УДК 517.988

В.В. Семенов, С.В. Денисов, Д.С. Сирык, О.С. Харьков

СХОДИМОСТЬ МЕТОДА ЭКСТРАПОЛЯЦИИ ИЗ ПРОШЛОГО И МЕТОДА ОПЕРАТОРНОЙ ЭКСТРАПОЛЯЦИИ*

Ключевые слова: вариационное неравенство, псевдомонотонность, дивергенция Брэгмана, экстраполяции из прошлого, операторная экстраполяция, адаптивность, сходимость.

Введение

Многие актуальные задачи исследования операций и математической физики могут быть записаны в форме вариационных неравенств [1–4]. Решение последних является активно развивающимся направлением прикладного нелинейного анализа [3, 5–31]. Отметим, что часто негладкие задачи оптимизации могут эффективно решаться, если их переформулировать в виде седловых задач, а к последним применить алгоритмы решения вариационных неравенств [5]. С появлением генерирующих состязательных нейронных сетей (GAN — generative adversarial network) устойчивый интерес к алгоритмам решения вариационных неравенств возник и в среде специалистов в области машинного обучения [6].

Наиболее известный метод для решения вариационных неравенств — экстраградиентный алгоритм Корпелевич [9]. Исследованию этого алгоритма посвящено большое количество публикаций [10–17]. В частности, предлагались модификации экстраградиентного алгоритма с одним метрическим проектированием на допустимое множество [10, 11, 14]. Эффективный современный вариант экстраградиентного метода — проксимальный зеркальный метод Немировского [5]. Его можно интерпретировать как вариант экстраградиентного метода с проектированием, понимаемым в смысле дивергенции Брэгмана [18, 19]. Он позволяет иногда эффективно использовать структуру допустимого множества задачи. Например, для симплекса в качестве расстояния можно взять дивергенцию Кульбака–Лейблера (дивергенция Брэгмана, построенная по отрицательной энтропии) и получить явно вычисляемый оператор проектирования на симплекс. Также интересный метод двойственной экстраполяции для решения вариационных неравенств предложил Ю.Е. Нестеров [20]. Адаптивный вариант проксимального зеркального метода Немировского изучен в [17].

---

* Работа выполнена при финансовой поддержке МОН Украины (проект «Математичне моделювання та оптимізація динамічних систем для оборони, медицини та екології», номер госрегистрации 0219U008403) и НАН Украины (проект «Нові методи дослідження коректності та розв'язання задач дискретної оптимізації, варіаційних нерівностей та їх застосування», номер госрегистрации 0119U101608).

В начале 1980-х годов Л.Д. Попов предложил интересную модификацию классического алгоритма Эрроу–Гурвица для поиска седловых точек выпукло-вогнутых функций [21]. В работе [22] исследована модификация метода Попова для решения вариационных неравенств с монотонными операторами, а в статье [23] предложен двухэтапный проксимальный алгоритм для решения задачи равновесного программирования (задач о равновесии), являющийся адаптацией метода [21] к общим неравенствам Ки Фаня. В [24–28] исследован двухэтапный проксимальный зеркальный метод, модификация двухэтапного проксимального алгоритма [23] с использованием дивергенции Брэгмана вместо евклидового расстояния. Заметим, что в последнее время алгоритм Попова для вариационных неравенств хорошо известен в среде специалистов по машинному обучению под названием «Extrapolation from the Past» [6]. Дальнейшее развитие этого круга идей привело к появлению так называемого forward-reflected-backward algorithm [29] и близких методов [30, 31]. И снова заметим, что в сообществе специалистов по машинному обучению этот итерационный алгоритм известен как метод оптимистического градиентного спуска (OGDA — optimistic gradient descent ascent) [6].

Данная работа продолжает статьи [17, 24] и посвящена изучению трех новых алгоритмов с брэгмановской проекцией для решения вариационных неравенств в гильбертовом пространстве. Первый алгоритм — результат модификации двухэтапного брэгмановского метода [24] с помощью экономной регулировки величины шага, не требующей знания константы Липшица оператора. Второй алгоритм, называемый алгоритмом операторной экстраполяции, получен заменой в методе из [29] (forward-reflected-backward algorithm) евклидовой метрики на дивергенцию Брэгмана. Привлекательной чертой алгоритма является всего одно вычисление на итерационном шаге проекции Брэгмана на допустимое множество. Третий алгоритм — адаптивный вариант второго, где используемое правило обновления величины шага не требует знания липшицевых констант и вычислений значений оператора A в дополнительных точках. Для вариационных неравенств с псевдомонотонными, липшицевыми и секвенциально слабо непрерывными операторами, действующими в гильбертовом пространстве, доказаны теоремы о сходимости методов.

Вспомогательные сведения и постановка задачи

Всюду H — действительное гильбертово пространство со скалярным произведением $(\cdot,\cdot)$ и нормой $\|\cdot\|$.

Пусть функция $\varphi: H \to (-\infty, +\infty]$ строго выпуклая и дифференцируемая по Гато на $\operatorname{int}\operatorname{dom}\varphi$, где $\operatorname{dom}\varphi = \{x \in H : \varphi(x) < +\infty\}$. Дивергенция Брэгмана (расстояние Брэгмана), соответствующая функции φ, определяется формулой [7, 18, 19]

$$V(a, b) = \varphi(a) - \varphi(b) - (\nabla\varphi(b), a - b) \quad \forall a \in \operatorname{dom}\varphi, \ \ \forall b \in \operatorname{int}\operatorname{dom}\varphi.$$

Примеры практически важных дивергенций Брэгмана приведены в [19]. Для функции $\varphi(\cdot) = \frac{1}{2}\|\cdot\|^2$ имеем $V(x, y) = \frac{1}{2}\|x - y\|^2$. Для функции отрицательной энтропии

$$\varphi(x) = \sum_{i=1}^{m} x_i \ln x_i, \quad x \in \mathbb{R}_+^m = \{x \in \mathbb{R}^m : \ x_i \geq 0\},$$

получаем дивергенцию Кульбака–Лейблера

$$V(x, y) = \sum_{i=1}^{m} x_i \ln(x_i / y_i) - \sum_{i=1}^{m} (x_i - y_i), \quad x \in \mathbb{R}_+^m, \ \ y \in \mathbb{R}_{++}^m = \operatorname{int}(\mathbb{R}_+^m).$$

