n}||x_{n+1} - y_n|| = 0. \tag{19}$$

Indeed, if (18) is not satisfied, then $d(y_n, x_n) + d(x_{n+1}, y_n) \ge \mu n^{-1}$ for some $\mu > 0$ and all sufficiently large numbers n. Hence, the series $\sum_n (d(y_n, x_n) + d(x_{n+1}, y_n))$ diverges. We have obtained a contradiction. Formula (19) immediately follows from (18).

Remark 7. If $\sigma = 1$, then we can use the scheme

$$\begin{cases} x_{n+1} = \pi_{x_n}\left(-\dfrac{1}{3L} Ay_n\right), \\ y_{n+1} = \pi_{x_{n+1}}\left(-\dfrac{1}{3L} Ay_n\right). \end{cases}$$

Let us present some specific versions of Algorithm 1.

Consider variational inequality on a standard simplex: find

$$x \in S_m : (Ax, y - x) \ge 0 \quad \forall\, y \in S_m.$$

Choosing the Kullback–Leibler distance, we obtain the following version of the algorithm:

$$x_i^{n+1} = \frac{x_i^n \exp(-\lambda(Ay_n)_i)}{\sum_{j=1}^m x_j^n \exp(-\lambda(Ay_n)_j)}, \quad i = 1, \dots, m,$$

$$y_i^{n+1} = \frac{x_i^{n+1} \exp(-\lambda(Ay_n)_i)}{\sum_{j=1}^m x_j^{n+1} \exp(-\lambda(Ay_n)_j)}, \quad i = 1, \dots, m,$$

where $(Ay_n)_i \in \mathbb{R}$ is the ith coordinate of vector $Ay_n \in \mathbb{R}^m$, $\lambda > 0$.

Transport applications [29], machine learning, and game theory use variational inequalities on direct products of scaled simplexes

$$C = \prod_{k=1}^{p} r_k S_{m_k} \subseteq \mathbb{R}^{\sum_{k=1}^{p} m_k},$$

where $r_k S_{m_k} = \left\{ x \in \mathbb{R}^{m_k} : x_i \geq 0,\ \sum_{i=1}^{m_k} x_i = r_k \right\}$, $r_k > 0$, i.e., the problems: find

$$x \in \prod_{k=1}^{p} r_k S_{m_k} : (Ax,\, y-x) \geq 0 \ \ \forall\, y \in \prod_{k=1}^{p} r_k S_{m_k}. \tag{20}$$

Based on separable function

$$\varphi(x) = \sum_{k=1}^{p} \varphi_k(x_k) = \sum_{k=1}^{p} \sum_{i=1}^{m_k} \frac{x_{k,i}}{r_k} \ln \frac{x_{k,i}}{r_k},$$

where $x = (x_1, \dots, x_p) = \left(\underbrace{x_{1,1}, x_{1,2}, \dots, x_{1,m_1}}_{x_1}, \dots, \underbrace{x_{p,1}, x_{p,2}, \dots, x_{p,m_p}}_{x_p} \right) \in \mathbb{R}^{\sum_{k=1}^{p} m_k}$, let us construct the Bregman distance on $\prod_{k=1}^{p} r_k S_{m_k}$:

$$d(x, y) = \sum_{k=1}^{p} d_k(x_k, y_k) = \sum_{k=1}^{p} \sum_{i=1}^{m_k} \frac{x_{k,i}}{r_k} \ln \frac{x_{k,i}}{y_{k,i}}.$$

Algorithm 1 for inequality (20) with such distance becomes

$$x_{k,i}^{n+1} = r_k \frac{x_{k,i}^{n} \exp(-\lambda r_k (Ay_n)_{k,i})}{\sum_{j=1}^{m_k} x_{k,j}^{n} \exp(-\lambda r_k (Ay_n)_{k,j})},\ k = 1, \dots, p,\ i = 1, \dots, m_k,$$

$$y_{k,i}^{n+1} = r_k \frac{x_{k,i}^{n+1} \exp(-\lambda r_k (Ay_n)_{k,i})}{\sum_{j=1}^{m_k} x_{k,j}^{n+1} \exp(-\lambda r_k (Ay_n)_{k,j})},\ k = 1, \dots, p,\ i = 1, \dots, m_k,$$

where $(Ay_n)_{k,i}$ is the coordinate, with number $\sum_{t=1}^{k-1} m_t + i$, of vector $Ay_n \in \mathbb{R}^{\sum_{k=1}^{p} m_k}$, $\lambda > 0$.

