ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2019, том 59, № 5, с. 889–894

УДК 519.217

АДАПТИВНЫЙ ПРОКСИМАЛЬНЫЙ МЕТОД ДЛЯ ВАРИАЦИОННЫХ НЕРАВЕНСТВ[1)]

© 2019 г. А. В. Гасников[1,2,3,*], П. Е. Двуреченский[3,4,**], Ф. С. Стонякин[5,1,***], А. А. Титов[1,****]

([1] 114070 Долгопрудный, М.о., Институтский пер., 9, МФТИ, Россия;

[2] 125319 Москва, Кочновский пр., 3, Высшая школа экономики, Россия;

[3] 127051 Москва, Большой Каретный пер., 19, Ин-т проблем передачи информации РАН, Россия;

[4] 10117 Berlin, Mohrenstr, 39, Weierstrass Institute for Applied Analysis and Stochastics, Германия;

[5] 295007 Симферополь, пр-т акад. Вернадского, 4, Крымский федеральный ун-т, Россия)

*e-mail: gasnikov@yandex.ru;

**e-mail: pavel.dvurechensky@gmail.com;

***e-mail: fedyor@mail.ru;

****e-mail: a.a.titov@phystech.edu

Поступила в редакцию 11.12.2017 г.
Переработанный вариант 19.12.2018 г.
Принята к публикации 19.12.2018 г.

Предлагается новый аналог проксимального зеркального метода А.С. Немировского с адаптивным выбором констант в минимизируемых прокс-отображениях на каждой итерации для вариационных неравенств с липшицевым полем. Получены оценки необходимого числа итераций для достижения заданного качества решения вариационного неравенства. Показано, как можно обобщить предлагаемый подход на случай гельдерова поля. Рассмотрена модификация предлагаемого алгоритма в случае неточного оракула для оператора поля. Библ. 17.

Ключевые слова: вариационное неравенство, проксимальный метод, адаптивный метод, гельдеров оператор поля, неточный оракул.

DOI: 10.1134/S0044466919050077

Вариационные неравенства (ВН) нередко возникают в самых разных проблемах оптимизации и имеют многочисленные приложения в математической экономике, математическом моделировании транспортных потоков, теории игр и других разделах математики (см., например [1]). Исследования в области методики решения ВН активно продолжаются (см., например, [2]–[12]).

Наиболее известным аналогом градиентного метода для ВН является экстраградиентный метод Г.М. Корпелевич [13]. Одним из современных вариантов экстраградиентного метода является проксимальный зеркальный метод А.С. Немировского [10]. Недавно Ю.Е. Нестеровым в [14] предложен новый алгоритм решения задач выпуклой минимизации с адаптивным выбором констант, который в случае липшицевости градиента целевой функции не требует знания этой константы Липшица (см. также [15], разд. 5). В настоящей статье мы на базе разработанного нами критерия (5) предлагаем похожий аналог проксимального зеркального метода А.С. Немировского для решения вариационных неравенств (алгоритм 1). Также мы распространяем предлагаемую методику на постановку задачи нахождения решения вариационного неравенства с неточным оракулом для оператора поля (алгоритм 2).

[1)]Исследование Ф.С. Стонякина, связанное с доказательством теоремы 2, алгоритмом 1, а также замечанием 3, выполнено при финансовой поддержке гранта РНФ (проект 18-71-00048). Исследование А.В. Гасникова, связанное с алгоритмом 2 и доказательством теоремы 3, выполнено при финансовой поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских ученых-докторов наук МД-1320.2018.1, а также при поддержке программы 5 топ 100 НИУ ВШЭ. Исследование П.Е. Двуреченского, связанное с алгоритмом 2 и доказательством теоремы 3, выполнено при поддержке грантов РФФИ 18-29-03071_мк и 18-31-20005 мол_а_вед.