Имеет место полезное 3-точечное тождество [19]:

$$V(a,c)-V(a,b)-V(b,c)=(\nabla\varphi(b)-\nabla\varphi(c),\, a-b).$$

Пусть K — непустое, замкнутое и выпуклое подмножество $\operatorname{int}\operatorname{dom}\varphi$. Если функция φ сильно выпукла с константой $\mu>0$ на множестве K, то из определения дивергенции Брэгмана следует оценка

$$V(a,b)\geq\frac{\mu}{2}\|a-b\|^2 \quad \forall a\in K,\ \forall b\in K.$$

Рассмотрим сильно выпуклые задачи минимизации вида

$$P_x^K(a)=\arg\min_{y\in K}\{-(a,\, y-x)+V(y,x)\} \quad \forall a\in H,\ \forall x\in\operatorname{int}\operatorname{dom}\varphi. \tag{1}$$

Задача (1) имеет единственное решение $z\in K$ [19], причем

$$-(a,\, y-z)+(\nabla\varphi(z)-\nabla\varphi(x),\, y-z)\geq 0 \quad \forall y\in K.$$

Последнее неравенство, учитывая 3-точечное тождество, можно записать в виде

$$-(a,\, y-z)+V(y,x)-V(y,z)-V(z,x)\geq 0 \quad \forall y\in K.$$

Точка $P_x^K(a)$ в случае $\varphi(\cdot)=\frac{1}{2}\|\cdot\|^2$ совпадает с евклидовой метрической проекцией $P_K(x+a)=\arg\min_{y\in K}\|y-(x+a)\|$. Напомним, что проекцией Брэгмана на множество K называют оператор, $\Pi_K:\operatorname{int}\operatorname{dom}\varphi\to K$, задаваемый формулой [7, 18]

$$\Pi_K(x)=\arg\min_{y\in K}V(y,x), \quad x\in\operatorname{int}\operatorname{dom}\varphi.$$

Ясно, что для $P_x^K(a)$ имеет место представление

$$P_x^K(a)=\Pi_K((\nabla\varphi)^{-1}(\nabla\varphi(x)+a)).$$

Пусть C — непустое подмножество пространства H, A — оператор, действующий из H в H. Рассмотрим вариационное неравенство:

$$\text{найти } x\in C:\ (Ax,\, y-x)\geq 0 \quad \forall y\in C, \tag{2}$$

множество решений которого обозначим S.

Предположим, что выполнены следующие условия:

• множество $C\subseteq H$ — выпуклое и замкнутое;

• оператор $A:H\to H$ — псевдомонотонный, секвенциально слабо непрерывный и липшицевый с константой $L>0$ на C;

• множество S — не пусто.

Напомним, что оператор называют псевдомонотонным на C, если для всех $x,\ y\in C$ из $(Ax,\, y-x)\geq 0$ следует $(Ay,\, y-x)\geq 0$.

Рассмотрим дуальное вариационное неравенство:

$$\text{найти } x\in C:\ (Ay,\, x-y)\leq 0 \quad \forall y\in C. \tag{3}$$

Множество решений (3) обозначим S^d. Заметим, что множество S^d выпуклое и замкнутое [3]. Неравенство (3) иногда называют слабой или дуальной постановкой (2), а решения (3) — слабыми решениями (2) [3]. Для псевдомонотонных операторов A всегда имеем $S\subseteq S^d$. В наших же условиях — $S^d=S$ [3].

Далее будем предполагать, что выполнено следующее условие.

А1. Функция φ сильно выпукла с константой 1 на множестве $C \subseteq \operatorname{int}\operatorname{dom}\varphi$ и равномерно непрерывно дифференцируема по Фреше на ограниченных подмножествах C, причем оператор $\nabla\varphi$ секвенциально слабо непрерывен.

Опишем первый из предлагаемых алгоритмов для решения вариационного неравенства (2).

Адаптивный алгоритм экстраполяции из прошлого

Для решения вариационного неравенства (2) в конечномерном линейном пространстве в [24] предложен алгоритм

$$\begin{cases} y_n = P_{x_n}^C(-\lambda A y_{n-1}), \\ x_{n+1} = P_{x_n}^C(-\lambda A y_n), \end{cases} \tag{4}$$

где величина λ задавалась, исходя из требования $\lambda \in \left(0, \frac{\sqrt{2}-1}{L}\right)$, т.е. использовалась информация о константах липшицевости оператора A. Известно [24], что последовательности (x_n) и (y_n), порожденные (4), сходятся к решению (1). В случае компактности множества $C \subseteq E$ и монотонности оператора A справедлива оценка

$$G(z_N) = \max_{y \in C}(Ay, z_N - y) = O\left(\frac{1}{N}\right),$$

где $z_N = \frac{1}{N}\sum_{n=1}^{N} y_n$ — усредненный выход работы алгоритма (4) [27].

Отталкиваясь от итерационной схемы (4) и работ [14, 16, 17], ниже предлагаем двухэтапный брэгмановский алгоритм, не требующий знания липшицевых констант и вычислений значений оператора A в дополнительных точках (т.е. без процедур типа линейного поиска). Опишем его подробнее.

Алгоритм 1. Экстраполяция из прошлого с адаптивной регулировкой. Выбираем элемент $y_0 \in H$, $x_1 \in \operatorname{int}\operatorname{dom}\varphi$, $\tau \in \left(0, \frac{1}{3}\right)$ и число $\lambda_1 > 0$. Полагаем $n = 1$.

1. Вычислить

$$y_n = P_{x_n}^C(-\lambda_n A y_{n-1}).$$

2. Вычислить

$$x_{n+1} = P_{x_n}^C(-\lambda_n A y_n).$$

Если $y_n = x_n = x_{n+1}$, то СТОП, иначе перейти к 3.

3. Вычислить

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau\frac{\|y_n - y_{n-1}\|}{\|Ay_n - Ay_{n-1}\|}\right\}, & \text{если } Ay_{n-1} \neq Ay_n, \\ \lambda_n, & \text{иначе.} \end{cases}$$

Положить $n := n+1$ и перейти к 1.

Докажем сходимость алгоритма 1 для вариационных неравенств с псевдомонотонными операторами, действующими в общих гильбертовых пространствах.

Слабая сходимость алгоритма 1. Прежде всего отметим, что задаваемая правилом пересчета последовательность (λ_n) невозрастающая и ограничена снизу числом $\min\{\lambda_1, \tau L^{-1}\}$. Следовательно, существует $\lim\limits_{n\to\infty} \lambda_n > 0$.

