Журнал обчислювальної та прикладної математики — 2023, № 2 — Journal of Numerical & Applied Mathematics

УДК 519.85

MSC 65K05

EFFICIENCY BOUNDS FOR ALGORITHMS WITH BREGMAN DIVERGENCE

O. S. Kharkov

Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, E-mail: olehharek@gmail.com

ОЦІНКИ ЕФЕКТИВНОСТІ ДЛЯ МЕТОДІВ З ДИВЕРГЕНЦІЄЮ БРЕГМАНА

О. С. Харьков

Факультет комп'ютерних наук та кібернетики, КНУ імені Тараса Шевченка, Київ, Україна, E-mail: olehharek@gmail.com

Abstract. In this paper, variants of extrapolation from the past algorithm and operator extrapolation algorithm with Bregman divergence for solving variational inequalities with monotone and Lipschitz-continuous operators, which act in a finite-dimensional real linear space, are investigated. Main results: efficiency bounds for the gap function.
Keywords: variational inequality, Bregman divergence, Extrapolation from the Past method, Operator Extrapolation method, gap function.

Анотація. У статті досліджено варіанти алгоритмів екстраполяції з минулого та операторної екстраполяції з дивергенцією Брегмана для розв'язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі. Основні результати: оцінки ефективності в термінах функції зазору.
Ключові слова: варіаційна нерівність, дивергенція Брегмана, метод екстраполяції з минулого, метод операторної екстраполяції, функція зазору.

Вступ

Варіаційні нерівності з монотонними операторами є загальним класом задач з опуклою структурою [1].

Окремі задачі опуклої недиференційовної оптимізації можуть ефективно розв'язуватися, якщо їх переформулювати у вигляді сідлових (мінімаксних) задач і застосувати алгоритми розв'язання варіаційних нерівностей.

З появою генеруючих змагальних нейронних мереж та інших моделей змагального навчання стійкий інтерес до алгоритмів розв'язання варіаційних нерівностей виник і в середовищі спеціалістів в галузі машинного навчання.

Дослідження алгоритмів розв'язання варіаційних нерівностей та близьких задач триває. Г. М. Корпелевич в 1970-х роках запропонувала екстраградієнтний метод. В 1980 році Л. Д. Попов запропонував для пошуку сідлових точок опукло-угнутих функцій цікаву модифікацію методу Ерроу–Гурвіца, який став джерелом багатьох сучасних алгоритмів. Ефективним сучасним варіантом екстраградієнтного методу є проксимальний дзеркальний метод (MirrorProx) А. С. Немировського [2].

У статті досліджено варіанти алгоритмів екстраполяції з минулого та операторної екстраполяції з дивергенцією Брегмана для розв'язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі. Основні результати: оцінки ефективності в термінах функції зазору.

1. Допоміжні відомості

Розглянемо варіанти алгоритмів з дивергенцією Брегмана для розв'язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі.

Нехай E — скінченновимірний дійсний лінійний простір. Введемо у цьому просторі норму $\|\cdot\|$ (не обов'язково евклідову). E^* — спряжений простір. Для $a \in E^*$ та $b \in E$ через $\langle a, b\rangle$ позначимо значення линійної функції a в точці b. Спряжена норма на E^* позначається $\|\cdot\|_*$.

Нехай $C \subseteq E$ — непорожня замкнена та опукла множина, A — оператор, що діє з E в E^*. Розглянемо варіаційну нерівність [1]:

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

множину розв'язків якої позначимо S.

Введемо необхідні для алгоритмів конструкції.

Нехай $\varphi : E \to \mathbb{R} \cup \{+\infty\}$ — власна замкнена опукла функція, що диференційовна на $\mathrm{dom}(\partial\varphi)$. Дивергенція Брегмана (відстань Брегмана), що відповідає функції φ, задається формулою [3]

$$V(a, b) = \varphi(a) - \varphi(b) - \langle \nabla\varphi(b), a - b\rangle \quad \forall a \in \mathrm{dom}(\varphi), \ b \in \mathrm{dom}(\partial\varphi).$$

Припустимо, що виконані такі умови:

- функція φ власна замкнена та опукла;
- функція φ диференційовна на $\mathrm{dom}(\partial\varphi)$;
- $C \subseteq \mathrm{dom}(\varphi)$;
- на множині C функція φ є сильно опуклою відносно норми $\|\cdot\|$ з константою сильної опуклості $\sigma > 0$.