Для некоторого оператора $g : Q \to \mathbb{R}^n$, заданного на выпуклом компакте $Q \subset \mathbb{R}^n$, будем рассматривать *сильные вариационные неравенства* вида

$$\left\langle g(x_*), x_* - x \right\rangle \le 0, \tag{1}$$

где g удовлетворяет условию Липшица. Отметим, что в (1) требуется найти $x_* \in Q$ (это x_* и называется решением ВН), для которого

$$\max_{x \in Q} \left\langle g(x_*), x_* - x \right\rangle \le 0. \tag{2}$$

В случае монотонного поля g наш подход позволяет рассматривать также *слабые вариационные неравенства*

$$\left\langle g(x), x_* - x \right\rangle \le 0. \tag{3}$$

Обычно в (3) требуется найти $x_* \in Q$, для которого (3) верно при всех $x \in Q$.

Начнем с некоторых вспомогательных понятий, соглашений и обозначений. Всюду далее будем полагать, что множество Q выпукло и компактно в пространстве $\mathbb{R}^n$ с нормой $\|\cdot\|$ (вообще говоря, не евклидовой), а $\|\cdot\|_*$ – сопряженная к $\|\cdot\|$ норма. Пусть задана 1-сильно выпуклая функция d относительно $\|\cdot\|$, которая дифференцируема во всех точках $x \in Q$. Можно ввести соответствующую d *дивергенцию Брэгмана*

$$V(x, y) = d(x) - d(y) - \left\langle \nabla d(y), x - y \right\rangle \quad \forall x, y \in Q, \tag{4}$$

где $\langle \cdot, \cdot \rangle$ – скалярное произведение в $\mathbb{R}^n$.

Для решения задачи (1), (2) мы предлагаем следующий адаптивный проксимальный зеркальный метод (АПЗМ). Пусть заданы число $\varepsilon > 0$ (точность решения) и начальное приближение $x^0 = \arg\min_{x \in Q} d(x) \in Q$.

Опишем $(N+1)$-ю итерацию предлагаемого алгоритма ($N = 0, 1, 2, \ldots$), положив изначально $N := 0$.

Алгоритм 1: АПЗМ

Шаг 1. $N := N + 1$, $L^{N+1} := L^N / 2$.

Шаг 2. Вычисляем

$$y^{N+1} := \arg\min_{x \in Q} \left\{ \left\langle g(x^N), x - x^N \right\rangle + L^{N+1} V(x, x^N) \right\},$$

$$x^{N+1} := \arg\min_{x \in Q} \left\{ \left\langle g(y^{N+1}), x - x^N \right\rangle + L^{N+1} V(x, x^N) \right\}.$$

Шаг 3. Если верно

$$\left\langle g(y^{N+1}) - g(x^N), y^{N+1} - x^{N+1} \right\rangle \le L^{N+1} V(y^{N+1}, x^N) + L^{N+1} V(x^{N+1}, y^{N+1}), \tag{5}$$

то переходим к следующей итерации (шаг 1).

Шаг 4. Если не выполнено (5), то увеличиваем L^{N+1} в 2 раза: $L^{N+1} := 2L^{N+1}$ и переходим к шагу 2.

Шаг 5. Критерий остановки метода:

$$\sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \ge \frac{R^2}{\varepsilon}, \qquad \text{где} \qquad R^2 = \max_{x \in Q} V(x, x^0).$$

Теорема 1. *После остановки алгоритма* 1 *справедлива оценка*:

$$\sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \left\langle g(y^{k+1}), y^{k+1} - x \right\rangle \le V(x, x^0) - V(x, x^N). \tag{6}$$

Доказательство. Непосредственно можно проверить равенства:

$$\left\langle \nabla_x V(x, x^k)\Big|_{x=x^{k+1}}, x - x^{k+1} \right\rangle = V(x, x^k) - V(x, x^{k+1}) - V(x^{k+1}, x^k), \tag{7}$$

$$\left\langle \nabla_x V(x, x^k)\Big|_{x=y^{k+1}}, x - y^{k+1} \right\rangle = V(x, x^k) - V(x, y^{k+1}) - V(y^{k+1}, x^k). \tag{8}$$

