Журнал обчислювальної та прикладної математики | 2022, № 2 | Journal of Numerical & Applied Mathematics

УДК 519.85
MSC 37C75, 65K05

WEAK CONVERGENCE OF THE OPERATOR EXTRAPOLATION METHOD FOR VARIATIONAL INEQUALITIES IN UNIFORMLY CONVEX BANACH SPACES

S. V. Denisov, V. V. Semenov, O. S. Kharkov
Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, E-mail: {denisov.univ, volodya.semenov, olehharek}@gmail.com

СЛАБКА ЗБІЖНІСТЬ МЕТОДУ ОПЕРАТОРНОЇ ЕКСТРАПОЛЯЦІЇ ДЛЯ ВАРІАЦІЙНИХ НЕРІВНОСТЕЙ В РІВНОМІРНО ОПУКЛИХ БАНАХОВИХ ПРОСТОРАХ

С. В. Денисов, В. В. Семенов, О. С. Харьков
Факультет комп'ютерних наук та кібернетики, КНУ імені Тараса Шевченка, Київ, Україна, E-mail: {denisov.univ, volodya.semenov, olehharek}@gmail.com

Abstract. This work is devoted to the study of new iterative algorithms for solving variational inequalities in uniformly convex Banach spaces. The first algorithm is a modification of the forward-reflected-backward algorithm, which uses the Alber generalized projection instead of the metric one. The second algorithm is an adaptive version of the first one, where the monotone step size update rule is used, which does not require knowledge of Lipschitz constants and linear search procedure.
Keywords: variational inequality, monotone operator, Alber generalized projection, 2-uniformly convex Banach space, uniformly smooth Banach space, algorithm, weak convergence, gap function.

Анотація. Робота присвячена дослідженню нових ітераційних алгоритмів для розв'язання варіаційних нерівностей в рівномірно опуклих банахових просторах. Перший алгоритм – модифікація «forward-reflected-backward algorithm», що використовує узагальнену проєкцію Альбера замість метричної. Другий алгоритм є адаптивним варіантом першого, де використовується правило поновлення величини кроку, що не вимагає знання ліпшицевих констант та лінійного пошуку.
Ключові слова: варіаційна нерівність, монотонний оператор, узагальнена проєкція Альбера, 2-рівномірно опуклий банаховий простір, рівномірно гладкий банаховий простір, алгоритм, слабка збіжність, функція зазору.

Вступ

Варіаційні нерівності дають уніфікований засіб формулювання багатьох актуальних задач математичної фізики, оптимального керування та дослідження операцій [1, 2]. Створення та дослідження алгоритмів розв'язання варіаційних нерівностей та близьких задач є напрямом прикладного нелінійного аналізу, що активно розвивається [3]. Зауважимо, що часто негладкі задачі опуклої оптимізації можуть ефективно розв'язуватися, якщо їх переформулювати у вигляді сідлових задач і застосувати алгоритми розв'язання варіаційних нерівностей [4]. З появою генеруючих змагальних нейронних мереж (generative adversarial network, GAN) та інших моделей змагального навчання стійкий інтерес до алгоритмів розв'язання варіаційних нерівностей виник і в середовищі спеціалістів в галузі машинного навчання [5].

В даній роботі, що продовжує цикл статей [6–8], запропоновано та досліджено нові ітераційні алгоритми для розв'язання варіаційних нерівностей в рівномірно опуклих банахових просторах. Перший алгоритм — модифікація методу «forward-reflected-backward algorithm» [9], що використовує узагальнену проєкцію Альбера замість метричної. Другий алгоритм є адаптивним варіантом першого, де використовується правило поновлення величини кроку, що не вимагає знання ліпшицевих констант та лінійного пошуку.

Для варіаційних нерівностей з монотонними, ліпшицевими операторами, що діють в 2-рівномірно опуклому та рівномірно гладкому банаховому просторі, доведено теореми про слабку збіжність методів. Також для першого алгоритму доведено оцінку ефективності в термінах функції зазору.

Допоміжні відомості

Спочатку нагадаємо декілька понять та фактів геометрії банахових просторів та нелінійного аналізу, що необхідні для формулювання та доведення результатів [10–14].