Для порожденных алгоритмом 1 последовательностей (x_n) и (y_n) имеют место неравенства

$$-\lambda_n (Ay_{n-1}, y-y_n) \le V(y, x_n) - V(y_n, x_n) - V(y, y_n) \quad \forall y \in C, \tag{5}$$

$$-\lambda_n (Ay_n, y-x_{n+1}) \le V(y, x_n) - V(x_{n+1}, x_n) - V(y, x_{n+1}) \quad \forall y \in C. \tag{6}$$

Неравенство (6) обосновывает правила остановки алгоритма 1. Действительно, при $x_{n+1} = x_n = y_n$ из (6) вытекает $(Ax_n, y-x_n) \ge 0$ для всех $y \in C$, т.е., $x_n \in S$.

Лемма 1. Для порожденных алгоритмом 1 последовательностей (x_n) и (y_n) имеет место неравенство

$$V(z, x_{n+1}) \le V(z, x_n) - \left(1 - 2\tau \frac{\lambda_n}{\lambda_{n+1}}\right) V(x_{n+1}, y_n) -$$

$$- \left(1 - \tau \frac{\lambda_n}{\lambda_{n+1}}\right) V(y_n, x_n) + \tau \frac{\lambda_n}{\lambda_{n+1}} V(x_n, y_{n-1}), \tag{7}$$

где $z \in S$.

Доказательство. Пусть $z \in S$. Из (6) следует

$$V(z, x_{n+1}) \le V(z, x_n) - V(x_{n+1}, x_n) + \lambda_n (Ay_n, z - x_{n+1}) =$$

$$= V(z, x_n) - V(x_{n+1}, x_n) + \lambda_n (Ay_n, y_n - x_{n+1}) + \lambda_n (Ay_n, z - y_n).$$

Учтя псевдомонотонность оператора A, получим

$$V(z, x_{n+1}) \le V(z, x_n) - V(x_{n+1}, x_n) + \lambda_n (Ay_n, y_n - x_{n+1}). \tag{8}$$

Из (5) следует

$$-\lambda_n (Ay_{n-1}, x_{n+1} - y_n) \le V(x_{n+1}, x_n) - V(y_n, x_n) - V(x_{n+1}, y_n). \tag{9}$$

Оценив сверху $-V(x_{n+1}, x_n)$ в (8) с помощью (9), получим

$$V(z, x_{n+1}) \le V(z, x_n) - V(x_{n+1}, y_n) - V(y_n, x_n) + \lambda_n (Ay_n - Ay_{n-1}, y_n - x_{n+1}). \tag{10}$$

Используя правило вычисления λ_{n+1}, оценим сверху слагаемое $\lambda_n (Ay_n - Ay_{n-1}$, $y_n - x_{n+1})$. Имеем

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

$$\le \tau \frac{\lambda_n}{\lambda_{n+1}} \| y_n - y_{n-1} \| \| x_{n+1} - y_n \| \le$$

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

$$\le \tau \frac{\lambda_n}{\lambda_{n+1}} V(y_n, x_n) + \tau \frac{\lambda_n}{\lambda_{n+1}} V(x_n, y_{n-1}) + 2\tau \frac{\lambda_n}{\lambda_{n+1}} V(x_{n+1}, y_n). \tag{11}$$

Применив (11) в (10), получим (7). ■

Докажем первый из основных результатов работы.

 ISSN 0572-2691

Теорема 1. Пусть множество $C \subseteq H$ выпуклое и замкнутое, оператор $A: H \to H$ псевдомонотонный, липшицевый с константой $L > 0$ и секвенциально слабо непрерывный, $S \neq \varnothing$. Предположим, что выполнено A1. Тогда последовательности (x_n) и (y_n), порожденные алгоритмом 1, слабо сходятся к некоторой точке $z \in S$.

Доказательство. Пусть $z' \in S$. Положим

$$a_n = V(z', x_n) + \tau \frac{\lambda_n}{\lambda_{n+1}} V(x_n, y_{n-1}),$$

$$b_n = \left(1 - \tau \frac{\lambda_n}{\lambda_{n+1}}\right) V(y_n, x_n) + \left(1 - \tau \frac{\lambda_{n+1}}{\lambda_{n+2}} - 2\tau \frac{\lambda_n}{\lambda_{n+1}}\right) V(x_{n+1}, y_n).$$

Неравенство (7) принимает вид $a_{n+1} \le a_n - b_n$. Поскольку существует $\lim\limits_{n\to\infty} \lambda_n > 0$, то

$$1 - \tau \frac{\lambda_n}{\lambda_{n+1}} \to 1 - \tau \in (0,1) \text{ и } 1 - \tau \frac{\lambda_{n+1}}{\lambda_{n+2}} - 2\tau \frac{\lambda_n}{\lambda_{n+1}} \to 1 - 3\tau \in (0,1) \text{ при } n \to \infty.$$

Можем сделать вывод, что существует предел

$$\lim_{n\to\infty} \left(V(z', x_n) + \tau \frac{\lambda_n}{\lambda_{n+1}} V(x_n, y_{n-1}) \right)$$

и

$$\sum_{n=1}^{\infty} \left(\left(1 - \tau \frac{\lambda_n}{\lambda_{n+1}}\right) V(y_n, x_n) + \left(1 - \tau \frac{\lambda_{n+1}}{\lambda_{n+2}} - 2\tau \frac{\lambda_n}{\lambda_{n+1}}\right) V(x_{n+1}, y_n) \right) < +\infty.$$

Отсюда получаем

$$\lim_{n\to\infty} V(y_n, x_n) = \lim_{n\to\infty} V(x_{n+1}, y_n) = 0 \tag{12}$$

и сходимость числовых последовательностей $(V(z', x_n))$ для всех $z' \in S$. Из (12) следует

$$\lim_{n\to\infty} \| y_n - x_n \| = \lim_{n\to\infty} \| x_{n+1} - y_n \| = 0. \tag{13}$$

Из неравенства $V(z', x_n) \ge \frac{1}{2} \| z' - x_n \|^2$ и (13) следует ограниченность последовательностей (x_n), (y_n), а из неравенств (8), (9) и предельных соотношений (13) — $\lim\limits_{n\to\infty} V(x_{n+1}, x_n) = 0$ и естественно

$$\lim_{n\to\infty} \| x_{n+1} - x_n \| = 0. \tag{14}$$

Рассмотрим подпоследовательность (x_{n_k}), слабо сходящуюся к некоторой точке $z \in H$. Ясно, что $z \in C$. Из соотношений (13), (14) следует, что (y_{n_k}), (x_{n_k+1}) слабо сходятся к z. Покажем, что $z \in S$. Предположим, что $Az \neq 0$ (в противном случае доказывать нечего, т.е. $z \in S$). Имеем