Далее, для всякого $x \in Q$ и $k = \overline{0, N-1}$ верны неравенства:

$$\left\langle \nabla_x \left(\left\langle g(x^k), x - x^k \right\rangle + L^{k+1} V(x, x^k) \right)\Big|_{x=y^{k+1}}, x - y^{k+1} \right\rangle \ge 0,$$

$$\left\langle \nabla_x \left(\left\langle g(y^{k+1}), x - x^k \right\rangle + L^{k+1} V(x, x^k) \right)\Big|_{x=x^{k+1}}, x - x^{k+1} \right\rangle \ge 0.$$

Поэтому

$$\left\langle g(y^{k+1}), x^{k+1} - x \right\rangle \le L^{k+1} V(x, x^k) - L^{k+1} V(x, x^{k+1}) - L^{k+1} V(x^{k+1}, x^k)$$

и

$$\left\langle g(x^k), y^{k+1} - x \right\rangle \le L^{k+1} V(x, x^k) - L^{k+1} V(x, y^{k+1}) - L^{k+1} V(y^{k+1}, x^k),$$

откуда для всякого $k = \overline{0, N-1}$ с учетом (5) мы имеем

$$\begin{aligned} \left\langle g(y^{k+1}), y^{k+1} - x \right\rangle &= \left\langle g(y^{k+1}), x^{k+1} - x \right\rangle + \left\langle g(x^k), y^{k+1} - x^{k+1} \right\rangle + \left\langle g(y^{k+1}) - g(x^k), y^{k+1} - x^{k+1} \right\rangle \le \\ &\le L^{k+1} V(x, x^k) - L^{k+1} V(x, x^{k+1}) - L^{k+1} V(x^{k+1}, x^k) + L^{k+1} V(x^{k+1}, x^k) - \\ &- L^{k+1} V(x^{k+1}, y^{k+1}) - L^{k+1} V(y^{k+1}, x^k) + L^{k+1} V(y^{k+1}, x^k) + L^{k+1} V(x^{k+1}, y^{k+1}), \end{aligned}$$

т.е. верно неравенство

$$\frac{1}{L^{k+1}} \left\langle g(y^{k+1}), y^{k+1} - x \right\rangle \le V(x, x^k) - V(x, x^{k+1}). \tag{9}$$

После суммирования неравенств (9) по $k = \overline{0, N-1}$ получим

$$\sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \left\langle g(y^{k+1}), y^{k+1} - x \right\rangle \le V(x, x^0) - V(x, x^N),$$

что и требовалось.

Допустим, что векторное поле g удовлетворяет условию

$$\|g(x) - g(y)\| \le L \|x - y\| \quad \forall x, y \in Q. \tag{10}$$

Всюду далее для краткости будем обозначать

$$S_N = \sum_{k=0}^{N-1} \frac{1}{L^{k+1}}.$$

Условие (10) означает, что для произвольных $x, y, z \in Q$

$$\left\langle g(y) - g(z), y - z \right\rangle \le \|g(y) - g(x)\|_* \|y - z\| \le L \|y - x\| \|y - z\|,$$

откуда в силу справедливого для всех $a, b \in \mathbb{R}$ неравенства $ab \le \frac{a^2}{2} + \frac{b^2}{2}$ при произвольных $x, y, z \in Q$ верно

$$\left\langle g(y) - g(x), y - z \right\rangle \le \frac{L}{2} \|y - x\|^2 + \frac{L}{2} \|y - z\|^2 \le LV(y, x) + LV(z, y).$$

Тогда ввиду (6) для всякого $x \in Q$

$$\min_{k=\overline{0,N-1}} \left\langle g(y^{k+1}), y^{k+1} - x \right\rangle \le \frac{1}{S_N} \sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \left\langle g(y^{k+1}), y^{k+1} - x \right\rangle \le \frac{1}{S_N} (V(x, x^0) - V(x, x^N)) \le \frac{R^2}{S_N},$$

