Стонякин Ф. С.

Лекция. Вариационные неравенства и седловые задачи: разрешимость, связь множеств сильных и слабых решений. Проекционный и экстраградиентный методы.

19 мая 2025 г.

1. Постановка задачи, примеры, сходимость и качество решения

Сильное решение: $x_* \in Q$:

$$\langle g(x_*), x_* - x\rangle \leqslant 0 \quad \forall\, x \in Q. \tag{1}$$

Слабое решение: $x_* \in Q$:

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

Если оператор g **монотонен**

$$\langle g(y) - g(x), y - x\rangle \geqslant 0 \quad \forall\, x \in Q, \tag{3}$$

то

$$\langle g(x_*) - g(x), x_* - x\rangle \geqslant 0 \quad \forall\, x \in Q \tag{4}$$

и

$$\langle g(x_*), x_* - x\rangle \geqslant \langle g(x), x_* - x\rangle. \tag{5}$$

Для монотонных операторов всякое сильное решение будет слабым решением.

Выпукло-вогнутые седловые задачи

$$
\begin{gathered}
f(u_*, v) \leqslant f(u_*, v_*) \leqslant f(u, v_*), \\
x = (u_x, v_x), \quad g(x) = (\nabla_u f(x), -\nabla_v f(x)), \\
\begin{cases} \langle \nabla_u f(x_*), u_* - u \rangle \leqslant 0, \\ \langle \nabla_v f(x_*), v_* - v \rangle \geqslant 0. \end{cases}
\end{gathered} \tag{6}
$$

$$\langle \nabla_u f(x_*), u_* - u \rangle \leqslant 0,$$

$$\langle -\nabla_v f(x_*), v_* - v \rangle \leqslant 0.$$

$$\langle g(x_*), x_* - x \rangle \leqslant 0.$$

Но чаще проще найти слабое решение x_*: $\quad \langle g(x), x_* - x \rangle \leqslant 0$!

2. ВН и неподвижные точки проекционных операторов. О разрешимости ВН [1)]

$$P_\alpha(x) := Pr_Q(x - \alpha g(x)), \quad \alpha > 0. \tag{7}$$

$$P_\alpha(x_*) = x_* = Pr_Q(x_* - \alpha g(x_*)). \tag{8}$$

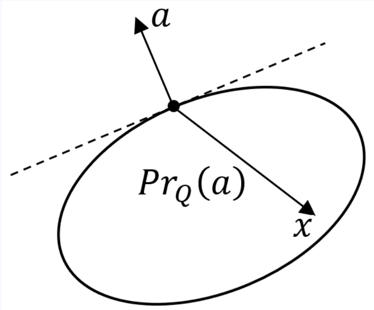

Рис. 1

Q — выпуклое замкнутое множество.

$Pr_Q(a)$ — проекция a на Q $\Leftrightarrow$ $\langle a - Pr_Q(a), x - Pr_Q(a)\rangle \leqslant 0.$

[1)] Пособие В.Г. Жадана

$P_1(x_*) = x_* = Pr_Q(x_* - g(x_*))$ тогда и только тогда, когда

$$\langle g(x_*), x_* - x\rangle = \langle g(x_*) + Pr_Q(x_* - g(x_*)) - x_*, Pr_Q(x_* - g(x_*)) - x\rangle =$$

$$= \langle x_* - g(x_*) - Pr_Q(x_* - g(x_*)), x - Pr_Q(x_* - g(x_*))\rangle \leqslant 0,$$

т. е. x_* — решение ВН.

Теорема 1. *Если Q — замкнутое выпуклое множество, то x_* есть решение ВН $\langle g(x_*), x_* - x\rangle \leqslant 0$ тогда и только тогда, когда $Pr_Q(x_* - \alpha g(x_*)) = x_*$ для всякого $\alpha > 0$.*

Следствие 1. *Если Q — выпуклый компакт, то для непрерывного $g : Q \to \mathbb{R}^n$ вариационное неравенство $\langle g(x_*), x_* - x\rangle \leqslant 0$ имеет решения.*

3. Взаимосвязь сильного и слабого решений [2)]

Какова взаимосвязь сильного и слабого решений ВН?

