Збіжність методу операторної екстраполяції для варіаційних нерівностей

https://doi.org/10.31713/MCIT.2024.100

Володимир Семенов
КНУ імені Тараса Шевченка
м. Київ, Україна
semenov.volodya@gmail.com

Олександра Коваленко
КНУ імені Тараса Шевченка
м. Київ, Україна
alexandra.kovalenko@gmail.com

Денис Чергикало
КНУ імені Тараса Шевченка
м. Київ, Україна
denischergicalo@gmail.com

Анотація – **Розглядаються варіаційні нерівності з монотонними операторами, що діють в гільбертовому просторі. Доведено теореми збіжності та оцінки швидкості збіжності для різних варіантів методу операторної екстраполяції. Також запропоновано адаптивний та регуляризований варіанти методу.**

Ключові слова – **варіаційна нерівність; збіжність; метод операторної екстраполяції; монотонність**

I. Вступ

Розглянемо варіаційну нерівність:

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

де C – непорожня підмножина гільбертового простору H, A – оператор, що діє з H в H. Множину розв'язків (1) позначимо S.

Припустимо, що виконані такі умови: множина $C \subseteq E$ – опукла та замкнена; оператор $A : H \to H$ – монотонний та ліпшицевий на C з константою $L > 0$; множина S непорожня.

Задача (1) – зручна форма запису різних задач, що виникають в математичній фізиці, дослідженні операцій та машинному навчанні [1–3].

Окремі задачі негладкої оптимізації можна розв'язувати, якщо їх формулювати у вигляді сідлових задач і застосовувати алгоритми розв'язання варіаційних нерівностей [2]. А з початком використання GANs (генерувальних змагальних нейронних мереж) та інших моделей змагального навчання зацікавленість в алгоритмах розв'язання сідлових задач та варіаційних нерівностей з'явилася серед багатьох спеціалістів з машинного навчання [3].

Одним з методів апроксимації розв'язків варіаційних нерівностей (1) є алгоритм операторної екстраполяції [4–7].

Мета повідомлення – ознайомити з результатами про характер збіжності нових варіантів алгоритму операторної екстраполяції для задач з монотонними операторами [5–7].

Робота виконана за фінансової підтримки МОН України (проєкт «Обчислювальні алгоритми і оптимізація для штучного інтелекту, медицини та оборони», 0122U002026) та НАН України (проєкт «Нові субградієнтні та екстраградієнтні методи для негладких задач регресії», 0124U002162).

II. Операторна екстраполяція

Для розв'язання варіаційної нерівності (1) в роботі [4] запропоновано наступний алгоритм.

Алгоритм 1 ([4]).

Ініціалізація. *Обираємо* $x_0 = x_1 \in H$, $\lambda_n > 0$.

Ітерації. *Генеруємо послідовність* (x_n) *:*

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

де P_C *– оператор метричного проектування на* C.

Умова зупинки має вигляд

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

Зауважимо, що алгоритм 1 автори [4] називали «forward-reflected-backward algorithm».

Відносно додатніх параметрів λ_n припустимо виконання такої умови:

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

III. Слабка збіжість

Має місце

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

$$\|x_{n+1} - z\|^2 + 2\lambda_n \langle Ax_n - Ax_{n+1}, x_{n+1} - z \rangle +$$
$$+ \lambda_n L \|x_{n+1} - x_n\|^2 \leq$$

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

де $z \in S$.

Теорема 1. *Нехай C – непорожня опукла та замкнена підмножина гільбертового простору H, $A: H \to H$ – монотонний та ліпшицевий на множині C оператор, $S \neq \varnothing$, також виконується умова (2). Тоді послідовність (x_n), що породжена алгоритмом 1, слабко збігаються до деякої точки $z \in S$.*

У статті [5] аналогічний теоремі 1 результат отримано для задач в 2-рівномірно опуклих та рівномірно гладких банахових просторах

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 (1) будемо оцінювати за допогою невід'ємної функції зазору (gap function) [2]

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

Очевидно, що для коректності означення функції зазору необхідна обмеженість допустимої множини C. Якщо оператор A є монотонний та $x \in C$ – розв'язок задачі (1), то $\operatorname{gap}(x) = 0$. Навпаки, якщо для $x \in C$ маємо $\operatorname{gap}(x) = 0$, то x – розв'язок (1).

У випадку обмеженості допустимої множини C доведено, що алгоритму 1 необхідно зробити $O\left(LD^2\varepsilon^{-1}\right)$ ітерацій для отримання точки $x \in C$ з $\operatorname{gap}(x) \leq \varepsilon$, $\varepsilon > 0$, де $D = \operatorname{diam} C < +\infty$.

Теорема 2. *Нехай (x_n) – послідовність, що породжена алгоритмом 1. Нехай $\lambda_n \in \left(0, 1/2L\right]$. Тоді для послідовності чезарівських середніх*

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

має місце нерівність

$$\operatorname{gap}\left(z_{N+1}\right) \leq \frac{1}{2\sum_{n=1}^{N}\lambda_n}\sup_{y \in C}\left\|y - x_1\right\|^2.$$

Теорема 3. *Нехай (x_n) – послідовність, що породжена алгоритмом 1 с $\lambda_n = 1/2L$. Тоді для послідовності середніх*

$$z_{N+1} = \tfrac{1}{N}\sum_{n=1}^{N} x_{n+1}$$

має місце оцінка

$$\operatorname{gap}\left(z_{N+1}\right) \leq \frac{LD^2}{N}.$$

IV. Сильна збіжність

У випадку сильної монотонності оператора A можна гарантувати сильну збіжність алгоритму 1. Більш того, для окремих модифікацій алгоритму операторної екстраполяції доведено оцінки лінійної швидкості збіжності [6].