Consider the smooth convex minimization problem

$$f(x) \to \min,\ x \in C,\ g_k(x) \leq 0,\ k = 1, \dots, p,$$

where $C \subseteq E$ is a convex closed set and f and g_k are convex differentiable functions. Let us introduce the Lagrange function $L(x, y) = f(x) + \sum_{k=1}^{p} y_k g_k(x)$ and consider the saddle problem: find

$$x' \in C,\ y' \in \mathbb{R}_{+}^{p} : L(x, y') \leq L(x', y') \leq L(x', y) \ \ \forall\, x \in C,\ \forall\, y \in \mathbb{R}_{+}^{p}. \tag{21}$$

Problem (21) is equivalent to the variational inequality

$$\left(\nabla f(x') + \sum_{k=1}^{p} y_k' \nabla g_k(x'), x - x' \right) - \sum_{k=1}^{p} g_k(x')(y_k - y_k') \geq 0 \ \ \forall\, x \in C,\ \forall\, y \in \mathbb{R}_{+}^{p}. \tag{22}$$

To solve (22), let us write the iterative process:

$$
\begin{cases}
x_{n+1} = \pi_{x_n}\left(-\lambda \left\{ \nabla f(\bar{x}_n) + \sum_{k=1}^{p} \bar{y}_k^n \nabla g_k(\bar{x}_n) \right\} \right), \\
y_{n+1} = [y_n + \lambda g(\bar{x}_n)]_+ , \\
\bar{x}_{n+1} = \pi_{x_{n+1}}\left(-\lambda \left\{ \nabla f(\bar{x}_n) + \sum_{k=1}^{p} \bar{y}_k^n \nabla g_k(\bar{x}_n) \right\} \right), \\
\bar{y}_{n+1} = [y_{n+1} + \lambda g(\bar{x}_n)]_+ ,
\end{cases}
$$

where $g(x) = (g_1(x), g_2(x), \ldots, g_p(x))$, $[\cdot]_+$ is the Euclidean projection onto the nonnegative orthant $\mathbb{R}_+^p$, $\pi_x : E^* \to C^0$ is mirror-prox constructed based on some Bregman distance d on C.

CONCLUSIONS

We have proposed a new version of the mirror descent method to solve variational inequalities with pseudomonotone operators. It can be interpreted as a modification of the two-stage Popov algorithm with the use of projection on feasible set in the sense of Bregman distance. As well as other mirror descent schemes, the method allows efficient account for the structure of feasible set of the problem. The main theoretical result is the proved theorem about the convergence of the method.

An obvious shortcoming of Algorithm 1 is the assumption that the Lipschitz constant of the operator is known or admits a simple estimate. Moreover, in certain problems, operators may not satisfy the global Lipschitz condition (in the majority of studies on the algorithms of solution of variational inequalities, Lipschitz operators are considered). It is important to propose a modification of Algorithm 1 with dynamic adjustment of step for variational inequalities with non-Lipschitz operator and to analyze its convergence.

We plan to consider a randomized version of Algorithm 1 and carry out the appropriate convergence analysis, which will expand the application of this version of the mirror descent method to solve huge-dimensional variational inequalities. Randomized versions of the mirror descent method constructed on the basis of the extragradient Korpelevich algorithm were studied in [27, 28]. Obtaining similar results for equilibrium programming problems is also of interest [17].

REFERENCES

1. I. V. Konnov, Combined Relaxation Methods for Variational Inequalities, Berlin–Heidelberg–New York, Springer-Verlag (2001).
2. P. I. Stetsyuk, “An approximate method of ellipsoids,” Cybern. Syst. Analysis, Vol. 39, No. 3, 435–439 (2003).
3. V. V. Semenov, “On the parallel proximal decomposition method for solving the problems of convex optimization,” J. Autom. Inform. Sci., Vol. 42, Iss. 4, 13–18 (2010).
4. V. V. Semenov, “Strongly convergent algorithms for variational inequality problem over the set of solutions of the equilibrium problems,” in: M. Z. Zgurovsky and V. A. Sadovnichiy (eds.), Continuous and Distributed Systems. Solid Mechanics and its Applications, Vol. 211, Springer Intern. Publ. (Switzerland), New York–Heidelberg (2014), pp. 131–146.
5. G. M. Korpelevich, “The extra-gradient method to find saddle points and solve other problems,” Ekonomika i Mat. Metody, Vol. 12, No. 4, 747–756 (1976).
6. E. N. Khobotov, “Modification of the extragradient method to solve variational inequalities and some optimization problems,” Zhur. Vych. Mat. Mat. Fiz., Vol. 27, No. 10, 1462–1473 (1987).
7. P. Tseng, “A modified forward-backward splitting method for maximal monotone mappings,” SIAM J. on Optimization, Vol. 38, 431–446 (2000).