Зрозуміло, що

$$V(a, b) \geq \frac{\sigma}{2}\|a - b\|^2 \quad \forall a \in C \quad \forall b \in C \cap \mathrm{dom}(\partial\varphi).$$

та

$$V(a,b)=0 \quad \Leftrightarrow \quad a=b.$$

Має місце корисна 3-точкова тотожність [3]:

$$V(a,c)=V(a,b)+V(b,c)+\langle \nabla\varphi(b)-\nabla\varphi(c), a-b\rangle,$$

де $a\in\mathrm{dom}(\varphi)$, $b,c\in\mathrm{dom}(\partial\varphi)$.

Розглянемо два основних приклади дивергенцій Брегмана. Для

$$\varphi(\cdot)=\tfrac{1}{2}\|\cdot\|_2^2,$$

де $\|\cdot\|_2$ — евклідова норма, маємо

$$V(x,y)=\tfrac{1}{2}\|x-y\|_2^2.$$

Для ймовірносного симплексу

$$\Delta^m=\left\{x\in\mathbb{R}^m: \ x_i\ge 0, \ \sum_{i=1}^m x_i=1\right\}$$

та функції від'ємної ентропії

$$\varphi(x)=\sum_{i=1}^m x_i\ln x_i$$

(за нерівністю Пінскера [4] вона сильно опукла з константою 1 відносно ℓ_1-норми на симплексі Δ^m) отримуємо дивергенцію Кульбака–Лейблера [3]

$$V(x,y)=\sum_{i=1}^m x_i\ln\left(\frac{x_i}{y_i}\right),$$

де $x\in\Delta^m$, $y\in\Delta_+^m=\{x\in\mathbb{R}^m: \ x_i>0, \ \sum_{i=1}^m x_i=1\}$.

Розглянемо сильно опуклі задачі мінімізації вигляду

$$P_x^C(a)=\mathrm{argmin}_{y\in C}\{-\langle a,y-x\rangle+V(y,x)\}, \tag{2}$$

де $a\in E^*$, $x\in\mathrm{dom}(\partial\varphi)$. Відомо [3], що задача (2) має єдиний розв'язок $z\in C\cap\mathrm{dom}(\partial\varphi)$, причому

$$-\langle a,y-z\rangle+\langle\nabla\varphi(z)-\nabla\varphi(x), y-z\rangle\ge 0 \quad \forall y\in C.$$

Останню нерівність, ураховуючи 3-точкову тотожність, можна записати у вигляді

$$V(y,z)\le V(y,x)-V(y,z)-\langle a,y-z\rangle \quad \forall y\in C.$$

Зауваження 1. Точка $P_x^C(a)$ в евклідовому випадку співпадає з евклідовою метричною проєкцією

$$P_C(x+a)=\mathrm{argmin}_{y\in C}\|y-(x+a)\|_2.$$

Зауваження 2. Для ймовірносного симплексу Δ^m та дивергенції Кульбака–Лейблера маємо [3]

$$P_x^{\Delta^m}(a)=\left(\frac{x_1e^{a_1}}{\sum_{j=1}^m x_je^{a_j}}, \frac{x_2e^{a_2}}{\sum_{j=1}^m x_je^{a_j}}, \dots, \frac{x_me^{a_m}}{\sum_{j=1}^m x_je^{a_j}}\right), \quad a\in\mathbb{R}^m, \ x\in\Delta_+^m.$$

Зауваження 3. Часто відображення $P_x^C : E^* \to C \cap \operatorname{dom}(\partial\varphi)$ називають прокс-відображенням [5] та позначають $P_x^C(a)$ через $\operatorname{Mirr}_x(a)$ [6].