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

Отсюда

$$(Ay_{n_k}, y - y_{n_k}) + (Ay_{n_k}, y_{n_k} - x_{n_k+1}) + \\ + \frac{1}{\lambda_{n_k+1}} (\nabla \varphi(x_{n_k+1}) - \nabla \varphi(x_{n_k}), y - x_{n_k+1}) \ge 0 \quad \forall y \in C. \tag{15}$$

Совершив в (15) предельный переход при $k \to \infty$ с учетом равномерной непрерывности $\nabla\varphi$ на ограниченных подмножествах C и (13), (14), получим

$$\varliminf_{k\to\infty} (Ay_{n_k}, y - y_{n_k}) \geq 0. \tag{16}$$

Из (16) следует существование для каждого $k \in \mathbb{N}$ такого наименьшего числа $m_k \in \mathbb{N}$, что

$$(Ay_{n_i}, y - y_{n_i}) \geq -2^{-k} \text{ для всех } i \geq m_k.$$

Последовательность (m_k) возрастающая. Из секвенциальной слабой непрерывности оператора A и слабой полунепрерывности нормы получаем

$$0 < \| Az \| \leq \varliminf_{k\to\infty} \| Ay_{n_{m_k}} \|. \tag{17}$$

Можно считать, что $Ay_{n_{m_k}} \neq 0$. Положим $v_k = \| Ay_{n_{m_k}} \|^{-2} Ay_{n_{m_k}}$. Очевидно, что $(Ay_{n_{m_k}}, v_k) = 1$ для всех $k \in \mathbb{N}$. Имеем

$$(Ay_{n_{m_k}}, y + 2^{-k} v_k - y_{n_{m_k}}) \geq 0.$$

Из псевдомонотонности оператора A следует

$$(A(y + 2^{-k} v_k), y + 2^{-k} v_k - y_{n_{m_k}}) \geq 0.$$

Отсюда получаем

$$(Ay, y - y_{n_{m_k}}) \geq (Ay - A(y + 2^{-k} v_k), y + 2^{-k} v_k - y_{n_{m_k}}) - 2^{-k} (Ay, v_k). \tag{18}$$

Из (17) следует ограниченность сверху последовательности $\| v_k \| = \| Ay_{n_{m_k}} \|^{-1}$. Итак, $\lim_{k\to 0} 2^{-k} \| v_k \| = 0$. Совершив предельный переход в (18), получим

$$(Ay, y - z) = \lim_{k\to\infty} (Ay, y - y_{n_{m_k}}) = \varliminf_{k\to\infty} (Ay, y - y_{n_{m_k}}) \geq 0 \quad \forall y \in C.$$

Следовательно, $z \in S$.

Покажем, что последовательность (x_n) слабо сходится к z. Тогда из (13) следует и слабая сходимость последовательности (y_n) к z. Рассуждаем от противного: пусть существует подпоследовательность (x_{m_k}) такая, что $x_{m_k} \to z'$ слабо и $z \neq z'$. Ясно, что $z' \in S$ и $V(z, z') > 0$. Последовательность $(V(z, x_n))$ сходится. Значит, используя 3-точечное тождество и A1, получаем

$$\lim_{n\to\infty} V(z, x_n) = \lim_{k\to\infty} V(z, x_{m_k}) = \varliminf_{k\to\infty} V(z, x_{m_k}) =$$

$$= \varliminf_{k\to\infty} \{V(z, z') + V(z', x_{m_k}) + (\nabla\varphi(z') - \nabla\varphi(x_{m_k}), z - z')\} \geq V\left(z, z'\right) +$$

$$+ \varliminf_{k\to\infty} V(z', x_{m_k}) + \varliminf_{k\to\infty} (\nabla\varphi(z') - \nabla\varphi(x_{m_k}), z - z') > \varliminf_{k\to\infty} V(z', x_{m_k}) = \lim_{n\to\infty} V(z', x_n).$$

Поменяв местами z и z', взяв вместо (x_{m_k}) подпоследовательность (x_{n_k}), слабо сходящуюся к z, приходим к неравенству

$$\lim_{n\to\infty} V(z, x_n) > \lim_{n\to\infty} V(z', x_n) > \lim_{n\to\infty} V(z, x_n),$$

что невозможно. Таким образом, $z = z'$. ■

Замечание 1. Если оператор A монотонный, то теорема 1 справедлива без предположения о секвенциальной слабой непрерывности оператора A. Это имеет место и для теорем 2, 3, приведенных ниже.

Алгоритм операторной экстраполяции

Рассмотрим новый алгоритм, являющийся модификацией так называемого forward-reflected-backward algorithm, предложенного в [29] для решения операторных включений с суммой максимального монотонного и липшицевого монотонного операторов. Вместо евклидовой метрики и проекции используем дивергенцию Брэгмана и соответствующую проекцию. Привлекательными чертами данного алгоритма является всего одно вычисление значения оператора $P_{x_n}^C$ на итерационном шаге и построение одной последовательности аппроксимаций (x_n), а не двух, как в алгоритме 1.

Алгоритм 2. Операторная экстраполяция. Выбираем элемент $x_0 \in H$, $x_1 \in \operatorname{int}\operatorname{dom}\varphi$, последовательность (λ_n), удовлетворяющую условию $\{\inf_n \lambda_n, \sup_n \lambda_n\} \subseteq \left(0, \frac{1}{2L}\right)$. Полагаем $n = 1$.

1. Вычислить

$$x_{n+1} = P_{x_n}^C(-\lambda_n A x_n - \lambda_{n-1}(Ax_n - Ax_{n-1})). \tag{19}$$

2. Если $x_{n-1} = x_n = x_{n+1}$, то СТОП, иначе перейти к 1.

Замечание 2. Воспользовавшись проекцией Брэгмана, формулу (19) запишем в виде:

$$x_{n+1} = \Pi_C((\nabla\varphi)^{-1}(\nabla\varphi(x_n) - \lambda_n Ax_n - \lambda_{n-1}(Ax_n - Ax_{n-1}))).$$

Замечание 3. Частным случаем алгоритма 2 является популярный среди специалистов в области машинного обучения алгоритм оптимистического градиентного спуска (OGDA — optimistic gradient descent ascent) [6].

Замечание 4. При отсутствии ограничений $(C = H)$ и с $\varphi(\cdot) = \frac{1}{2}\|\cdot\|^2$ алгоритм операторной экстраполяции совпадает с алгоритмом экстраполяции из прошлого.

