

Адаптивный проксимальный метод для вариационных неравенств

*Ф. С. Стонякин^{1, 2}, А. В. Гасников^{2, 3, 4},
П. Е. Двуреченский^{4, 5}, А. А. Титов²*

¹Крымский федеральный университет имени В. И. Вернадского,
пр-кт Академика Вернадского, 4, Симферополь, РФ, 295007;

²Московский физико-технический институт,
Институтский пер., 9, Долгопрудный, РФ, 141070;

³Национальный исследовательский университет "Высшая школа экономики",
Кочновский проезд, 3, Москва, РФ, 125319;

⁴Институт проблем передачи информации имени А. А. Харкевича РАН,
Большой Каретный пер., 19, Москва, РФ, 127051;

⁵Weierstrass Institute for Applied Analysis and Stochastics,
Mohrenstr. 39, Berlin, Germany, 10117.

Аннотация

В статье предложен 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 гёльдерова поля. Рассмотрена модификация предлагаемого алгоритма в случае неточного оракула для оператора поля.

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 ре-

шения задач выпуклой минимизации с адаптивным выбором констант, который в случае липшицевости градиента целевой функции не требует знания этой константы Липшица (см. также раздел 5 пособия [15]). 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).

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

$$\langle g(x_*), x_* - x \rangle \leq 0, \quad (1)$$

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

$$\max_{x \in Q} \langle g(x_*), x_* - x \rangle \leq 0. \quad (2)$$

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

$$\langle g(x), x_* - x \rangle \leq 0. \quad (3)$$

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

Начнём 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 $\|\cdot\|_*$ — сопряжённая 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*

$$V(x, y) = d(x) - d(y) - \langle \nabla d(y), x - y \rangle \quad \forall x, y \in Q, \quad (4)$$

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 задано 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 ($N = 0, 1, 2, \dots$), положив изначально $N := 0$.

АЛГОРИТМ 1: АПЗМ

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

2. Вычисляем

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

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

3. Если верно $\langle g(y^{N+1}) - g(x^N), y^{N+1} - x^{N+1} \rangle \leq$

$$\leq L^{N+1} V(y^{N+1}, x^N) + L^{N+1} V(x^{N+1}, y^{N+1}), \quad (5)$$

то переходим к следующей итерации (пункт 1).

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

5. Критерий остановки метода: $\sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \geq \frac{R^2}{\varepsilon}$, где $R^2 = \max_{x \in Q} V(x, x^0)$.

Теорема 1. После остановки Алгоритма 1 справедлива оценка:

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

Доказательство. Непосредственно можно проверить равенства:

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

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

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

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

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

Поэтому

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

и

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

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

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

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

т.е. верно

$$\frac{1}{L^{k+1}} \langle g(y^{k+1}), y^{k+1} - x \rangle \leq V(x, x^k) - V(x, x^{k+1}). \quad (9)$$

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

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

что и требовалось. \square

Допустим, что векторное поле g удовлетворяет условию

$$\|g(x) - g(y)\|_* \leq L\|x - y\| \quad \forall x, y \in Q. \quad (10)$$

Всюду далее для краткости будем обозначать

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

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

$$\langle g(y) - g(z), y - z \rangle \leq \|g(y) - g(x)\|_* \cdot \|y - z\| \leq L\|y - x\| \cdot \|y - z\|,$$

откуда в силу справедливого для всех $a, b \in \mathbb{R}$, $\delta > 0$ неравенства $ab \leq \frac{a^2}{2} + \frac{b^2}{2}$ при произвольных $x, y, z \in Q$ верно

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

Тогда ввиду (6) для всякого $x \in Q$

$$\begin{aligned}
\min_{k=0, N-1} \langle g(y^{k+1}), y^{k+1} - x \rangle & \leq \frac{1}{S_N} \sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \langle g(y^{k+1}), y^{k+1} - x \rangle \leq \\
& \leq \frac{1}{S_N} (V(x, x^0) - V(x, x^N)) \leq \frac{R^2}{S_N},
\end{aligned}$$

где $R^2 = \max_{x \in Q} V(x, x^0)$.

Если потребовать, чтобы

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

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

$$\frac{R^2}{S_N} \leq \varepsilon \text{ верно при } N \geq \frac{2LR^2}{\varepsilon}.$$

Таким образом, справедлива следующая

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

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

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

Замечание 1. Если $g \neq 0$, то выполнения условия $L^0 \leq 2L$ можно добиться, выбрав

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

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 гёльдерова поля g . Точнее говоря, если справедливо

$$\|g(x) - g(y)\|_* \leq L_\nu \|x - y\|^\nu \quad (13)$$

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

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

и

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

Замечание 4. Для слабых ВН (4) имеем:

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

поэтому критерий (11) можно заменить на

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

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

В завершении 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

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

Замечание 5. Заметим, что мы здесь считаем, что δ_c — контролируемая погрешность оракула (её 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 $\bar{\delta}_u > 0$ и в произвольной точке $x \in Q$ можно вычислить величину $\bar{g}(x)$, удовлетворяющую условию $\|\bar{g}(x) - g(x)\|_* \leq \bar{\delta}_u$. В таком случае $\bar{g}(x) = \tilde{g}(x, \delta_c, \delta_u)$ удовлетворяет (17), если положить $\delta_u = \bar{\delta}_u$, а также $\delta_c = 0$ и $L(\delta_c) \equiv L$.