Припустимо, що виконані такі умови:

- множина $C \subseteq E$ – опукла та замкнена;
- оператор $A : E \to E^*$ – монотонний на C, тобто
$$\langle Ax - Ay, x - y \rangle \geq 0 \quad \forall x, y \in C,$$
та ліпшицевий на C (з константою $L > 0$), тобто
$$\|Ax - Ay\|_* \leq \|x - y\| \quad \forall x, y \in C;$$
- множина S непорожня.

Розглянемо дуальну варіаційну нерівність:

$$x \in C : \ \langle Ay, x - y \rangle \leq 0 \quad \forall y \in C. \tag{3}$$

Множину розв'язків задачі (3) позначимо S^d. Множина S^d опукла та замкнена. Для монотонних операторів A завжди маємо $S \subseteq S^d$. В наших умовах маємо $S^d = S$.

2. Алгоритми

Перший з розглянутих алгоритмів має вигляд.

Алгоритм 1. Екстраполяція з минулого.

Для $x_1 = y_0 \in C \cap \operatorname{dom}(\partial\varphi)$ генерируємо послідовність елементів x_n, $y_n \in C \cap \operatorname{dom}(\partial\varphi)$ за допомогою ітераційної схеми

$$\begin{cases} y_n = P_{x_n}^C\left(-\lambda_n A y_{n-1}\right), \\ x_{n+1} = P_{x_n}^C\left(-\lambda_n A y_n\right), \end{cases}$$

де $\lambda_n > 0$.

Зауваження 4. Алгоритм 1 запропоновано в [7]. В роботах [7–10] досліджувалась його збіжність та прононувались деякі модифікації.

При виконанні для деякого $n \in \mathbb{N}$ в алгоритмі 1 рівностей

$$y_n = y_{n-1} = x_n \quad \text{або} \quad x_{n+1} = x_n = y_n \tag{4}$$

має місце включення $y_n \in S$. Дійсно, рівність

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

рівносильна нерівності

$$\langle A y_n, y - x_{n+1} \rangle + \frac{\langle \nabla\varphi(x_{n+1}) - \nabla\varphi(x_n), y - x_{n+1} \rangle}{\lambda_n} \geq 0 \quad \forall y \in C.$$

З другої рівності (4) випливає

$$\langle A y_n, y - y_n \rangle \geq 0 \ \ \forall y \in C,$$

тобто, $y_n \in S$.

Аналогічно, з

$$\langle Ay_{n-1}, y - y_n \rangle + \frac{\langle \nabla\varphi(y_n) - \nabla\varphi(x_n), y - y_n \rangle}{\lambda_n} \geq 0 \quad \forall y \in C$$

при першій рівності в (4) отримуємо $y_n \in S$.

Далі припустимо, що для всіх номерів $n \in \mathbb{N}$ умова (4) не має місця.

Другий з розглянутих алгоритмів має вигляд.

Алгоритм 2. Операторна екстраполяція.

Обираємо $x_0 = x_1 \in C \cap \operatorname{dom}(\partial\varphi)$, $\lambda_n, \mu_n > 0$. Покладаємо $n = 1$.

1: Обчислити

$$x_{n+1} = P^C_{x_n}\left(-\lambda_n A x_n - \mu_n (A x_n - A x_{n-1})\right).$$

2: Якщо $x_{n+1} = x_n = x_{n-1}$, то СТОП, інакше покласти $n := n + 1$ та перейти до **1**.

Зауваження 5. Алгоритм 2 запропоновано в [11]. Алгоритм 2 є модифікацією методу з попереднього розділу, що використовує дивергенцію Брегмана замість квадрату евклідової норми.

Правило зупинки в алгоритмі 2 обґрунтовується так: при виконанні

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

маємо

$$x_n = P^C_{x_n}\left(-\lambda_n A x_n\right),$$

звідки $x_n \in S$.

Наведемо дві версії алгоритму 2.