Для последовательности (x_n), порожденной алгоритмом 2, имеет место неравенство

$$-(\lambda_n Ax_n + \lambda_{n-1}(Ax_n - Ax_{n-1}), y - x_{n+1}) \le$$
$$\le V(y, x_n) - V(x_{n+1}, x_n) - V(y, x_{n+1}) \quad \forall y \in C. \tag{20}$$

Неравенство (20) дает обоснование правила остановки алгоритма 2. Действительно, при $x_{n-1} = x_n = x_{n+1}$ из (20) вытекает $(Ax_n, y - x_n) \ge 0$ для всех $y \in C$, т.е. $x_n \in S$.

Перейдем к доказательству сходимости алгоритма 2.

Лемма 2. Для последовательности (x_n), порожденной алгоритмом 2, имеет место неравенство

$$V(z, x_{n+1}) + \lambda_n(Ax_n - Ax_{n+1}, x_{n+1} - z) + \lambda_n L V(x_{n+1}, x_n) \le$$
$$\le V(z, x_n) + \lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z) + \lambda_{n-1} L V(x_n, x_{n-1}) -$$
$$-(1 - \lambda_{n-1}L - \lambda_n L) V(x_{n+1}, x_n),$$

где $z \in S$.

Доказательство. Пусть $z \in S$. Имеем

$$V(z, x_{n+1}) \le V(z, x_n) - V(x_{n+1}, x_n) + (\lambda_n Ax_n + \lambda_{n-1}(Ax_n - Ax_{n-1}), z - x_{n+1}). \quad (21)$$

Воспользуемся псевдомонотонностью оператора A. Имеем

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

$$+\lambda_{n-1}(Ax_n - Ax_{n-1}, z - x_{n+1}) + \underbrace{\lambda_n (Ax_{n+1}, z - x_{n+1})}_{\le 0} \le \lambda_n (Ax_n - Ax_{n+1}, z - x_{n+1}) +$$

$$+\lambda_{n-1}(Ax_n - Ax_{n-1}, z - x_n) + \lambda_{n-1}(Ax_n - Ax_{n-1}, x_n - x_{n+1}). \quad (22)$$

Применив (22) в (21), получим

$$V(z, x_{n+1}) \le V(z, x_n) - V(x_{n+1}, x_n) + \lambda_n (Ax_n - Ax_{n+1}, z - x_{n+1}) +$$

$$+\lambda_{n-1}(Ax_n - Ax_{n-1}, z - x_n) + \lambda_{n-1}(Ax_n - Ax_{n-1}, x_n - x_{n+1}). \quad (23)$$

Оценим сверху слагаемое $\lambda_{n-1}(Ax_n - Ax_{n-1}, x_n - x_{n+1})$. Имеем

$$\lambda_{n-1}(Ax_n - Ax_{n-1}, x_n - x_{n+1}) \le \lambda_{n-1} \| Ax_n - Ax_{n-1} \| \| x_n - x_{n+1} \| \le$$

$$\le \lambda_{n-1} L \| x_n - x_{n-1} \| \| x_n - x_{n+1} \| \le \frac{\lambda_{n-1}L}{2} \| x_n - x_{n-1} \|^2 + \frac{\lambda_{n-1}L}{2} \| x_n - x_{n+1} \|^2 \le$$

$$\le \lambda_{n-1} L V(x_n, x_{n-1}) + \lambda_{n-1} L V(x_{n+1}, x_n). \quad (24)$$

Применив (24) в (23), приходим к неравенству

$$V(z, x_{n+1}) + \lambda_n (Ax_n - Ax_{n+1}, x_{n+1} - z) + \lambda_n L V(x_{n+1}, x_n) \le$$

$$\le V(z, x_n) + \lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z) + \lambda_{n-1} L V(x_n, x_{n-1}) -$$

$$-(1 - \lambda_{n-1}L - \lambda_n L) V(x_{n+1}, x_n),$$

что и требовалось доказать. ■

Имеет место следующая теорема.

Теорема 2. Пусть множество $C \subseteq H$ выпуклое и замкнутое, оператор $A : H \to H$ псевдомонотонный, липшицевый с константой $L > 0$ и секвенциально слабо непрерывный, $S \ne \varnothing$. Предположим, что выполнено A1 и последовательность (λ_n) удовлетворяет условию $\{\inf_n \lambda_n, \sup_n \lambda_n\} \subseteq \left(0, \frac{1}{2L}\right)$. Тогда порожденная алгоритмом 2 последовательность (x_n) слабо сходятся к некоторой точке $z \in S$.

Доказательство. Пусть $z' \in S$. Берем такое $\delta > 0$, что $1 - \lambda_{n-1}L - \lambda_n L \ge \delta$ для всех $n \in \mathbb{N}$. Положим

$$a_n = V(z', x_n) + \lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z') + \lambda_{n-1} L V(x_n, x_{n-1}), \ b_n = \delta V(x_{n+1}, x_n).$$

Из леммы 2 вытекает неравенство $a_{n+1} \le a_n - b_n$. Покажем, что $a_n \ge 0$ для всех $n \in \mathbb{N}$. Имеем

$$a_n = V(z', x_n) + \lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z') + \lambda_{n-1} L V(x_n, x_{n-1}) \ge$$

$$\ge \frac{1}{2} \| x_n - z' \|^2 - \lambda_{n-1} \| Ax_{n-1} - Ax_n \| \| x_n - z' \| + \frac{\lambda_{n-1}L}{2} \| x_{n-1} - x_n \|^2 \ge$$

$$\ge \frac{1}{2} \| x_n - z' \|^2 - \lambda_{n-1} L \| x_{n-1} - x_n \| \| x_n - z' \| + \frac{\lambda_{n-1}L}{2} \| x_{n-1} - x_n \|^2 \ge$$

$$\ge \frac{1 - \lambda_{n-1}L}{2} \| x_n - z' \|^2 \ge 0.$$

Теперь можем сделать вывод, что существует предел

$$\lim_{n\to\infty}(V(z', x_n)+\lambda_{n-1}(Ax_{n-1}-Ax_n, x_n-z')+\lambda_{n-1}LV(x_n, x_{n-1}))$$

и $\sum_{n=1}^{\infty} V(x_{n+1}, x_n) < +\infty.$ Отсюда получаем ограниченность последовательности (x_n) и

$$\lim_{n\to\infty} V(x_{n+1}, x_n) = \lim_{n\to\infty} \|x_{n+1}-x_n\| = 0. \tag{25}$$

Поскольку

$$\lim_{n\to\infty}(\lambda_{n-1}(Ax_{n-1}-Ax_n, x_n-z')+\lambda_{n-1}LV(x_n, x_{n-1}))=0,$$

имеет место сходимость числовых последовательностей $(V(z', x_n))$ для всех $z' \in S$.