где

$$R^2 = \max_{x\in Q} V(x, x^0).$$

Если потребовать, чтобы

$$\max_{x\in Q} \min_{\overline{k=0,N-1}} \left\{\left\langle g(y^{k+1}), y^{k+1} - x\right\rangle\right\} \le \frac{1}{S_N} \max_{x\in Q} \sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \left\langle g(y^{k+1}), y^{k+1} - x\right\rangle \le \varepsilon, \tag{11}$$

то в случае $L^0 \le 2L$ можно оценить количество итераций работы алгоритма 1. Действительно, $L^0 \le 2L$ означает, что $L^{k+1} \le 2L$ для всякого $k = \overline{0, N-1}$ и поэтому верно

$$\frac{R^2}{S_N} \le \varepsilon \quad \text{при} \quad N \ge \frac{2LR^2}{\varepsilon}.$$

Таким образом, справедлива следующая

Теорема 2. *Если выполнено условие* (10) *для поля* g, *то алгоритм* 1 *работает не более*

$$N = \left\lceil \frac{2LR^2}{\varepsilon} \right\rceil \tag{12}$$

итераций. Если выбрать $L^0 \le 2L$, *то верно* (11).

Замечание 1. Если $g \not\equiv 0$, то выполнения условия $L^0 \le 2L$ можно добиться, выбрав

$$L^0 := \frac{\|g(x) - g(y)\|_*}{\|x - y\|} \quad \text{при} \quad g(x) \ne g(y).$$

Замечание 2. Заметим, что оценка числа итераций (12) с точностью до числового множителя оптимальна для вариационных неравенств [9], [16], [17].

Замечание 3. Отметим, что похожими рассуждениями можно получить аналог теоремы 2 для гельдерова поля g. Точнее говоря, если справедливо

$$\|g(x) - g(y)\|_* \le L_\nu \|x - y\|^\nu \tag{13}$$

при $\nu \in [0;1]$, $x, y \in Q$, причем $L_0 < +\infty$ (другие константы L_ν могут быть бесконечными), то утверждение теоремы 2 выполняется, если

$$L^0 \le 2L = 2 \inf_{\nu\in[0;1]} L_\nu \left(\frac{2L_\nu}{\varepsilon}\right)^{\frac{1-\nu}{1+\nu}} \tag{14}$$

и

$$N = \left\lceil \inf_{\nu\in[0;1]} \left(\frac{2L_\nu R^{1+\nu}}{\varepsilon}\right)^{\frac{2}{1+\nu}} \right\rceil. \tag{15}$$

Замечание 4. Для слабых ВН (4) имеем

$$\left\langle g(x), y^{k+1} - x\right\rangle = \left\langle g(y^{k+1}), y^{k+1} - x\right\rangle + \left\langle g(x) - g(y^{k+1}), y^{k+1} - x\right\rangle \le \left\langle g(y^{k+1}), y^{k+1} - x\right\rangle,$$

поэтому критерий (11) можно заменить на

$$\max_{x\in Q} \left\langle g(x), \tilde{y} - x\right\rangle \le \varepsilon, \quad \text{где} \quad \tilde{y} = \frac{\sum_{k=0}^{N-1} \frac{1}{L^{k+1}} y^{k+1}}{S_N}. \tag{16}$$

Отметим, что именно оценку вида (16) обычно используют как критерий качества решения слабого ВН (см., например [10]).

В завершение отметим, что рассматриваемую в работе методику можно распространить на случай неточного оракула $\tilde{g}(x, \delta_c, \delta_u)$ для поля g. Поясним смысл, который мы вкладываем в это понятие.

Будем полагать, что существует фиксированное $\delta_u > 0$ такое, что для всякого $\delta_c > 0$ существует константа $L(\delta_c) \in (0, +\infty)$, для которой $\forall x, y, z \in Q$ верно

$$\langle \tilde{g}(y, \delta_c, \delta_u) - \tilde{g}(x, \delta_c, \delta_u), y - z \rangle \le L(\delta_c)\big(V(y, x) + V(z, y)\big) + \delta_c + \delta_u. \tag{17}$$

Замечание 5. Заметим, что мы здесь считаем, что δ_c – контролируемая погрешность оракула (ее возможно бесконечно уменьшать и для заданной точности решения $\varepsilon > 0$ мы всюду будем полагать, что $\delta_c = \frac{\varepsilon}{2}$). С другой стороны, погрешность δ_u мы полагаем неконтролируемой.

