Швидкість збіжності нових алгоритмів для варіаційних нерівностей

https://doi.org/10.31713/MCIT.2023.048

Олег Харьков
Факультет комп'ютерних наук та кібернетики,
КНУ імені Тараса Шевченка,
м. Київ, Україна
olehharek@gmail.com

Володимир Семенов
Факультет комп'ютерних наук та кібернетики,
КНУ імені Тараса Шевченка,
м. Київ, Україна
semenov.volodya@gmail.com

Анотація—**Досліджено збіжність алгоритмів для розв'язання варіаційних нерівностей в гільбертових просторах. Доведено лінійну швидкість збіжності алгоритмів екстраполяції з минулого та операторної екстраполяції для варіаційних нерівностей з операторами, що задовольняють умову типу узагальненої сильної монотонності.**

Ключові слова—варіаційна нерівність; лінійна збіжність; метод екстраполяції з минулого; метод операторної екстраполяції.

I. Вступ

Нехай H – дійсний гільбертовий простір з скалярним добутком $(\cdot,\cdot)$ та породженою нормою $\|\cdot\|$. Розглянемо варіаційну нерівність:

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

де C – непорожня опукла та замкнена підмножина простору H, A – оператор, що діє в просторі H. Будемо вважати, що оператор A є ліпшицевим на множині C (з константою $L>0$), тобто

$$\|Ax - Ay\| \le L\|x-y\| \quad \forall x, y \in C.$$

Задача (1) – зручна загальна форма запису різних задач, що виникають в математичній фізиці, дослідженні операцій та машинному навчанні [1, 2]. Окремі задачі негладкої оптимізації можна ефективно розв'язувати, якщо їх формулювати у вигляді сідлових задач і застосовувати алгоритми розв'язання варіаційних нерівностей. А з початком широкого використання генерувальних змагальних нейронних мереж та інших моделей змагального або робастного навчання інтерес до алгоритмів розв'язання сідлових задач та варіаційних нерівностей з'явився серед спеціалістів з машинного навчання [1]. Популярними методами апроксимації розв'язків варіаційних нерівностей (1) є алгоритм екстраполяції з минулого [1, 2] та алгоритм операторної екстраполяції [3, 4].

Мета повідомлення – ознайомити з результатами про лінійну швидкість збіжності алгоритмів для задач з операторами, що задовольняють умову типу узагальненої сильної монотонності. Для сильно монотонних та ліпшицевих операторів подібні результати отримано в [2, 5].

Припустимо, що існує єдиний розв'язок $z \in C$ варіаціної нерівності (1), а оператор A задовольняє умову

$$(Ax, x-z) \ge \mu\|x-z\|^2 \quad \forall x \in C \qquad (2)$$

для деякого $\mu > 0$.

Умова (2) випливає з сильної монотонності оператора A. Але існують немонотонні оператори з властивістю (2). Наприклад [6],

$$Ax = (2-\|x\|)x, \quad x \in C = \left\{x \in \ell_2 : \|x\| \le \tfrac{3}{2}\right\}.$$

II. Екстраполяція з минулого

Розглянемо такий варіант алгоритму екстраполяції з минулого. Для $x_1 = y_0 \in H$ генеруємо послідовність елементів x_n, y_n:

$$\begin{cases} y_n = P_C\left(x_n - \dfrac{1}{4L}Ay_{n-1}\right), \\ x_{n+1} = P_C\left(x_n - \dfrac{1}{4L}Ay_n\right), \end{cases} \qquad (3)$$

де P_C – оператор метричного проектування на C.

Має місце

Теорема 1. *Нехай C – непорожня опукла замкнена підмножина гільбертового простору H, $A: H \to H$ – ліпшицевий (з константою $L>0$) на множині C оператор, існує єдиний розв'язок $z \in C$ варіаціної нерівності (1) та виконується умова (2). Тоді для породжених алгоритмом (3) послідовностей (x_n), (y_n) виконується оцінка*

$$\|x_{n+1}-z\|^2 + \frac{1}{4}\|y_n - x_{n+1}\|^2 \le \left(1 - \frac{\mu}{4L}\right)^n \|x_1 - z\|^2, \quad n \ge 1.$$

Доведення. Для схеми

$$\begin{cases} y_n = P_C\left(x_n - \lambda Ay_{n-1}\right), \\ x_{n+1} = P_C\left(x_n - \lambda Ay_n\right), \end{cases}$$

де $\lambda > 0$, запишемо нерівність [1]:

$$\|x_{n+1}-z\|^2 \le \|x_{n+1}-z\|^2 - \|x_n - y_n\|^2 -$$

$$-\left\|x_{n+1}-y_n\right\|^2+2\lambda\left(Ay_n,z-y_n\right)+$$

$$+2\lambda\left(Ay_{n-1}-Ay_n,x_{n+1}-y_n\right).$$

З умови (2) та нерівності

$$\left\|a+b\right\|^2\le 2\left\|a\right\|^2+2\left\|b\right\|^2$$

випливає

$$\left(Ay_n,y_n-z\right)\ge\mu\left\|y_n-z\right\|^2\ge$$

$$\ge\mu\left(\frac{1}{2}\left\|x_n-z\right\|^2-\left\|y_n-x_n\right\|^2\right),$$