Покажем, что все слабые частичные пределы последовательности (x_n) принадлежат множеству S. Рассмотрим подпоследовательность (x_{n_k}), слабо сходящуюся к некоторой точке $z \in H$. Ясно, что $z \in C$. Покажем, что $z \in S$. Предположим, что $Az \neq 0$. Имеем

$$(Ax_{n_k}, y-x_{n_k}) + (Ax_{n_k}, x_{n_k}-x_{n_k+1}) + \frac{\lambda_{n_k-1}}{\lambda_{n_k}}(Ax_{n_k}-Ax_{n_k-1}, y-x_{n_k+1}) +$$

$$+\frac{1}{\lambda_{n_k}}(\nabla\varphi(x_{n_k+1})-\nabla\varphi(x_{n_k}), y-x_{n_k+1}) \geq 0 \quad \forall y \in C. \tag{26}$$

Совершив в (26) предельный переход при $k \to \infty$ с учетом равномерной непрерывности $\nabla\varphi$ на ограниченных подмножествах C, липшицевости оператора A и (25), получим

$$\varliminf_{k\to\infty}(Ax_{n_k}, y-x_{n_k}) \geq 0 \quad \forall y \in C.$$

Далее, рассуждая, как в доказательстве теоремы 1, приходим к включению $z \in S$.

Слабая сходимость всей последовательности (x_n) к элементу z следует из сходимости числовых последовательностей $(V(z', x_n))$ для всех $z' \in S$ и принадлежности множеству S всех слабых частичных пределов последовательности (x_n). ■

Алгоритм 2 использует информацию о константе Липшица оператора A. Рассмотрим ниже вариант алгоритма 2 с простым правилом обновления параметров λ_n без знания константы Липшица.

Адаптивный алгоритм операторной экстраполяции

Предлагаем следующий адаптивный вариант алгоритма 2.

Алгоритм 3. Операторная экстраполяция с адаптивной регулировкой. Выбираем элемент $x_0 \in H$, $x_1 \in \operatorname{int}\operatorname{dom}\varphi$, $\tau \in \left(0, \frac{1}{2}\right)$ и числа $\lambda_0, \lambda_1 > 0$. Полагаем $n = 1$.

1. Вычислить

$$x_{n+1} = P_{x_n}^{C}(-\lambda_n Ax_n - \lambda_{n-1}(Ax_n - Ax_{n-1})).$$

2. Если $x_{n-1} = x_n = x_{n+1}$, то СТОП, иначе перейти к 3.

3. Вычислить

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

Положить $n := n+1$ и перейти к 1.

Задаваемая правилом пересчета последовательность (λ_n) невозрастающая и ограничена снизу числом $\min\{\lambda_1, \tau L^{-1}\}$. Следовательно, существует $\lim\limits_{n\to\infty} \lambda_n > 0$.

Лемма 3. Для порожденной алгоритмом 3 последовательности (x_n) имеет место неравенство

$$V(z, x_{n+1}) + \lambda_n (Ax_n - Ax_{n+1}, x_{n+1} - z) + \tau \frac{\lambda_n}{\lambda_{n+1}} V(x_{n+1}, x_n) \leq$$

$$\leq V(z, x_n) + \lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z) + \tau \frac{\lambda_{n-1}}{\lambda_n} V(x_n, x_{n-1}) -$$

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

где $z \in S$.

Доказательство. Пусть $z \in S$. Как и в доказательстве леммы 2, приходим к неравенству

$$V(z, x_{n+1}) \leq V(z, x_n) - V(x_{n+1}, x_n) + \lambda_n (Ax_n - Ax_{n+1}, z - x_{n+1}) +$$

$$+ \lambda_{n-1}(Ax_n - Ax_{n-1}, z - x_n) + \lambda_{n-1}(Ax_n - Ax_{n-1}, x_n - x_{n+1}).$$

Используя правило пересчета λ_n, оценим сверху слагаемое $\lambda_{n-1}(Ax_n - Ax_{n-1}$, $x_n - x_{n+1})$. Имеем

$$\lambda_{n-1}(Ax_n - Ax_{n-1}, x_n - x_{n+1}) \leq \lambda_{n-1} \| Ax_n - Ax_{n-1} \| \| x_n - x_{n+1} \| \leq$$

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

$$\leq \tau \frac{\lambda_{n-1}}{\lambda_n} V(x_n, x_{n-1}) + \tau \frac{\lambda_{n-1}}{\lambda_n} V(x_{n+1}, x_n).$$

Приходим к неравенству

$$V(z, x_{n+1}) + \lambda_n (Ax_n - Ax_{n+1}, x_{n+1} - z) + \tau \frac{\lambda_n}{\lambda_{n+1}} V(x_{n+1}, x_n) \leq$$

$$\leq V(z, x_n) + \lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z) + \tau \frac{\lambda_{n-1}}{\lambda_n} V(x_n, x_{n-1}) -$$

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

что и требовалось доказать. ■

Сформулируем и докажем теорему о сходимости алгоритма 3.

Теорема 3. Пусть множество $C \subseteq H$ — выпуклое и замкнутое, оператор $A : H \to H$ — псевдомонотонный, липшицевый с константой $L > 0$ и секвенциально слабо непрерывный, $S \neq \varnothing$. Предположим, что выполнено предположение A1. Тогда порожденная алгоритмом 3 последовательность (x_n) слабо сходится к некоторой точке $z \in S$.

Доказательство. Пусть $z' \in S$. Положим

ISSN 0572-2691

$$a_n = V(z', x_n) + \lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z') + \tau \frac{\lambda_{n-1}}{\lambda_n} V(x_n, x_{n-1}),$$

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

Неравенство леммы 3 принимает вид $a_{n+1} \le a_n - b_n$. Поскольку существует $\lim\limits_{n\to\infty} \lambda_n > 0$, то

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

Покажем, что $a_n \ge 0$ для всех достаточно больших $n \in \mathbb{N}$. Имеем

$$a_n = V(z', x_n) + \lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z') + \tau \frac{\lambda_{n-1}}{\lambda_n} V(x_n, x_{n-1}) \ge$$

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

$$\ge \frac{1}{2}\|x_n - z'\|^2 - \tau \frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|\|x_n - z'\| + \tau \frac{\lambda_{n-1}}{2\lambda_n}\|x_{n-1} - x_n\|^2 \ge$$

$$\ge \frac{1 - \tau\lambda_{n-1}\lambda_n^{-1}}{2}\|x_n - z'\|^2.$$

Поскольку существует такое $n_0 \in \mathbb{N}$, что $1 - \tau\lambda_{n-1}\lambda_n^{-1} > 0$ для всех $n \ge n_0$, то $a_n \ge 0$ начиная с n_0.