Розглянемо варіаційну нерівність на стандартному симплексі:

$$\text{знайти} \;\; x \in \Delta^m : \;\; (Ax, y - x) \geq 0 \;\; \forall y \in \Delta_m.$$

Обираючи дивергенцію Кульбака–Лейблера та $\mu_n = \lambda_n = \lambda > 0$, одержуємо таку версію:

$$x_i^{n+1} = \frac{x_i^n e^{-\lambda(2Ax_n - Ax_{n-1})_i}}{\sum_{j=1}^d x_j^n e^{-\lambda(2Ax_n - Ax_{n-1})_j}}, \;\; i = 1, ..., m,$$

де $(a)_i \in \mathbb{R}$ — i-та координата вектора $a \in \mathbb{R}^m$.

Розглянемо варіаційну нерівність на добутку стандартних симплексів:

$$\text{знайти} \;\; x \in \Delta^{m_1} \times \Delta^{m_2} : \;\; (Ax, y - x) \geq 0 \;\; \forall y \in \Delta^{m_1} \times \Delta^{m_2}. \tag{5}$$

За сепарабельною функцією

$$\varphi(x) = \varphi_1(x_1) + \varphi_2(x_2) = \sum_{i=1}^{m_1} x_{1,i} \ln x_{1,i} + \sum_{i=1}^{m_2} x_{2,i} \ln x_{2,i},$$

де $x = (x_1, x_2) = (\underbrace{x_{1,1}, x_{1,2}, \dots x_{1,m_1}}_{x_1}, \underbrace{x_{2,1}, x_{2,2}, \dots x_{2,m_2}}_{x_2}) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, побудуємо дивергенцію Брегмана на $C = \Delta^{m_1} \times \Delta^{m_2}$:

$$V(x, y) = V_1(x_1, y_1) + V_2(x_2, y_2) = \sum_{i=1}^{m_1} x_{1,i} \ln \frac{x_{1,i}}{y_{1,i}} + \sum_{i=1}^{m_2} x_{2,i} \ln \frac{x_{2,i}}{y_{2,i}}.$$

Алгоритм 2 для нерівності (5) з таким вибором дивергенції та $\mu_n = \lambda_n = \lambda > 0$ приймає вигляд:

$$x_{k,i}^{n+1} = \frac{x_{k,i}^n \exp\left(-\lambda\left(2Ax_n - Ax_{n-1}\right)_{k,i}\right)}{\sum_{j=1}^{m_k} x_{k,j}^n \exp\left(-\lambda\left(2Ax_n - Ax_{n-1}\right)_{k,j}\right)}, \quad k = 1, 2, \ i = 1, \dots, m_k,$$

де $(a)_{k,i}$ — $\left(\sum_{t=1}^{k-1} m_t + i\right)$-та координата вектора $a \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$.

Якість наближеного розв'язку $x \in C$ варіаційної нерівності (1) будемо оцінювати за допомогою невід'ємної функції зазору [2]

$$\operatorname{gap}(x) = \sup_{y \in C} \langle Ay, x - y \rangle. \tag{6}$$

Очевидно, що для коректності означення функції зазору (6) необхідна обмеженість допустимої множини C. Якщо $x \in C$ — розв'язок (1), то $\operatorname{gap}(x) = 0$. Навпаки, якщо для $x \in C$ маємо $\operatorname{gap}(x) = 0$, то x — розв'язок (1).

3. Сублінійні оцінки ефективності

У випадку обмеженості множини C доведемо, що алгоритмам необхідно зробити $O\left(\frac{1}{\varepsilon}\right)$ ітерацій для отримання точки $x \in C$ з

$$\operatorname{gap}(x) = \sup_{y \in C} \langle Ay, x - y \rangle \le \varepsilon.$$

Почнемо з аналізу алгоритму 1.