Нехай E — дійсний банаховий простір з нормою $\|\cdot\|$, E^* – спряжений до E простір, $\langle x^*, x\rangle$ — значення функціоналу $x^* \in E^*$ на елементі $x \in E$. Норму в E^* будемо позначати $\|\cdot\|_*$.

Нехай $S_E = \{x \in E : \ \|x\| = 1\}$. Банаховий простір E називають строго опуклим, якщо для всіх $x, y \in S_E$ та $x \neq y$ маємо $\left\|\frac{x+y}{2}\right\| < 1$. Модуль опуклості простору E визначається таким чином

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

Банаховий простір E називають рівномірно опуклим, якщо $\delta_E(\varepsilon) > 0$ для всіх $\varepsilon \in (0, 2]$. Банаховий простір E називають 2-рівномірно опуклим, якщо існує таке $c > 0$, що $\delta_E(\varepsilon) \geq c\varepsilon^2$ для всіх $\varepsilon \in (0, 2]$. Очевидно, що 2-рівномірно опуклий простір є рівномірно опуклим. Відомо, що рівномірно опуклий банаховий простір рефлексивний [12].

Банаховий простір E називають гладким, якщо границя

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

існує для всіх $x, y \in S_E$. Банаховий простір E називають рівномірно гладким, якщо границя (1) існує рівномірно по $x, y \in S_E$.

Відомо, що гільбертові простори та простори L_p $(1 < p \leq 2)$ є 2-рівномірно опуклими та рівномірно гладкими (простори L_p рівномірно гладкі для $p \in (1, \infty)$) [12].

Багатозначний оператор $J : E \to 2^{E^*}$, що діє за правилом

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

називають нормалізованим дуальным відображенням [12]. Відомо, що [12]: якщо простір E гладкий, то відображення J однозначне; якщо простір E строго опуклий, то відображення J ін'єктивне та строго монотонне; якщо простір E рефлексивний, то відображення J сюр'єктивне; якщо простір E рівномірно гладкий, то відображення J рівномірно неперервне на обмежених підмножинах E.

Нехай E — гладкий банаховий простір. Розглянемо введений Я. Альбером [10, 11] функціонал

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

З означення ϕ випливає корисна 3-точкова тотожність:

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

Якщо простір E строго опуклий, то для $x, y \in E$ маємо

$$\phi(x, y) = 0 \Leftrightarrow x = y.$$

Лема 1 (Aoyama K., Kohsaka F. [13]). *Нехай E — 2-рівномірно опуклий та гладкий банаховий простір. Тоді для деякого $\mu \geq 1$ виконується нерівність*

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

Для банахових просторів ℓ_p, L_p та W_p^m $(1 < p \leq 2)$ маємо $\mu = \frac{1}{p-1}$ [14]. Для гільбертового простору нерівність (2) перетворюється на тотожність з $\mu = 1$.

Нехай K — непорожня замкнена та опукла підмножина рефлексивного, строго опуклого та гладкого простору E. Відомо [10, 11], що для кожного $x \in E$ існує єдиний елемент $z \in K$ такий, що

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

Цей елемент z позначають $\Pi_K x$, а відповідний оператор $\Pi_K : E \to K$ називають узагальненою проєкцією E на K (узагальненою проєкцією Альбера) [10, 11]. Зауважимо, що якщо E гільбертовий простір, то Π_K співпадає з метричною проєкцією на множину K.

Лема 2 (Alber Y. I. [11]). *Нехай K – замкнена та опукла підмножина рефлексивного, строго опуклого та гладкого простору E, $x \in E$, $z \in K$. Тоді*

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

Нерівність леми 2 рівносильна такій [11]:

$$\phi\left(y, \Pi_K x\right) + \phi\left(\Pi_K x, x\right) \leq \phi\left(y, x\right) \quad \forall y \in K.$$

Основним елементом розглянутих нижче алгоритмів є обчислення за даними точками $x \in E$ та $x^* \in E^*$ нової точки

$$x^+ = \Pi_K J^{-1}\left(Jx - x^*\right).$$

З леми 2 та 3-точкової тотожності випливає фундаментальна для аналізу алгоритмів нерівність

$$\phi\left(y, x^+\right) \leq \phi\left(y, x\right) - \phi\left(x^+, x\right) + 2\left\langle x^*, y - x^+\right\rangle \quad \forall y \in K.$$

Постановка задачі

Розглянемо варіаційну нерівність:

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

де C — непорожня підмножина 2-рівномірно опуклого та рівномірно гладкого банахового простору E, A — оператор, що діє з E в E^*. Множину розв'язків (3) позначимо S.