Теперь можем сделать вывод, что существует предел

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

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

Отсюда получаем ограниченность последовательности (x_n) и $\lim\limits_{n\to\infty} V(x_{n+1}, x_n) =$ $= \lim\limits_{n\to\infty}\|x_{n+1} - x_n\| = 0.$ Поскольку

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

сходятся последовательности $(V(z', x_n))$ для всех $z' \in S$.

Далее, повторяя рассуждения доказательства теоремы 2, приходим к результату. ■

Замечание 5. Правило обновления на шаге 3 алгоритма 3 можно поменять на следующее

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau \frac{\sqrt{2V\left(x_{n+1}, x_n\right)}}{\|Ax_{n+1} - Ax_n\|}\right\}, & \text{если } Ax_{n+1} \ne Ax_n, \\ \lambda_n & \text{иначе.} \end{cases} \tag{27}$$

Теорема 3 справедлива и для варианта алгоритма 3 с правилом (27).

Заключение

В статье изучены три новых алгоритма с брэгмановской проекцией для приближенного решения вариационных неравенств.

Первый алгоритм получен введением новой адаптивной регулировки величины шага в двухэтапный брэгмановский метод [24]. Второй алгоритм получен заменой в методе из [29] (forward-reflected-backward algorithm) евклидовой метрики на дивергенцию Брэгмана. Частным случаем этого алгоритма является популярный среди специалистов в области машинного обучения алгоритм оптимистического градиентного спуска (OGDA) [6]. Третий алгоритм является адаптивным вариантом второго. Привлекательная черта второго и третьего алгоритмов — необходимость на итерационном шаге всего одного вычисления проекции Брэгмана на допустимое множество. Используемые правила обновления величины шага не требуют знания липшицевых констант и вычислений значений оператора A в дополнительных точках (т.е. не используются процедуры типа линейного поиска).

Для вариационных неравенств с псевдомонотонными, липшицевыми и секвенциально слабо непрерывными операторами, действующими в гильбертовом пространстве, доказаны теоремы о слабой сходимости методов.

В одной из ближайших работ мы планируем рассмотреть мультиблочный вариант адаптивного метода операторной экстраполяции. Интересной является задача анализа сходимости распределенных и/или стохастических версий алгоритмов, например с редукцией дисперсии.

В.В. Семенов, С.В. Денисов, Д.С. Сірик, О.С. Харьков

ЗБІЖНІСТЬ МЕТОДУ ЕКСТРАПОЛЯЦІЇ З МИНУЛОГО ТА МЕТОДУ ОПЕРАТОРНОЇ ЕКСТРАПОЛЯЦІЇ

Одним з популярних напрямів сучасного прикладного нелінійного аналізу є дослідження варіаційних нерівностей. Багато актуальних проблем дослідження операцій і математичної фізики можна записати у формі варіаційних нерівностей. З появою генеруючих змагальних нейронних мереж інтерес до алгоритмів розв'язання варіаційних нерівностей виник і в середовищі фахівців машинного навчання. Дана робота присвячена дослідженню трьох нових алгоритмів з брегманівською проекцією для розв'язання варіаційних нерівностей в гільбертовому просторі. Перший алгоритм — результат модифікації двоетапного брегманівського методу за допомогою економного регулювання величини кроку, що не вимагає знання ліпшицевої константи оператора. Другий алгоритм — алгоритм операторної екстраполяції, отриманий заміною в методі Маліцького–Тама евклідової метрики на дивергенцію Брегмана. Приваблива риса алгоритму — всього одне обчислення на ітераційному кроці проекції Брегмана на допустиму множину. Третій алгоритм — адаптивний варіант другого, де використовується правило поновлення величини кроку, що не вимагає знання ліпшицевих констант і обчислень значень оператора в додаткових точках. Для варіаційних нерівностей з псевдомонотонними, ліпшицевими та секвенційно слабко неперервними операторами, що діють в гільбертовому просторі, доведено теореми про збіжність методів.

Ключові слова: варіаційна нерівність, псевдомонотонність, дивергенція Брегмана, екстраполяція з минулого, операторна екстраполяція, адаптивність, збіжність.

V.V. Semenov, S.V. Denisov, D.S. Siryk, O.S. Kharkov

CONVERGENCE OF THE EXTRAPOLATION METHOD FROM THE PAST AND THE OPERATOR EXTRAPOLATION METHOD

One of the popular areas of modern applied nonlinear analysis is the study of variational inequalities. Many important problems of operations research and mathematical physics can be written in the form of variational inequalities. With the advent of generating adversarial neural networks, interest in algorithms for solving variational inequalities arose in the ML-community. This paper is devoted to the study of three new algorithms with Bregman projection for solving variational inequalities in Hilbert space. The first algorithm is the result of a modification of the two-stage Bregman method by low-cost adjusting the step size that without the prior knowledge of the Lipschitz constant of operator. The second algorithm, which we call the operator extrapolation algorithm, is obtained by replacing the Euclidean metric in the Malitsky–Tam method with the Bregman divergence. An attractive feature of the algorithm is only one computation at the iterative step of the Bregman projection onto the feasible set. The third algorithm is an adaptive version of the second, where the used rule for updating the step size does not require knowledge of Lipschitz constants and the calculation of operator values at additional points. For variational inequalities with pseudo-monotone, Lipschitz-continuous, and sequentially weakly continuous operators acting in a Hilbert space, convergence theorems are proved.

Keywords: variational inequality, pseudo-monotonicity, Bregman divergence, extrapolation from the past, operator extrapolation, adaptivity, convergence.