Необходимость получить сильное решение (или близкое к нему в каком-то смысле) возникает из приложений. А для слабых решений больше известно теоретических результатов.

Теорема 2. *Пусть Q — выпуклое замкнутое подмножество $\mathbb{R}^n$, а $g: Q \to \mathbb{R}^n$ есть непрерывное и монотонное отображение. Тогда точка $x_* \in Q$ есть сильное решение тогда и только тогда, когда она есть слабое решение.*

Необходимость.

$$\langle g(x_*), x_* - x\rangle \leqslant 0 \Rightarrow \langle g(x), x_* - x\rangle \leqslant 0 \quad \forall\, x \in Q. \tag{9}$$

Достаточность.

Пусть $x_*: \ \langle g(x), x_* - x\rangle \leqslant 0$ для всякого $x \in Q$.

Тогда для некоторого $x \in Q$, $x \neq x_*$ введём $x_\lambda := \lambda x_* + (1-\lambda)x$, $0 \leqslant \lambda \leqslant 1$. При $\lambda \to 0$ $x_\lambda \to x$; при $\lambda \to 1$ $x_\lambda \to x_*$:

$$x_\lambda - x_* = (1-\lambda)(x - x_*) \tag{10}$$

[2)] Пособие В.Г. Жадана

и тогда

$$\langle g(x_\lambda), x - x_* \rangle = \frac{1}{1-\lambda} \langle g(x_\lambda), x_\lambda - x_* \rangle \geqslant 0 \quad \forall\, \lambda \in [0;1) \Rightarrow$$

$$\Rightarrow \langle g(x_\lambda), x_* - x \rangle \leqslant 0 \quad \forall\, \lambda \in [0;1). \tag{11}$$

Переходим к пределу при $\lambda \to 1$ и тогда $x_\lambda \to x_*$ означает, что $\langle g(x_*), x_* - x \rangle \leqslant 0$ при всяком $x \in Q$.

Теорема 3. *Пусть Q — выпуклое замкнутое множество $\mathbb{R}^n$, а $g:\ Q \to \mathbb{R}^n$ есть непрерывный монотонный оператор. Тогда множество решений ВН выпукло.*

Пусть x_* и y_* — два разных решения.
Тогда $\langle g(x), x_* - x \rangle \leqslant 0$ и $\langle g(x), y_* - x \rangle \leqslant 0$,

$$\langle g(x), \lambda x_* + (1-\lambda) y_* - x \rangle \leqslant 0. \tag{12}$$

Утверждение 1. *Если оператор g строго монотонен, т. е.*

$$\langle g(y) - g(x), y - x \rangle > 0 \quad \forall\, x, y \in Q, \quad x \neq y, \tag{13}$$

то решение единственно.

$$\langle g(x_*), x_* - y_* \rangle \leqslant 0,$$
$$\langle g(y_*), y_* - x_* \rangle \leqslant 0,$$
$$\langle g(x_*) - g(y_*), x_* - y_* \rangle \leqslant 0,$$

что противоречит свойству монотонности оператора g.

Как измерять сходимость и качество решения?

- По аргументу $\|x_k - x^*\|_2^2$ — всё аналогично выпуклой оптимизации, самый надежный критерий.
- По функции. В задачах выпуклой минимизации часто измеряют по $f(x_k) - f(x_*)$. Попробуем сконструировать что-то аналогичное для ВН и седловых задач.
- Попробуем: $f(u_k, v_k) - f(u_*, v_*)$.Насколько хорош такой критерий?

Рассмотрим простую седловую задачу $\min\limits_u \max\limits_v (u-1)\cdot(v+1)$. Решение этой задачи $u = 1,\ v = -1,\ f(1, -1) = 0$.
Пусть начальная точка $u_0 = 0,\ v_0 = 0,\ f(0, 0) = -1$. Тогда $f(u_0, y_0) - f(u_*, v_*)$ отрицательно. Такой критерий не подходит.