Для породжених алгоритмом 1 послідовностей (x_n) та (y_n) мають місце нерівності

$$-\lambda_n \langle Ay_{n-1}, y - y_n \rangle \le V(y, x_n) - V(y_n, x_n) - V(y, y_n) \quad \forall y \in C. \tag{7}$$

$$-\lambda_n \langle Ay_n, y - x_{n+1} \rangle \le V(y, x_n) - V(x_{n+1}, x_n) - V(y, x_{n+1}) \quad \forall y \in C. \tag{8}$$

З (8) випливає

$$\begin{aligned} V(y, x_{n+1}) &\le V(y, x_n) - V(x_{n+1}, x_n) + \lambda_n \langle Ay_n, y - x_{n+1} \rangle = \\ &= V(y, x_n) - V(x_{n+1}, x_n) + \lambda_n \langle Ay_n, y_n - x_{n+1} \rangle + \\ &\qquad + \lambda_n \langle Ay_n, y - y_n \rangle. \end{aligned}$$

Урахувавши монотонность оператора A, отримаємо

$$\begin{aligned} V(y, x_{n+1}) \le V(y, x_n) - V(x_{n+1}, x_n) + \lambda_n \langle Ay_n, y_n - x_{n+1} \rangle + \\ + \lambda_n \langle Ay, y - y_n \rangle. \end{aligned} \tag{9}$$

З (7) випливає

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

Оцінимо зверху $-V(x_{n+1}, x_n)$ в (9) за допомогою (10). Отримаємо

$$V(y, x_{n+1}) \le V(y, x_n) - V(x_{n+1}, y_n) - V(y_n, x_n) + \\ + \lambda_n \langle Ay_n - Ay_{n-1}, y_n - x_{n+1} \rangle + \\ + \lambda_n \langle Ay, y - y_n \rangle . \quad (11)$$

Оцінимо зверху доданок $\lambda_n \langle Ay_n - Ay_{n-1}, y_n - x_{n+1} \rangle$ в (11). Маємо

$$\lambda_n \langle Ay_n - Ay_{n-1}, y_n - x_{n+1} \rangle \le \lambda_n L \|y_{n-1} - y_n\| \|x_{n+1} - y_n\| \le \\ \le \frac{\lambda_n L}{2} \|y_{n-1} - y_n\|^2 + \frac{\lambda_n L}{2} \|y_n - x_{n+1}\|^2 \le \\ \le \lambda_n L \|y_{n-1} - x_n\|^2 + \lambda_n L \|y_n - x_n\|^2 + \frac{\lambda_n L}{2} \|y_n - x_{n+1}\|^2 \le \\ \le \frac{2\lambda_n L}{\sigma} V(x_n, y_{n-1}) + \frac{2\lambda_n L}{\sigma} V(y_n, x_n) + \frac{\lambda_n L}{\sigma} V(x_{n+1}, y_n).$$

Приходимо до нерівності

$$V(y, x_{n+1}) \le V(y, x_n) - V(y_n, x_n) - V(x_{n+1}, y_n) + \\ + \frac{2\lambda_n L}{\sigma} V(x_n, y_{n-1}) + \frac{2\lambda_n L}{\sigma} V(y_n, x_n) + \frac{\lambda_n L}{\sigma} V(x_{n+1}, y_n) + \\ + \lambda_n \langle Ay, y - y_n \rangle . \quad (12)$$

Перепишемо (12) у вигляді

$$\lambda_n \langle Ay, y_n - y \rangle \le \left(V(y, x_n) + \frac{2\lambda_n L}{\sigma} V(x_n, y_{n-1}) \right) - \\ - \left(V(y, x_{n+1}) + \frac{2\lambda_{n+1} L}{\sigma} V(x_{n+1}, y_n) \right) - \\ - \left(1 - \frac{2\lambda_{n+1} L}{\sigma} - \frac{\lambda_n L}{\sigma} \right) V(x_{n+1}, y_n) - \\ - \left(1 - \frac{2\lambda_n L}{\sigma} \right) V(y_n, x_n). \quad (13)$$