Припустимо, що виконані такі умови: множина $C \subseteq E$ — опукла та замкнена; оператор $A : E \to E^*$ — монотонний та ліпшицевий на C (з константою $L > 0$); множина S не порожня.

Варіаційну нерівність (3) можна сформулювати як задачу пошуку нерухомої точки [11]:

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

де $\lambda > 0$. Формулювання (4) корисне, оскільки дає очевидну алгоритмічну ідею. Схема

$$x_{n+1} = \Pi_C J^{-1}\left(Jx_n - \lambda Ax_n\right) \tag{5}$$

вивчалась для обернено сильно монотонних операторів $A : E \to E^*$. Але для ліпшицевих монотонних операторів схема (5) в загальному випадку не збігається.

Однією з основних теоретичних задач є оцінка числа ітерацій алгоритму, що необхідне для отримання наближеного розв'язку заданої якості. Якість наближеного розв'язку $x \in C$ варіаційної нерівності (3) будемо оцінювати за допогою невід'ємної функції зазору

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

Очевидно, що для коректності означення функції зазору (6) необхідна обмеженість допустимої множини C.

Лема 3. *Нехай оператор A — монотонний. Якщо $x \in C$ — розв'язок* (3), *то*

$$\operatorname{gap}(x) = 0.$$

Навпаки, якщо для $x \in C$ маємо

$$\operatorname{gap}(x) = 0,$$

то x — розв'язок (3).

Алгоритм операторної екстраполяції

Для розв'язання варіаційної нерівності (3) пропонуємо наступний алгоритм операторної екстраполяції.

Алгоритм 1.

0. Обираємо $x_0 = x_1 \in E$, $\lambda_n > 0$. Покладаємо $n = 1$.

1. Обчислити

$$m_n = \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right),$$

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

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

Алгоритм 1 є модифікацією «forward-reflected-backward algorithm» [9] для задач в банахових просторах, що використовує узагальнену проєкцію Альбера [11] замість метричної.

Збіжність «forward-reflected-backward algorithm» в гільбертовому просторі доведена в [9].

Правило зупинки обґрунтовується тотожністю (4), що рівносильна варіаційній нерівності (3). Дійсно, при виконанні

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

маємо

$$x_n = \Pi_C J^{-1}\left(Jx_n - \lambda_n Ax_n\right),$$

звідки $y_n \in S$.

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

$$\operatorname{gap}(x) \le \varepsilon,$$

де $\varepsilon > 0$, $D = \sup_{a,b \in C} \phi(a,b) < +\infty$.

Теорема 1. *Нехай (x_n) — послідовність, що породжена алгоритмом 1. Припустимо, що $\lambda_n \in \left(0, \frac{1}{2\mu L}\right]$. Тоді для послідовності чезарівсъких середніх $z_{N+1} = \frac{\sum_{n=1}^{N}\lambda_n x_{n+1}}{\sum_{n=1}^{N}\lambda_n}$ має місце нерівність*

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

Наслідок 1. *Нехай (x_n) — послідовність, що породжена алгоритмом 1 с $\lambda_n = \frac{1}{2\mu L}$. Тоді для послідовності середніх $z_{N+1} = \frac{1}{N}\sum_{n=1}^{N} x_{n+1}$ має місце оцінка*

$$\operatorname{gap}(z_{N+1}) \le \frac{L\mu}{N}\sup_{y \in C}\phi(y, x_1).$$

Перейдемо до питання слабкої збіжності алгоритму 1. В якості функції Ляпунова оберемо

$$V_n = \phi\left(z, x_n\right) + 2\left\langle m_n, x_n - z\right\rangle + \mu\lambda_{n-1}L\phi\left(x_n, x_{n-1}\right),$$

де $z \in S$.