- Другой вариант $f(u_k, v_*) - f(u_*, v_k)$.

Рассмотрим простую седловую задачу $\min\limits_u \max\limits_v (u-1)\cdot(v+1)$. Решение этой задачи $u = 1,\ v = -1,\ f(1, -1) = 0$. Тогда $f(u, v_*) - f(u_*, v) = 0$. Такой критерий не подходит для выпукло-вогнутых седел, но подходит для сильно-выпукло-сильно-вогнутых.

- Лучший вариант: $\max\limits_v f(u_k,v) - \min\limits_u f(u,v_k)$.

$$\max_v f(u_k,v) \geqslant f(u_k,v_*) \geqslant f(u_*,v_*), \tag{14}$$

$$\min_u f(u,v_k) \leqslant f(u_*,v_k) \leqslant f(u_*,v_*). \tag{15}$$

Тогда

$$\max_v f(u_k,v) - \min_u f(u,v_k) \geqslant f(u_k,v_*) - f(u_*,v_k) \geqslant 0. \tag{16}$$

Если

$$\min_{u\in Q_u}\max_{v\in Q_v}(u-1)\cdot(v+1), \quad Q_u=\mathbb{R},\ Q_v=\mathbb{R}, \tag{17}$$

то

$$\max_v f(u_k,v) - \min_u f(u,v_k) = +\infty.$$

Поэтому вводят еще одно предположение ограниченности множества: Q — ограниченное, т. е. для любых $x,\ x'\in Q$

$$\|x-x'\|_2 \leqslant D_x. \tag{18}$$

Такое предположение нам понадобится для монотонных неравенств и выпукло-вогнутых седел. В сильно монотонном случае от него можно отказаться, т.к. возможно гарантировать обычную сходимость по аргументу.

Для ВН в качестве критерия можно воспользоваться $\max\limits_{x\in Q}\langle g(x_k), x_k - x\rangle$ (смысл более менее понятен: в случае если x_* — решение, имеем $\max\limits_{z\in Q}\langle g(x_*), x_* - x\rangle \leqslant 0$.

- В теоретическом анализе часто пользуются критерием: $\max\limits_{x\in Q}\langle g(x), x_k - x\rangle$.

4. Проекционный и экстраградиентный метод: сравнение

$$x_{k+1} = Pr_Q(x_k - h_k g(x_k)). \tag{19}$$

Пример 1.

$$f(x,y) = xy \quad \min_x \max_y f(x,y) \text{ — ?}$$

$$(x,y) \in \mathbb{R}^2, \quad (x_*, y_*) = (0,0) \text{ — седловая точка}$$

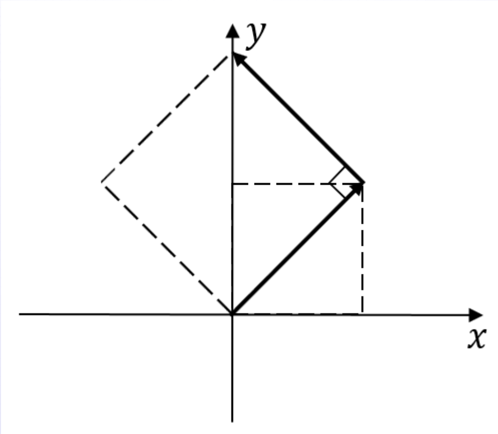

Рис. 2

$$g(x,y) = (f'_x, -f'_y) = (y, -x) \tag{20}$$

Korpelevich, G., M. Extragradient method for finding saddle points and other problems, Economics and Math. methods 12 (4), 747–756 (1976).

$$y_{k+1} = Pr_Q(x_k - h_k g(x_k)), \quad x_{k+1} = Pr_Q(x_k - h_k g(y_{k+1})). \tag{21}$$

Applications: game theory, problems of transport equilibrium.