Припустимо, що $\lambda_n \in \left(0, \frac{\sigma}{3L}\right]$. Тоді з (13) випливає

$$2\lambda_n \langle Ay, y_n - y \rangle \le \left(V(y, x_n) + \frac{2\lambda_n L}{\sigma} V(x_n, y_{n-1}) \right) - \\ - \left(V(y, x_{n+1}) + \frac{2\lambda_{n+1} L}{\sigma} V(x_{n+1}, y_n) \right). \quad (14)$$

Просумувавши (14) по n від 1 до N отримаємо

$$2 \sum_{n=1}^{N} \lambda_n \langle Ay, y_n - y \rangle \le V(y, x_1) + \frac{2\lambda_1 L}{\sigma} V(x_1, y_0),$$

та

$$\langle Ay, z_N - y \rangle \le \frac{V(y, x_1)}{2 \sum_{n=1}^{N} \lambda_n}, \quad (15)$$

де $z_N = \frac{\sum_{n=1}^{N} \lambda_n y_n}{\sum_{n=1}^{N} \lambda_n}$. Переходимо до супремуму по $y \in C$ в (15)

$$\operatorname{gap}(z_N) \le \frac{\sup_{y\in C} V(y, x_1)}{2\sum_{n=1}^{N} \lambda_n}.$$

Таким чином, має місце

Теорема 1. *Нехай $C \subseteq E$ — непорожня опукла замкнена обмежена множина, $A : E \to E^*$ — монотонний та L-ліпшицевий на множині C оператор. Нехай (y_n) — послідовність, що породжена алгоритмом 1 з $\lambda_n = \frac{\sigma}{3L}$, тобто,*

$$\begin{cases} x_1 = y_0 \in C, \\ y_n = P_C\left(x_n - \frac{\sigma}{3L} A y_{n-1}\right), \\ x_{n+1} = P_C\left(x_n - \frac{\sigma}{3L} A y_n\right). \end{cases}$$

Тоді для послідовності середніх $z_N = \frac{1}{N}\sum_{n=1}^{N} y_n$ має місце оцінка

$$\operatorname{gap}(z_N) \le \frac{L\frac{3}{2\sigma}\sup_{y\in C} V(y, x_1)}{N}.$$

Проаналізуємо алгоритм 2.

Нехай $\lambda_n \in \left(0, \frac{\sigma}{2L}\right]$, $\mu_n = \lambda_{n-1}$. Для послідовності (x_n) має місце нерівність

$$\begin{aligned} -\langle \lambda_n A x_n + \lambda_{n-1}(A x_n - A x_{n-1}), y - x_{n+1}\rangle \le \\ \le V(y, x_n) - V(x_{n+1}, x_n) - V(y, x_{n+1}) \quad \forall y \in C. \end{aligned} \tag{16}$$

Перепишемо (16) таким чином

$$\begin{aligned} V(y, x_n) - V(y, x_{n+1}) \ge \\ \ge \lambda_n \langle A x_{n+1}, x_{n+1} - y\rangle - \lambda_n \langle A x_{n+1} - A x_n, x_{n+1} - y\rangle + \\ + \lambda_{n-1}\langle A x_n - A x_{n-1}, x_n - y\rangle + \lambda_{n-1}\langle A x_n - A x_{n-1}, x_{n+1} - x_n\rangle + \\ + V(x_{n+1}, x_n). \end{aligned} \tag{17}$$

Сумуючи (17) по n від 1 до N, отримуємо

$$\begin{aligned} V(y, x_1) - V(y, x_{N+1}) \ge \\ \ge \sum_{n=1}^{N} \lambda_n \langle A x_{n+1}, x_{n+1} - y\rangle - \lambda_N \langle A x_{N+1} - A x_N, x_{N+1} - y\rangle + \\ + \sum_{n=1}^{N} \left(\lambda_{n-1}\langle A x_n - A x_{n-1}, x_{n+1} - x_n\rangle + V(x_{n+1}, x_n)\right). \end{aligned} \tag{18}$$