Пример 1. Пусть поле g удовлетворяет условию Липшица и заданы его неточные значения на Q, т.е. существует $\overline{\delta}_u > 0$ и в произвольной точке $x \in Q$ можно вычислить величину $\overline{g}(x)$, удовлетворяющую условию

$$\|\overline{g}(x) - g(x)\|_* \le \overline{\delta}_u.$$

В таком случае $\overline{g}(x) = \tilde{g}(x, \delta_c, \delta_u)$ удовлетворяет (17), если положить $\delta_u = \overline{\delta}_u$, а также

$$\delta_c = 0 \quad \text{и} \quad L(\delta_c) \equiv L.$$

При условиях (17) мы предлагаем следующую модификацию алгоритма 1. Пусть заданы $\varepsilon > 0$ (точность решения) и начальное приближение

$$x^0 = \arg\min_{x \in Q} d(x) \in Q.$$

Опишем $(N+1)$-ю итерацию предлагаемого алгоритма $(N = 0, 1, 2, \ldots)$, положив изначально $N := 0$.

Алгоритм 2: "Неточный" АПЗМ

Шаг 1. $N := N + 1$, $L^{N+1} := L^N/2$.

Шаг 2. Вычисляем:

$$y^{N+1} := \arg\min_{x \in Q}\left\{\left\langle \tilde{g}\left(x^N, \frac{\varepsilon}{2}, \delta_u\right), x - x^N \right\rangle + L^{N+1}V(x, x^N)\right\},$$

$$x^{N+1} := \arg\min_{x \in Q}\left\{\left\langle \tilde{g}\left(y^{N+1}, \frac{\varepsilon}{2}, \delta_u\right), x - x^N \right\rangle + L^{N+1}V(x, x^N)\right\}.$$

Шаг 3. Если верно

$$\left\langle \tilde{g}(y^{N+1}, \frac{\varepsilon}{2}, \delta_u) - \tilde{g}(x^N, \frac{\varepsilon}{2}, \delta_u), y^{N+1} - x^{N+1} \right\rangle \le L^{N+1}V(y^{N+1}, x^N) + L^{N+1}V(x^{N+1}, y^{N+1}) + \frac{\varepsilon}{2} + \delta_u,$$

то переходим к следующей итерации (шаг 1).

Шаг 4. Иначе увеличиваем в 2 раза константу $L^{N+1} := 2L^{N+1}$ и переходим к шагу 2.

Шаг 5. Критерий остановки метода

$$\sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \ge \frac{2R^2}{\varepsilon}, \quad \text{где} \quad R^2 = \max_{x \in Q} V(x, x^0).$$

Аналогично доказательству теоремы 2 проверяется

Теорема 3. *Пусть* $\tilde{g}$ *удовлетворяет* (17). *Тогда после остановки алгоритма* 2 *справедлива оценка*:

$$\frac{1}{S_N}\sum_{k=0}^{N-1}\frac{1}{L^{k+1}}\left\langle \tilde{g}\left(y^{k+1}, \frac{\varepsilon}{2}, \delta_u\right), y^{k+1} - x\right\rangle \le \frac{V(x, x^0) - V(x, x^N)}{S_N} + \frac{\varepsilon}{2} + \delta_u. \tag{18}$$

Неравенство (18) означает, что после работы алгоритма 2 будет верно:

$$\frac{1}{S_N}\sum_{k=0}^{N-1}\frac{1}{L^{k+1}}\left\langle \tilde{g}\left(y^{k+1}, \frac{\varepsilon}{2}, \delta_u\right), y^{k+1} - x\right\rangle \le \frac{2LR^2}{N} + \frac{\varepsilon}{2} + \delta_u, \tag{19}$$