Наприклад, розглянемо такий алгоритм.

Алгоритм 2 ([6]).

Для $x_1 = x_0 \in C$ генеруємо послідовність (x_n):

$$x_{n+1} = P_C\left(x_n - \tfrac{1}{2L}Ax_n - \tfrac{1}{2(L+\mu)}\left(Ax_n - Ax_{n-1}\right)\right),$$

де $\mu, L > 0$ – константи сильної монотонності та ліпшицевості оператора A.

Має місце

Теорема 4. *Нехай C – непорожня опукла замкнена підмножина гільбертового простору H, $A: H \to H$ – ліпшицевий та сильно монотонний на множині C оператор. Тоді для породженої алгоритмом 2 послідовності (x_n) виконується оцінка*

$$\left\|x_{n+1} - z\right\|^2 \leq \left(1 - \tfrac{\mu}{L+\mu}\right)^n 2\left\|x_1 - z\right\|^2, \quad n \geq 1,$$

де $z \in C$ – єдиний розв'язок задачі (1).

У випадку відсутності сильної монотонності можна гарантувати лише слабку збіжність. Але спираючись на відомий ітераційний алгоритм Гальперна для апроксимації нерухомих точок нерозтягуючих операторів [1] побудовано такий регуляризований варіант алгоритму 1.

Алгоритм 3 ([7]).

Ініціалізація. *Задаємо елементи $y \in H$, $x_0, x_1 \in C$, послідовність чисел $\lambda_n > 0$ та таку послідовність (α_n), що $\alpha_n \in (0,1)$, $\lim_{n\to\infty}\alpha_n = 0$, $\sum_{n=1}^{\infty}\alpha_n = +\infty$.*

Ітерації. *Генеруємо послідовність (x_n):*

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

Відносно додатніх параметрів λ_n припустимо виконання умови (2).

Сформулюємо результат про сильну збіжність алгоритму 3.

Лема 2. *Для послідовності (x_n), що породжена алгоритмом 3, виконується нерівність*

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

$$-\left(1-\alpha_n\right)\left(\tfrac{1}{2}-\lambda_{n-1}L\right)\left\|x_n-x_{n-1}\right\|^2,$$

де $z\in S$.

Теорема 4. *Нехай C – непорожня опукла та замкнена підмножина гільбертового простору H, $A:H\to H$ – монотонний та ліпшицевий на множині C оператор, $S\neq\varnothing$, $y\in H$, виконується умова (2). Тоді послідовність $\left(x_n\right)$, що породжена алгоритмом 3, сильно збігається до точки $z=P_S y$.*

Параметри λ_n в алгоритмі 3 задавались виходячи з нерівності (2). Тобто, явно використовувалась інформація про константу ліпшицевості оператора A. Алгоритм 3 та адаптивна схема з статті [5] дозволяє побудувати такий алгоритм з адаптивним вибором величини λ_n, що не вимагає знання ліпшицевих констант операторів та процедур типу лінійного пошуку.

Алгоритм 4.

Ініціалізація. *Задаємо елементи $y\in H$, $x_0,x_1\in C$, числа $\tau\in\left(0,\tfrac{1}{2}\right)$, $\lambda_1,\lambda_0>0$, таку послідовність $\left(\alpha_n\right)$, що $\alpha_n\in\left(0,1\right)$, $\lim_{n\to\infty}\alpha_n=0$, $\sum_{n=1}^{\infty}\alpha_n=+\infty$.*

Ітерації. *Генеруємо послідовність $\left(x_n\right)$:*

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

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

Має місце аналогічний теоремі 4 про сильну збіжність алгоритму 4 до точки $z=P_S y$.

ЛІТЕРАТУРА

[1] Семенов В.В. Варіаційні нерівності: теорія та алгоритми. Київ: ВПЦ «Київський університет», 2021, 167 с.

[2] 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 Optim., vol. 15, 2004, pp. 229–251.

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

[4] Y. Malitsky, M.K. Tam, "A Forward-Backward Splitting Method for Monotone Inclusions Without Cocoercivity," SIAM J. on Optim., vol. 30, 2020, pp. 1451–1472.

[5] V.V. Semenov, S.V. Denisov, G.V. Sandrakov, O.S. Kharkov, "Convergence of the Operator Extrapolation Method for Variational Inequalities in Banach Spaces," Cybernetics and Systems Analysis, vol. 58, 2022, pp. 740–753.

[6] V.V. Semenov, O.S. Kharkov, "Convergence Rate of the Extrapolation from the Past and Operator Extrapolation Algorithms," Cybernetics and Systems Analysis, vol. 60, 2024, pp. 783–791.

[7] V.V. Semenov, O.S. Kharkov, "Strong Convergence of the Regularized Operator Extrapolation Algorihm for Variational Inequalities," Cybernetics and Systems Analysis, vol. 60, 2024, pp. 392–402.