Ліпшицевість оператора A дає

$$\sum_{n=1}^{N}\left(\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_{n+1}-x_n\right\rangle + V\left(x_{n+1}, x_n\right)\right) \geq$$

$$\geq \sum_{n=1}^{N}\left(-\frac{\lambda_{n-1}L}{2}\|x_n - x_{n-1}\|^2 - \frac{\lambda_{n-1}L}{2}\|x_n - x_{n+1}\|^2 + \frac{\sigma}{2}\|x_{n+1}-x_n\|^2\right) \geq$$

$$\geq \sum_{n=1}^{N}\left(-\frac{\sigma}{4}\|x_n - x_{n-1}\|^2 - \frac{\sigma}{4}\|x_n - x_{n+1}\|^2 + \frac{\sigma}{2}\|x_{n+1}-x_n\|^2\right) =$$

$$= \sum_{n=1}^{N}\left(-\frac{\sigma}{4}\|x_n - x_{n-1}\|^2 + \frac{\sigma}{4}\|x_{n+1}-x_n\|^2\right) = \frac{\sigma}{4}\left\|x_{N+1}-x_N\right\|^2.$$

Використовуючи останню оцінку в (18), отримуємо

$$V\left(y, x_1\right) - V\left(y, x_{N+1}\right) \geq$$

$$\geq \sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle - \lambda_N\left\langle Ax_{N+1}-Ax_N, x_{N+1}-y\right\rangle +$$

$$+ \frac{\sigma}{4}\left\|x_{N+1}-x_N\right\|^2 \geq$$

$$\geq \sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle - \lambda_N L\left\|x_{N+1}-x_N\right\|\left\|x_{N+1}-y\right\| +$$

$$+ \frac{\sigma}{4}\left\|x_{N+1}-x_N\right\|^2 \geq$$

$$\geq \sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle - \frac{\lambda_N L}{2}\left\|x_{N+1}-y\right\|^2.$$

Приходимо до нерівності

$$\sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle - \frac{\lambda_N L}{2}\left\|x_{N+1}-y\right\|^2 +$$

$$+ V\left(y, x_{N+1}\right) \leq V\left(y, x_1\right) \quad \forall y \in C. \quad (19)$$

Використовуючи монотонність оператора A, отримуємо

$$\sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle \geq \sum_{n=1}^{N}\lambda_n\left\langle Ay, x_{n+1}-y\right\rangle =$$

$$= \left(\sum_{n=1}^{N}\lambda_n\right)\left\langle Ay, z_{N+1}-y\right\rangle, \quad (20)$$

де

$$z_{N+1} = \frac{\sum_{n=1}^{N}\lambda_n x_{n+1}}{\sum_{n=1}^{N}\lambda_n}.$$

Враховуючи оцінку (20) в (19), приходимо до нерівності

$$\left(\sum_{n=1}^{N}\lambda_n\right)\langle Ay, z_{N+1}-y\rangle + \left(\frac{\sigma}{2}-\frac{\lambda_N L}{2}\right)\|x_{N+1}-y\|^2 \le V(y,x_1) \quad \forall y\in C,$$

звідки випливає

$$\operatorname{gap}(z_{N+1}) = \sup_{y\in C}\langle Ay, z_{N+1}-y\rangle \le \frac{\sup_{y\in C} V(y,x_1)}{\sum_{n=1}^{N}\lambda_n}.$$

Таким чином, має місце

Теорема 2. *Нехай (x_n) — послідовність, що породжена алгоритмом 2 з $\lambda_n=\mu_n=\frac{\sigma}{2L}$, тобто,*

$$x_{n+1} = P_{x_n}^C\left(-\tfrac{\sigma}{2L}(2Ax_n - Ax_{n-1})\right).$$

Тоді має місце оцінка

$$\operatorname{gap}(z_{N+1}) \le \frac{L\frac{2}{\sigma}\sup_{y\in C} V(y,x_1)}{N},$$

де $z_{N+1}=\frac{1}{N}\sum_{n=1}^{N}x_{n+1}$.