Має місце

Лема 4. *Для послідовності (x_n), що породжена алгоритмом 1, виконується нерівність*

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

де $z \in S$.

Теорема 2. *Нехай C — непорожня опукла та замкнена підмножина 2-рівномірно опуклого та рівномірно гладкого банахового простору E, $A : E \to E^*$ — монотонний та ліпшицевий на множині C оператор, $S \neq \emptyset$. Припустимо, що нормалізоване дуальне відображення J секвенційно слабко неперервне та послідовність (λ_n) така, що $0 < \inf_n \lambda_n \leq \sup_n \lambda_n < \frac{1}{2\mu L}$. Тоді послідовність (x_n), що породжена алгоритмом 1, слабко збігаються до деякої точки $z \in S$.*

У випадку $C = E$ варіаційна нерівність (3) набуває вигляду операторного рівняння

$$\text{знайти} \quad x \in E: \quad Ax = 0 \tag{7}$$

Для (7) алгоритм 1 дає такий ітераційний процес

$$Jx_{n+1} = Jx_n - \lambda_n Ax_n - \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), \tag{8}$$

який збіжний лише за умови монотонності оператора A.

Метод

$$Jx_{n+1} = Jx_n - \lambda_n Ax_n$$

у цьому випадку може слабко збігатись лише в ергодичному розумінні. А за об'ємом необхідних для здійснення ітераційного кроку обчислень процес (8) має переваги над екстаградієнтним методом

$$\begin{cases} Jx_{n+\frac{1}{2}} = Jx_n - \lambda_n Ax_n, \\ Jx_{n+1} = Jx_n - \lambda_n Ax_{n+\frac{1}{2}}, \end{cases}$$

та неявним методом, що використовує резольвенту

$$x_{n+1} = \left(J + \lambda_n A\right)^{-1} Jx_n.$$

Схему (8) можна подати у вигляді двоетапного процесу

$$Jx_n = Jy_n - \lambda_{n-1}Ax_{n-1},$$

$$Jy_{n+1} = Jy_n - \lambda_{n-1}Ax_n.$$

Тобто, за відсутності обмежень ($C = E$) алгоритм 1 співпадає з алгоритмом екстраполяції з минулого [7].

Адаптивний алгоритм операторної екстраполяції

Параметри λ_n алгоритму 1 задавалась виходячи з умови

$$0 < \inf_n \lambda_n \le \sup_n \lambda_n < \tfrac{1}{2\mu L}.$$

Тобто, використовувалась інформація про константи ліпшицевості оператора A.

Відштовхуючись від алгоритму 1 та робіт [6, 7] у статті [8] побудовано алгоритм з адаптивним вибором величини λ_n, що не вимагає знання ліпшицевих констант та процедур типу лінійного пошуку.

У даному повідомленні наведемо результат про збіжність даного алгоритму, отриманий за допомогою ляпуновського аналізу.

Припустимо, що відома лише константа $\mu \ge 1$ з леми 1.

Алгоритм 2.

0. Обираємо $x_0 = x_1 \in E$, $\tau \in \left(0, \frac{1}{2\mu}\right)$ та число $\lambda_1 = \lambda_0 > 0$. Покладаємо $n = 1$.

1. Обчислити

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

2. Якщо $x_{n+1} = x_n = x_{n-1}$, то СТОП, інакше перейти до **3**.

3. Обчислити

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau\frac{\|x_{n+1}-x_n\|}{\|Ax_{n+1}-Ax_n\|_*}\right\}, & \text{якщо} \quad Ax_{n+1} \ne Ax_n, \\ \lambda_n, & \text{інакше.} \end{cases}$$

Покласти $n := n + 1$ та перейти до 1.

Послідовність (λ_n), що задається на третьому етапі ітераційного кроку в алгоритмі 2, незростаюча та обмежена знизу числом $\min\left\{\lambda_1, \frac{\tau}{L}\right\}$. Отже, існує $\lim\limits_{n\to\infty} \lambda_n > 0$.