1. Kinderlehrer D. Stampacchia G. An introduction to variational inequalities and their applications. New York : Academic Press, 1980. Russian transl. M. : Mir, 1983. 256 p.
2. Nagurney A. Network economics: A variational inequality approach. Dordrecht : Kluwer Academic Publishers, 1999. 325 p. doi: https://doi.org/10.1007/978-1-4757-3005-0.
3. Konnov I.V. Combined relaxation methods for variational inequalities. Berlin; Heidelberg; New York : Springer-Verlag, 2001. 181 p. doi: https://doi.org/10.1007/978-3-642-56886-2.
4. Lyashko S.I., Klyushin D.A., Nomirovsky D.A., Semenov V.V. Identification of age-structured contamination sources in ground water. R. Boucekkline, N. Hritonenko, and Y. Yatsenko (eds.) Optimal Control of Age-Structured Populations in Economy, Demography, and the Environment, Routledge, London; New York, 2013. P. 277–292.
5. Nemirovski A. 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 Journal on Optimization.* 2004. **15.** P. 229–251. https://doi.org/10.1137/S1052623403425629.
6. Gidel G., Berard H., Vincent P., Lacoste-Julien S. A variational inequality perspective on generative adversarial networks. arXiv preprint arXiv. 2018. 1802.10551.
7. Reich S., Sabach S. A strong convergence theorem for a proximal-type algorithm in reflexive Banach spaces. *J. Nonlinear Convex Anal.* 2009. **10**. P. 471–485.
8. Semenov V.V. On the parallel proximal decomposition method for solving the problems of convex optimization. *Journal of Automation and Information Sciences.* 2010. **42**, N 4. P. 13–18. https://doi.org/10.1615/JAutomatInfScien.v42.i4.20.
9. Korpelevich G.M. An extragradient method for finding saddle points and for other problems. *Matecon.* 1976. **12**, N 4. P. 747–756.
10. Tseng P. A modified forward-backward splitting method for maximal monotone mappings. *SIAM Journal on Control and Optimization.* 2000. **38**. P. 431–446. https://doi.org/10.1137/S0363012998338806.
11. Censor Y., Gibali A., Reich S. The subgradient extragradient method for solving variational inequalities in Hilbert space. *Journal of Optimization Theory and Applications.* 2011. **148**. P. 318–335. https://doi.org/10.1007/s10957-010-9757-3.

12. Verlan D.A., Semenov V.V., Chabak L.M. A strongly convergent modified extragradient method for variational inequalities with non-lipschitz operators. *Journal of Automation and Information Sciences*. 2015. **47**, N 7. P. 31–46. https://doi.org/10.1615/JAutomatInfScien.v47.i7.40.
13. Semenov V.V. Modified extragradient method with Bregman divergence for variational inequalities. *Journal of Automation and Information Sciences*. 2018. **50**, N 8. P. 26–37. https://doi.org/10.1615/JAutomatInfScien.v50.i8.30.
14. Denisov S.V., Nomirovskii D.A., Rublyov B.V., Semenov V.V. Convergence of extragradient algorithm with monotone step size strategy for variational inequalities and operator equations. *Journal of Automation and Information Sciences*. 2019. **51**, N 6. P. 12–24. https://doi.org/10.1615/JAutomatInfScien.v51.i6.20.
15. Bot R.I., Csetnek E.R., Vuong P.T. The forward-backward-forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in Hilbert spaces. *European Journal of Operational Research*. 2020. **287**, N 1. P. 49–60. https://doi.org/10.1016/j.ejor.2020.04.035.
16. Vedel Y.I., Golubeva E.N., Semenov V.V., Chabak L.M. Adaptive extraproximal algorithm for the equilibrium problem in the hadamard spaces. *Journal of Automation and Information Sciences*. 2020. **52**, N 8. P. 46–58. https://doi.org/10.1615/JAutomatInfScien.v52.i8.40.
17. Denisov S.V., Semenov V.V., Stetsyuk P.I. Bregman extragradient method with monotone rule of step adjustment. *Cybernetics and Systems Analysis*. 2019. **55**, N 3. P. 377–383. https://doi.org/10.1007/s10559-019-00144-5.
18. Bregman L.M. 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*. 1967. **7**, N 3. P. 200–217. https://doi.org/10.1016/0041-5553(67)90040-7.
19. Beck A. First-order methods in optimization. philadelphia: society for industrial and applied mathematics, 2017. 479 p.
20. Nesterov Yu. Dual extrapolation and its applications to solving variational inequalities and related problems. *Mathematical Programming*. 2007. **109**, N 2-3. P. 319–344. https://doi.org/10.1007/s10107-006-0034-z.
21. Popov L.D. A modification of the Arrow-Hurwicz method for search of saddle points. *Mathematical notes of the Academy of Sciences of the USSR*. 1980. **28**, N 5. P. 845–848. https://doi.org/10.1007/BF01141092.
22. Malitsky Yu.V., Semenov V.V. An extragradient algorithm for monotone variational inequalities. *Cybernetics and Systems Analysis*. 2014. **50**, N 2. P. 271–277. https://doi.org/10.1007/s10559-014-9614-8.
23. Lyashko S.I., Semenov V.V. A new two-step proximal algorithm of solving the problem of equilibrium programming. In: Goldengorin B. (ed.) *Optimization and Its Applications in Control and Data Sciences. Springer Optimization and Its Applications*. Cham : Springer. 2016. **115**. P. 315–325. https://doi.org/10.1007/978-3-319-42056-1_10.
24. Semenov V.V. A version of the mirror descent method to solve variational inequalities. *Cybernetics and Systems Analysis*. 2017. **53**, N 2. P. 234–243. https://doi.org/10.1007/s10559-017-9923-9.
25. Gibali A. A new Bregman projection method for solving variational inequalities in Hilbert spaces. *Pure Appl. Funct. Anal.* 2018. **3**, N 3. P. 403–415.
26. Chabak L., Semenov V., Vedel Y. A new non-Euclidean proximal method for equilibrium problems. In: Chertov O., Mylovanov T., Kondratenko Y., Kacprzyk J., Kreinovich V., Stefanuk V. (eds.) Recent Developments in Data Science and Intelligent Analysis of Information. *ICDSIAI 2018. Advances in Intelligent Systems and Computing*. Cham : Springer. 2019. **836.** P. 50–58. https://doi.org/10.1007/978-3-319-97885-7_6.
27. Nomirovskii D.A., Rublyov B.V., Semenov V.V. Convergence of two-stage method with Bregman divergence for solving variational inequalities. *Cybernetics and Systems Analysis*. 2019. **55**, N 3. P. 359–368. https://doi.org/10.1007/s10559-019-00142-7.
28. Gibali A., Thong D.V. A new low-cost double projection method for solving variational inequalities. *Optim. Eng.* 2020. **21**. P. 1613–1634. https://doi.org/10.1007/s11081-020-09490-2.
29. Malitsky Y., Tam M.K. A forward-backward splitting method for monotone inclusions without cocoercivity. *SIAM Journal on Optimization*. 2020. **30**, N 2. P. 1451–1472. https://doi.org/10.1137/18M1207260.
30. Csetnek E.R., Malitsky Y., Tam M.K. Shadow Douglas-Rachford splitting for monotone inclusions. *Appl Math Optim*. 2019. **80**. P. 665–678. https://doi.org/10.1007/s00245-019-09597-8.
31. Cevher V., Vu B.C. A reflected forward-backward splitting method for monotone inclusions involving Lipschitzian operators. *Set-Valued and Variational Analysis*. 2021. **29**. P. 163–174. https://doi.org/10.1007/s11228-020-00542-4.

Получено 24.02.2021