Works in the field of methods for solving VI

1. F. Facchinei, J. S. Pang, Finite-Dimensional Variational Inequality and Complementar Problems, vols. 1 and 2, // Springer-Verlag, New York, 2003.
2. Beznosikov A., Polyak B., Gorbunov E., Kovalev D., Gasnikov A. Smooth Monotone Stochastic Variational Inequalities and Saddle Point Problems – Survey.
3. 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, 15, pp. 229 – 251.

4. Yu. E. Nesterov. Algorithmic convex optimization. // Diss. ... Doc. Phys.-Math. Sciences: 01.01.07. — Moscow, MIPT: 2013. — 367 p.

Lower bounds for variational inequalities and saddle point problems:

Bilinear saddle problems: $O\left(\frac{1}{\varepsilon}\right)$

Nemirovsky, A. S. (1992). Information-based complexity of linear operator equations. Journal of Complexity, 8(2). P. 153–175.

Численное сравнение проекционного и экстраградиентного методов

In *Stonyakin, F., Gasnikov, A., Dvurechensky, P., Titov, A., Alkousa, M.: Generalized Mirror Prox Algorithm for Monotone Variational Inequalities: Universality and Inexact Oracle. J Optim Theory Appl 194, 988–1013 (2022)* also it was studied the GMP with inexactness.

Let us consider the variational inequality with Lipschitz-continuous strongly monotone operator

$$g : Q \subset \mathbb{R}^n \to \mathbb{R}^n, \quad g(x) = x. \tag{22}$$

We compare the work of GMP with Modified Projection Method, which was proposed in *Khanh P.D., Vuong P.T.: Modified projection method for strongly pseudomonotone variational inequalities. J. Glob. Optim. 58(2), 341–350 (2014).*

We run Algorithms with different values of the accuracy $\varepsilon \in \{10^{-i}, i = 3, 4, \ldots, 10\}$ and for the dimension $n = 10^7$. We take $Q = \{x \in \mathbb{R}^n, \|x\|_2 \leqslant 2\}$. The results of the comparison illustrate the norm $\|x_{\text{out}} - x_*\|_2$, as a function of iterations, where x_{out} is the output of each algorithm. Note that $x_* = \mathbf{0} \in \mathbb{R}^n$. In the conducted experiments, at the first, we run Algorithm GMP, and calculate $\|x_{\text{out}} - x_*\|_2$ for the different previously mentioned values of ε and the corresponding number of iterations, resulted by the working of algorithm. For the calculated number

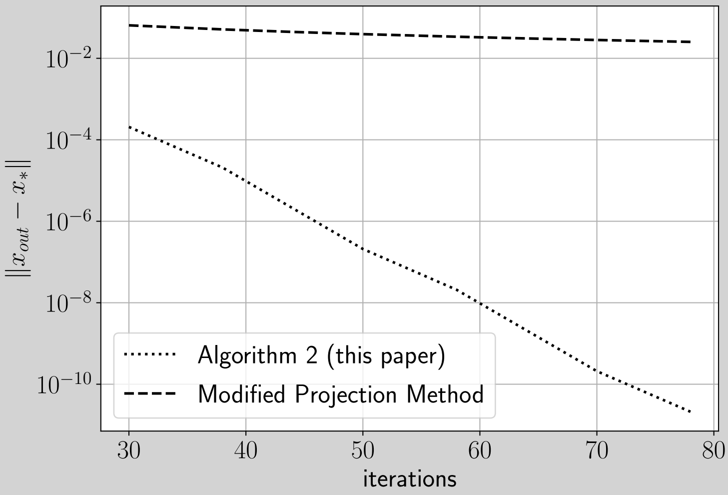

Рис. 3. Results of comparison with $n = 10^7$.

of iterations of GMP, we rum Modified Projection Method and calculate the corresponding values $\|x_{\text{out}} - x_*\|_2$.