где

$$R^2 = \max_{x,y\in Q} V(x,y).$$

При условии $L^0 \le 2L$ можно показать, что алгоритм 2 будет работать не более $N = \left\lceil \frac{4LR^2}{\varepsilon} \right\rceil$ итераций. После его остановки заведомо верно неравенство

$$\max_{x\in Q} \min_{k=0,N-1} \left\langle \tilde{g}\left(y^{k+1}, \frac{\varepsilon}{2}, \delta_u\right), y^{k+1} - x \right\rangle \le \frac{1}{S_N} \max_{x\in Q} \sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \left\langle \tilde{g}\left(y^{k+1}, \frac{\varepsilon}{2}, \delta_u\right), y^{k+1} - x \right\rangle \le \varepsilon + \delta_u,$$

которое может рассматриваться как критерий качества найденного решения с погрешностью δ_u.

СПИСОК ЛИТЕРАТУРЫ

1. *Facchinei F., Pang J.S.* Finite-Dimensional Variational Inequality and Complementarity Problems. New York: Springer, 2003. V. 1, 2. 693 p.
2. *Антипин А.С.* Методы решения вариационных неравенств со связанными ограничениями // Ж. вычисл. матем. и матем. физ. 2000. Т. 40. № 9. С. 1291–1307.
3. *Antipin A.* Gradient approach of computing fixed points of equilibrium problems // J. of Global Optimizat. 2002. V. 24. № 3. P. 285–309.
4. *Антипин А.С., Ячимович В., Ячимович М.* Динамика и вариационные неравенства // Ж. вычисл. матем. и матем. физ. 2017. Т. 57. № 5. С. 783–800.
5. *Коннов И.В.* Модель миграционного равновесия с обратными функциями полезности // Уч. зап. Казан. ун-та. Сер. Физ.-матем. науки. 2013. Т. 155. № 2. С. 91–99.
6. *Коннов И.В., Салахутдин Р.А.* Двухуровневый итеративный метод для нестационарных смешанных вариационных неравенств // Изв. вузов. Матем. 2017. № 10. С. 50–61.
7. *Меленьчук Н.В.* Методы и алгоритмы для решения задач математического моделирования на основе вариационных неравенств // Дисс. ... канд. физ.-матем. наук. Омск: Омский гос. техн. ун-т, 2011.
8. *Нестеров Ю.Е.* Алгоритмическая выпуклая оптимизация // Дисс. ... докт. физ.-матем. наук. М.: Моск. физ.-техн. ин-т (гос. ун-т), 2013.
9. *Guzman C., Nemirovski A.* On lower complexity bounds for large-scale smooth convex optimization // J. Complexity. 2015. V. 31. № 1. P. 1–14.
10. *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. V. 15. P. 229–251.
11. *Nesterov Yu.* Dual extrapolation and its application for solving variational inequalities and related problems // Math. Program. 2007. Ser. B. P. 319–344.
12. *Nesterov Yu., Scrimali L.* Solving strongly monotone variational and quasi-variational inequalities // Discrete and Continuous Dynamical Systems – A. 2011. V. 31. № 4. P. 1383–1396.
13. *Корпелевич Г.М.* Экстраградиентный метод для отыскания седловых точек и других задач // Экономика и матем. методы. Т. 12. № 4. С. 747–756.
14. *Nesterov Yu.* Universal gradient methods for convex optimization problems // Math. Program. 2015. V. 152. № 1–2. P. 381–404.
15. *Гасников А.В.* Современные численные методы оптимизации. Метод универсального градиентного спуска. М.: Изд-во МФТИ: 2018. 160 с. https://arxiv.org/abs/1711.00394.
16. *Немировский А.С., Юдин Д.Б.* Сложность задач и эффективность методов оптимизации. М.: Наука, 1979. 384 с.
17. *Nemirovski A.* Information-based complexity of convex programming. Technion–Israel Institute of Technology. Fall Semester 1994/95. 268 p. Available at http://www2.isye.gatech.edu/ nemirovs/Lec_EMCO.pdf.