8. Y. Censor, A. Gibali, and S. Reich, “The subgradient extragradient method for solving variational inequalities in Hilbert space,” J. of Optimization Theory and Applications, Vol. 148, 318–335 (2011).
9. S. I. Lyashko, V. V. Semenov, and T. A. Voitova, “Low-cost modification of Korpelevich’s methods for monotone equilibrium problems,” Cybern. Syst. Analysis, Vol. 47, No. 4, 631–639 (2011).
10. V. V. Semenov, “A strongly convergent splitting method for systems of operator inclusions with monotone operators,” J. Autom. Inform. Sci., Vol. 46, Iss. 5, 45–56 (2014).
11. V. V. Semenov, “Hybrid splitting methods for the system of operator inclusions with monotone operators,” Cybern. Syst. Analysis, Vol. 50, No. 5, 741–749 (2014).
12. D. A. Verlan, V. V. Semenov, and L. M. Chabak, “A strongly convergent modified extragradient method for variational inequalities with non-Lipschitz operators,” J. Autom. Inform. Sci., Vol. 47, Iss. 7, 31–46 (2015).
13. S. V. Denisov, V. V. Semenov, and L. M. Chabak, “Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators,” Cybern. Syst. Analysis, Vol. 51, No. 5, 757–765 (2015).
14. L. D. Popov, “Modification of the Arrow–Hurwitz method of finding saddle points,” Matem. Zametki, Vol. 28, No. 5, 777–784 (1980).
15. Yu. V. Malitsky and V. V. Semenov, “An extragradient algorithm for monotone variational inequalities,” Cybern. Syst. Analysis, Vol. 50, No. 2, 271–277 (2014).
16. Yu. V. Malitsky and V. V. Semenov, “A hybrid method without extrapolation step for solving variational inequality problems,” J. of Global Optimization, Vol. 61, Iss. 1, 193–202 (2015).
17. Ya. I. Vedel and V. V. Semenov, “A new two-stage proximal algorithm to solve the equilibrium problem,” Zhurn. Obch. ta Prykladnoi Matem., No. 1 (118), 15–23 (2015).
18. L. M. Bregman, “The relaxation method of finding common point of convex sets and its application to solve convex programming problems, Zhurn. Vych. Mat. Mat. Fiz., Vol. 7, No. 3, 620–631 (1967).
19. A. S. Nemirovski and D. B. Yudin, Complexity of Problems and Efficiency of Optimization Methods [in Russian], Nauka, Moscow (1979).
20. A. Ben-Tal, T. Margalit, and A. Nemirovski, “The ordered subsets mirror descent optimization method with applications to tomography,” SIAM J. on Optimization, Vol. 12, 79–108 (2001).
21. A. Beck and M. Teboulle, “Mirror descent and nonlinear projected subgradient methods for convex optimization,” Operations Research Letters, Vol. 31, No. 3, 167–175 (2003).
22. Z. Allen-Zhu and L. Orecchia, “Linear coupling: An ultimate unification of gradient and mirror descent,” e-print (2014), arXiv:1407.1537.
23. A. S. Anikin, A. V. Gasnikov, and A. Yu. Gornov, “Randomization and scarcity in huge-scale optimization problems on the example of the mirror descent method,” Tr. MFTI, Vol. 8, No. 1, 11–24 (2016).
24. 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. on Optimization, Vol. 15, 229–251 (2004).
25. A. Auslender and M. Teboulle, “Interior projection-like methods for monotone variational inequalities,” Mathematical Programming, Vol. 104, Iss. 1, 39–68 (2005).
26. Yu. Nesterov, “Dual extrapolation and its applications to solving variational inequalities and related problems,” Mathematical Programming, Vol. 109, Iss. 2–3, 319–344 (2007).
27. A. Juditsky, A. Nemirovski, and C. Tauvel, Solving variational inequalities with stochastic mirror-prox algorithm," Stochastic Systems, Vol. 1, No. 1, 17–58 (2011).
28. M. Baes, M. Burgisser, and A. Nemirovski, “A randomized mirror-prox method for solving structured large-scale matrix saddle-point problems,” SIAM J. on Optimization, Vol. 23, 934–962 (2013).
29. A. V. Gasnikov, A. A. Logunovskaya, and L. E. Morozova, “Relation of imitative logit dynamics in population game theory and with the mirror descent method in online optimization on the example of the shortest route problem,” Tr. MFTI, Vol. 7, No. 4, 104–113 (2015).