В якості функції Ляпунова оберемо

$$W_n = \phi\left(z, x_n\right) + 2\left\langle m_n, x_n - z\right\rangle + \tau\mu\frac{\lambda_{n-1}}{\lambda_n} L\phi\left(x_n, x_{n-1}\right),$$

де $z \in S$.

Лема 5. *Для послідовності (x_n), що породжена алгоритмом 2, виконується нерівність*

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

Теорема 3. *Нехай C — непорожня опукла та замкнена підмножина 2-рівномірно опуклого та рівномірно гладкого банахового простору E, $A : E \to E^*$ — монотонний та ліпшицевий оператор, $S \ne \emptyset$. Припустимо, що нормалізоване дуальне відображення J секвенційно слабко неперервне. Тоді послідовність (x_n), що породжена алгоритмом 2, слабко збігається до деякої точки $z \in S$.*

Для операторного рівняння (7) алгоритм 2 дає такий ітераційний процес

$$\begin{cases} Jx_{n+1} = Jx_n - \lambda_n Ax_n - \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), \\ \lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau\frac{\|x_{n+1}-x_n\|}{\|Ax_{n+1}-Ax_n\|_*}\right\}, & \text{якщо } Ax_{n+1} \neq Ax_n, \\ \lambda_n, & \text{інакше.} \end{cases} \end{cases}$$

Література

1. Nagurney A. Network economics: A variational inequality approach. Dordrecht: Kluwer Academic Publishers, 1999. 325 p.
2. Киндерлерер Д., Стампаккья Г. Введение в вариационные неравенства и их приложения. Москва: Мир, 1983. 256 с.
3. Facchinei F., Pang J.-S. Finite-Dimensional Variational Inequalities and Complementarily Problem. V. 2. New York: Springer, 2003. 666 p.
4. 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.
5. Gidel G., Berard H., Vincent P., Lacoste-Julien S. A Variational Inequality Perspective on Generative Adversarial Networks. *preprint arXiv:1802.10551.* 2018.
6. Vedel Y., Semenov V. Adaptive Extraproximal Algorithm for the Equilibrium Problem in Hadamard Spaces. In: Olenev N., Evtushenko Y., Khachay M., Malkova V. (eds.) Optimization and Applications. OPTIMA 2020. Lecture Notes in Computer Science, vol 12422. Springer, Cham, 2020. P. 287–300.
7. Semenov V. V., Denisov S. V., Kravets A. V. Adaptive Two-Stage Bregman Method for Variational Inequalities. *Cybernetics and Systems Analysis.* 2021. Vol. 57. Issue 6. P. 959–967.
8. Vedel Y., Semenov V., Denisov S. A Novel Algorithm with Self-adaptive Technique for Solving Variational Inequalities in Banach Spaces. In: Olenev N. N., Evtushenko Y. G., Jaćimović M., Khachay M., Malkova V. (eds.) Advances in Optimization and Applications. OPTIMA 2021. Communications in Computer and Information Science, vol 1514. Springer, Cham, 2021. P. 50–64.
9. Malitsky Y., Tam M. K. A Forward-Backward Splitting Method for Monotone Inclusions Without Cocoercivity. *SIAM J. on Optim.* 2020. Vol. 30. P. 1451–1472.
10. Alber Y., Ryazantseva I. Nonlinear Ill Posed Problems of Monotone Type. Dordrecht: Springer, 2006. 410 p.
11. Alber Y. I. Metric and generalized projection operators in Banach spaces: properties and applications. In: Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, vol. 178. New York: Dekker, 1996. P. 15–50.
12. Beauzamy B. Introduction to Banach Spaces and Their Geometry. Amsterdam: North-Holland, 1985. 307 p.
13. Aoyama K., Kohsaka F. Strongly relatively nonexpansive sequences generated by firmly nonexpansive-like mappings. *Fixed Point Theory Appl.* 2014. 95. https://doi.org/10.1186/1687-1812-2014-95.
14. Xu H. K. Inequalities in Banach spaces with applications. *Nonlinear Anal.* 1991. Vol. 16. Iss. 12. P. 1127–1138.

Надійшла: 22.03.2022 / Прийнята: 27.07.2022