5. Проекционный метод: теоретический результат о сходимости

Предположим, что имеется вариационное неравенство, в котором Q — выпуклое замкнутое подмножество $\mathbb{R}^n$, а $g(x)$ — непрерывное на $\mathbb{R}^n$ отображение. Согласно теореме 2.1.5 точка x_* — решение вариационного неравенства в том и только в том случае, когда она является неподвижной точкой проекционного отображения

$$P(x) = Pr_Q(x - \alpha g(x)), \ \alpha > 0,$$

т. е. если выполняется равенство

$$x_* = P(x_*) = Pr_Q(x_* - \alpha g(x_*)). \tag{23}$$

Рассмотрим теперь метод решения ВН, в котором используется отображение $P(x)$. Назовем его *проекционным методом*. Пусть задано начальное приближение $x_0 \in Q$. Последующие точки x_k, $k \geqslant 1$, находим с помощью следующего рекуррентного соотношения:

$$x_{k+1} = Pr_Q(x_k - \alpha g(x_k)), \ k = 0, 1, 2, \ldots, \tag{24}$$

где $\alpha > 0$ — параметр.

Теорема 4. *Пусть Q — выпуклое замкнутое подмножество в $\mathbb{R}^n$. Пусть, кроме того, $g(x)$ — сильно монотонное на Q отображение с константой $\mu > 0$, удовлетворяющее условию Липшица, т. е. для всех $x, y \in Q$ выполняется*

$$\langle g(x) - g(y), x - y \rangle \geqslant \mu \left\| x - y \right\|_2^2, \tag{25}$$

$$\left\| g(x) - g(y) \right\|_2 \leqslant L \left\| x - y \right\|_2. \tag{26}$$

Тогда при любом $x_0 \in Q$ и $0 < \alpha < \frac{2\mu}{L^2}$ последовательность $\{x_k\}$, порождаемая итерационным процессом (24), сходится к решению задачи ВН — точке $x_* \in Q$ — с линейной скоростью

$$\left\| x_{k+1} - x_* \right\|_2 \leqslant C \left\| x_k - x_* \right\|_2, \; 0 < C < 1. \tag{27}$$

Доказательство. Воспользуемся тем, что оператор проектирования $P(x) = Pr_Q(x - \alpha g(x))$ является нерастягивающим. Тогда из (23) и (24)

$$\begin{aligned} \left\| x_{k+1} - x_* \right\|_2^2 = \left\| Pr_Q(x_k - \alpha g(x_k)) - Pr_Q(x_* - \alpha g(x_*)) \right\|_2^2 \leqslant \\ \leqslant \left\| x_k - \alpha g(x_k) - x_* + \alpha g(x_*) \right\|_2^2 = \\ = \left\| x_k - x_* \right\|_2^2 + \alpha^2 \left\| g(x_k) - g(x_*) \right\|_2^2 - 2\alpha \langle g(x_k) - g(x_*), x_k - x_* \rangle. \end{aligned} \tag{28}$$

Но согласно (25)

$$\langle g(x_k) - g(x_*), x_k - x_* \rangle \geqslant \mu \left\| x_k - x_* \right\|_2^2, \tag{29}$$

а в силу (26)

$$\|g(x_k) - g(x_*)\|_2^2 \leqslant L^2 \|x_k - x_*\|_2^2 . \tag{30}$$

Поэтому после подстановки (29) и (30) в (28) приходим к

$$\|x_{k+1} - x_*\|_2^2 \leqslant \left(1 + \alpha^2 L^2 - 2\alpha\mu\right) \|x_k - x_*\|_2^2 .$$

При $0 < \alpha < \frac{2\mu}{L^2}$ выполняется неравенство

$$1 + \alpha^2 L^2 - 2\alpha\mu < 1.$$

Отсюда заключаем, что итерации, определяемые формулой (24), сходятся с линейной скоростью к x_*, причем в (27) константа C равняется

$$C = \left(1 + \alpha^2 L^2 - 2\alpha\mu\right)^{1/2} .$$

$\square$

СПАСИБО ЗА ВНИМАНИЕ!