а ліпшицевість A гарантує виконання нерівності

$$\left(Ay_{n-1}-Ay_n,x_{n+1}-y_n\right)\le L\left\|y_{n-1}-y_n\right\|\left\|x_{n+1}-y_n\right\|\le$$

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

Використавши вищенаведені оцінки, отримуємо

$$\left\|x_{n+1}-z\right\|^2\le\left(1-\lambda\mu\right)\left\|x_{n+1}-z\right\|^2-\left(1-2\lambda\mu\right)\left\|x_n-y_n\right\|^2-$$

$$-\left\|x_{n+1}-y_n\right\|^2+\lambda L\left\|y_{n-1}-y_n\right\|^2+\lambda L\left\|x_{n+1}-y_n\right\|^2 \quad (4).$$

Член $\left\|y_{n-1}-y_n\right\|^2$ оцінимо таким чином

$$\left\|y_{n-1}-y_n\right\|^2\le 2\left\|y_{n-1}-x_n\right\|^2+2\left\|y_n-x_n\right\|^2.$$

Ураховуючи цю оцінку в (4), приходимо до нерівності

$$\left\|x_{n+1}-z\right\|^2\le\left(1-\lambda\mu\right)\left\|x_{n+1}-z\right\|^2-$$

$$-\left(1-2\lambda L-2\lambda\mu\right)\left\|x_n-y_n\right\|^2-$$

$$-\left(1-2\lambda L\right)\left\|x_{n+1}-y_n\right\|^2+2\lambda L\left\|x_n-y_{n-1}\right\|^2.$$

Перепишемо останню нерівність у вигляді

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

$$\le\left(1-\lambda\mu\right)\left(\left\|x_n-z\right\|^2+\frac{2\lambda L}{1-\lambda\mu}\left\|x_n-y_{n-1}\right\|^2\right)-$$

$$-\left(1-2\lambda L-2\lambda\mu\right)\left\|x_n-y_n\right\|^2-$$

$$-\left(1-2\lambda L-\frac{2\lambda L}{1-\lambda\mu}\right)\left\|x_{n+1}-y_n\right\|^2.$$

Для $\lambda=\dfrac{1}{4L}$ отримуємо

$$\left\|x_{n+1}-z\right\|^2+\frac{L}{4L-2\mu}\left\|x_{n+1}-y_n\right\|^2\le$$

$$\le\left(1-\frac{\mu}{4L}\right)\left(\left\|x_n-z\right\|^2+\frac{L}{4L-2\mu}\left\|x_n-y_{n-1}\right\|^2\right).$$

Звідки випливає потрібна оцінка. ■

З теореми 1 випливають оцінки

$$\left\|x_{n+1}-z\right\|^2\le e^{-\frac{\mu}{4L}n}\left\|x_1-z\right\|^2,$$

$$\left\|x_{n+1}-y_n\right\|^2\le e^{-\frac{\mu}{4L}n}4\left\|x_1-z\right\|^2.$$

III. Операторна екстраполяція

Розглянемо тепер варіант алгоритму операторної екстраполяції. Для $x_1=x_0\in C$ генерируємо послідовність елементів x_n:

$$x_{n+1}=P_C\left(x_n-\frac{1}{2L}Ax_n-\frac{1}{2\left(L+\mu\right)}\left(Ax_n-Ax_{n-1}\right)\right). \quad (5)$$

Теорема 2. *Нехай C – непорожня опукла замкнена підмножина гільбертового простору H, $A:H\to H$ – ліпшицевий (з константою $L>0$) на множині C оператор, існує єдиний розв'язок $z\in C$ варіаціної нерівності (1) та виконується умова (2). Тоді для породженої алгоритмом (5) послідовності $\left(x_n\right)$ виконується оцінка*

$$\left\|x_{n+1}-z\right\|^2\le\left(1-\frac{\mu}{L+\mu}\right)^n 2\left\|x_1-z\right\|^2,\quad n\ge 1.$$

Зауважимо, що для задачі пошуку нуля сильно монотонного та ліпшицевого оператора $A:H\to H$ в [5] для алгоритму

$$x_{n+1}=x_n-\frac{1}{8L}Ax_n-\frac{1}{8L}\left(Ax_n-Ax_{n-1}\right)$$

отримана оцінка

$$\left\|x_{n+1}-z\right\|^2\le\left(1-\frac{\mu}{8L}\right)^n 16384\frac{L^2}{\mu^2}\left\|x_1-z\right\|^2,\quad n\ge 1.$$

Література

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

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

[3] 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.

[4] V. V. Semenov, S. V. Denisov, G. V. Sandrakov, and O. S. Kharkov, "Convergence of the Operator Extrapolation Method for Variational Inequalities in Banach Spaces," Cybernetics and Systems Analysis, vol. 58 (6), 2022, pp. 740–753.

[5] A. Mokhtari, A. Ozdaglar, and S. Pattathil, "A unified analysis of extra-gradient and optimistic gradient methods for saddle point problems: proximal point approach," arXiv preprint arXiv:1901.08511, 2019.

[6] P. D. Khanh, and P. T. Vuong, "Modified projection method for strongly pseudomonotone variational inequalities," J. Glob. Optim., vol. 58, 2014, pp. 341–350.