Зауваження 6. Припустимо, що замість ліпшицевості A виконана така умова:

- оператор $A: E\to E^*$ — відносно ліпшицевий на C (з константою $L>0$), тобто

  $$\langle Ax-Ay, x-z\rangle \le LV(x,y)+LV(z,x),$$

  де x, $y\in C\cap\operatorname{dom}(\partial\varphi)$, $z\in C$;

Якщо оператор $A: E\to E^*$ — L-ліпшицевий на C, то він є $\frac{L}{\sigma}$-відносно ліпшицевим на C [12]. Дійсно,

$$\langle Ax-Ay, x-z\rangle \le \|Ax-Ay\|_*\|x-z\| \le L\|x-y\|\|x-z\| \le$$
$$\le \frac{L}{2}\|x-y\|^2+\frac{L}{2}\|x-z\|^2 \le \frac{L}{\sigma}V(x,y)+\frac{L}{\sigma}V(z,x).$$

Актуальною є задача отримання оцінок для алгоритмів 1 та 2 в класі відносно ліпшицевих монотонних операторів.

Зауваження 7. У роботі [13] побудовані адаптивні варіанти алгоритмів 1, 2.

Заключні зауваження

У роботі досліджено варіанти алгоритмів екстраполяції з минулого та операторної екстраполяції з дивергенцією Брегмана для розв'язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі. Основні результати: $O\left(\frac{1}{\varepsilon}\right)$-оцінки ефективності в термінах функції зазору.

Робота виконана за фінансової підритмки МОН України (проєкт «Обчислювальні алгоритми і оптимізація для штучного інтелекту, медицини та оборони», 0122U002026).

Література

1. Семенов В. В. Варіаційні нерівності: теорія та алгоритми. Київ: ВПЦ «Київський університет», 2021. 167 с.
2. 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 J. on Optim.* 2004. Vol. 15. P. 229–251.
3. Beck A. First-Order Methods in Optimization. Philadelphia: Society for Industrial and Applied Mathematics, 2017. 479 p.
4. Vishnoi N. K. Algorithms for Convex Optimization. Cambridge: Cambridge University Press, 2021. 340 p.
5. Juditsky A., Nemirovski A., Tauvel C. Solving variational inequalities with Stochastic Mirror-Prox algorithm. *Stochastic Systems.* 2011. Vol. 1. No. 1. P. 17–58.
6. Allen-Zhu Z., Orecchia L. Linear coupling: An ultimate unification of gradient and mirror descent. arXiv preprint arXiv:1407.1537. 2014.
7. Semenov V. V. A variant of mirror descent method for solving variational inequalities. In: Polyakova, L. N. (ed.) Constructive Nonsmooth Analysis and Related Topics (dedicated to the memory of V. F. Demyanov). IEEE, 2017. P. 281–284.
8. Semenov V. V. A Version of the Mirror descent Method to Solve Variational Inequalities. *Cybernetics and Systems Analysis.* 2017. Vol. 53. P. 234–243.
9. Nomirovskii D. A., Rublyov V. V., Semenov V. V. Convergence of Two-Stage Method with Bregman Divergence for Solving Variational Inequalities. *Cybernetics and Systems Analysis.* 2019. Vol. 55. P. 359–368.
10. 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, vol. 836. Springer, Cham, 2019. P. 50–58.
11. Семенов В. Харьков О. Метод операторної екстраполяції для варіаційних нерівностей в банахових просторах. *Фізико-математичне моделювання та інформаційні технології.* 2023. Вип. 37. С. 118–122.
12. Cohen M. B., Sidford A., Tian K. Relative Lipschitzness in Extragradient Methods and a Direct Recipe for Acceleration. arXiv preprint arXiv: 2011.06572. 2021.
13. Семенов В. В., Денисов С. В., Сирык Д. С., Харьков О. С. Сходимость метода экстраполяции из прошлого и метода операторной экстраполяции. *Міжнародний науково-технічний журнал «Проблеми керування та інформатики».* 2021. Том 66. № 3. С. 58–72.

Надійшла: 10.07.2023 / Прийнята: 15.10.2023