При условиях (17) мы предлагаем следующую модификацию Алгоритма 1. Пусть задано $\varepsilon > 0$ (точность решения) и начальное приближение $x^0 = \arg \min_{x \in Q} d(x) \in Q$.

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

АЛГОРИТМ 2: «Неточный» АПЗМ

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

2. Вычисляем:

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

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

3. Если верно $\langle \tilde{g}(y^{N+1}, \frac{\varepsilon}{2}, \delta_u) - \tilde{g}(x^N, \frac{\varepsilon}{2}, \delta_u), y^{N+1} - x^{N+1} \rangle \leq$

$$\leq L^{N+1} V(y^{N+1}, x^N) + L^{N+1} V(x^{N+1}, y^{N+1}) + \frac{\varepsilon}{2} + \delta_u,$$

то переходим к следующей итерации (пункт 1).

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

5. Критерий остановки метода: $\sum_{k=0}^{N-1} \frac{1}{L^{k+1}} \geq \frac{2R^2}{\varepsilon}$, где $R^2 = \max_{x \in Q} V(x, x^0)$.

Аналогично доказательству теоремы 2 проверяется следующее утверждение.

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

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

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

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

где $R^2 = \max_{x, y \in Q} V(x, y)$. При условии $L^0 \leq 2L$ 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:

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

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

Финансовая поддержка. Исследование Ф.С. Стонякаина, связанное с доказательством теоремы 2, Алгоритмом 1, а также замечанием 3, выполнено за счёт гранта Российского научного фонда (проект 18-71-00048). Исследование А.В. Гасникова, связанное с Алгоритмом 2 и доказательством теоремы 3, выполнено при поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских учёных-докторов наук МД-1320.2018.1, а также при поддержке программы 5 топ 100 НИУ ВШЭ.

Список литературы

- [1] F. Facchinei, J. S. Pang, Finite-Dimensional Variational Inequality and Complementarity Problems, vols. 1 and 2, //Springer-Verlag, New York, 2003.
- [2] А. С. Антипин. Методы решения вариационных неравенств со связанными ограничениями, // Ж. вычисл. матем. и матем. физ., 40:9 (2000), С. 1291 – 1307.
- [3] А. Antipin, Gradient approach of computing fixed points of equilibrium problems, //Journal of Global Optimization, 2002, 24(3), pp. 285–309.
- [4] А. С. Антипин, В. Ячимович, М. Ячимович. Динамика и вариационные неравенства, // Ж. вычисл. матем. и матем. физ., 57:5 (2017), С. 783–800.
- [5] И. В. Коннов, Модель миграционного равновесия с обратными функциями полезности, // Учен. зап. Казан. ун-та. Сер. Физ.-матем. науки, 155, № 2, Изд-во Казанского ун-та, Казань, 2013, С. 91–99.
- [6] И. В. Коннов, Раис Ахмад Салахутдин, Двухуровневый итеративный метод для нестационарных смешанных вариационных неравенств, // Изв. вузов. Матем., 2017, № 10, С. 50–61.
- [7] Н. В. Меленьчук. Методы и алгоритмы для решения задач математического моделирования на основе вариационных неравенств, // Дисс. ... канд. физ.-мат. наук: 05.13.18 — Омск: 2011. — 123 с.
- [8] Ю. Е. Нестеров. Алгоритмическая выпуклая оптимизация. // Дисс. ... докт. физ.-мат. наук: 01.01.07 / Нестеров Юрий Евгеньевич; [Место защиты: Моск. физ.-техн. ин-т (гос. ун-т)]. — Москва, 2013. — 367 с.: ил. РГБ ОД, 71 15-1/114.
- [9] C. Guzman, A. Nemirovski, On lower complexity bounds for large-scale smooth convex optimization, //Journal of Complexity, 2015, 31, pp. 1–14.
- [10] 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 Journal on Optimization, 2004, 15, pp. 229–251.

- [11] Yu. Nesterov, Dual extrapolation and its application for solving variational inequalities and related problems. // Math. Program., 2007, Ser. B, pp.319 – 344.
- [12] Yu. Nesterov, L. Scrimali, Solving strongly monotone variational and quasi-variational inequalities. // Discrete and Continuous Dynamical Systems, 31.10.2139/ssrn.970903, 2007.
- [13] Г.М. Корпелевич. Экстраградиентный метод для отыскания седловых точек и других задач // Экономика и матем. методы, 12: 4 (1976), С. 747 – 756.
- [14] Yu. Nesterov, Universal gradient methods for convex optimization problems, //Math. Program., 2015, Ser. A, 152, pp. 381–404.
- [15] А. В. Гасников. Современные численные методы оптимизации. Метод универсального градиентного спуска, М. Изд-во МФТИ: 2018, доступно по ссылке: <https://arxiv.org/abs/1711.00394>.
- [16] А. С. Немировский, Д.Б. Юдин. Сложность задач и эффективность методов оптимизации, М.: Наука. Главная редакция физико-математической литературы, 1979. 384 с.
- [17] А. Nemirovski, Information-based complexity of convex programming, //Technion, Fall Semester 1994/95, available at http://www2.isye.gatech.edu/nemirovs/Lec_EMCO.pdf.