

# Универсальные методы для стохастических вариационных неравенств

Климза Антон Алексеевич

Московский физико-технический институт

*Курс:* Научный трек Иннпрака ФПМИ

*Эксперт:* д.ф.-м.н. А. В. Гасников

17 мая 2024

## Вариационное неравенство

Пусть дано выпуклое множество  $Z \in \mathbb{R}^n$  и оператор  $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Тогда хотим найти  $z^* \in Z$ , такую что:

$$\langle g(z^*), z - z^* \rangle \geq 0, \forall z \in Z.$$

## Источники

[1] Anton Rodomanov Ali Kavis Yongtao Wu Kimon Antonakopoulos Volkan Cevher. Universal Gradient Methods for Stochastic Convex Optimization. 2024.

[2] Fedor Stonyakin Alexander Gasnikov Pavel Dvurechensky Alexander Titov Mohammad Alkousa. Generalized Mirror Prox Algorithm for Monotone Variational Inequalities: Universality and Inexact Oracle. 2022.

# Постановка задачи

## Вариационное неравенство

Пусть дано выпуклое множество  $Z \in \mathbb{R}^n$  и оператор  $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Тогда хотим найти  $z^* \in Z$ , такую что:

$$\langle g(z^*), z - z^* \rangle \geq 0, \quad \forall z \in Z.$$

## Стохастический случай

$$g(z) = \mathbb{E}_{\xi} g(z, \xi),$$

$$\mathbb{E}_{\xi} \|g(z) - g(z, \xi)\|^2 \leq \sigma^2.$$

## Условие Гёльдера

$$\exists \nu \in [0, 1], L_{\nu} \geq 0 : \|g(x) - g(y)\|_* \leq L_{\nu} \|x - y\|^{\nu} \quad \forall x, y \in Z.$$

## Проксимальный зеркальный метод

$$w_k = \arg \min_{x \in Q} \left( \langle g(z_k), x - z_k \rangle + L_k \frac{1}{2} \|z_k - x\|^2 \right),$$

$$z_{k+1} = \arg \min_{x \in Q} \left( \langle g(w_k), x - w_k \rangle + L_k \frac{1}{2} \|z_k - x\|^2 \right).$$

Где  $L_k$  – константа Липшица.

Идея – изменять  $L_k$  на каждом шаге, подстраиваясь под гладкость задачи.

Формула пересчёта  $L_{k+1}$

$$(L_{k+1} - L_k) \frac{D^2}{2} = \left| -\langle g(w_k), z_{k+1} - w_k \rangle - L_{k+1} \frac{1}{2} \|z_k - z_{k+1}\|^2 \right|_+.$$

Оценка сходимости

$$-\frac{1}{k} \sum_{i=0}^k \langle g(w_k), z^* - w_k \rangle \leq \frac{2D^2 L_{k+1}}{k},$$

$$L_{k+1} \leq \frac{3-\nu}{2} \left( \frac{k}{D^2} \right)^{\frac{1-\nu}{2}} L_\nu.$$

Где  $L_\nu$  – константа Гёльдера,  $D$  – диаметр рассматриваемого множества решений  $Z$ .

---

**Algorithm 1** Универсальный проксимальный зеркальный метод (UMP)

---

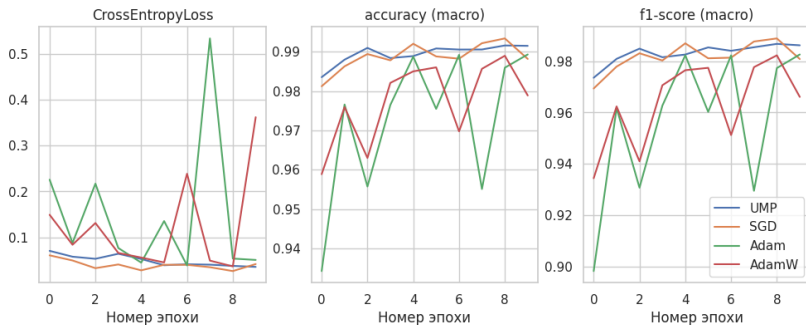
- 1: Set  $z_0 = \arg \min_{u \in Q} d(u)$ ,  $L_0 = \|g(z_0)\|$ .
  - 2: **for**  $k = 0, 1, \dots$  **do**
  - 3:    $w_k = \arg \min_{x \in Q} (\langle g(z_k), x \rangle + L_k \frac{1}{2} \|z_k - x\|^2)$ .
  - 4:    $z_{k+1} = \arg \min_{x \in Q} (\langle g(w_k), x \rangle + L_k \frac{1}{2} \|z_k - x\|^2)$ .
  - 5:    $L_{k+1} = L_k + \max \left( 0, \frac{-2\langle g(w_k), z_{k+1} - w_k \rangle - L_k \|w_k - z_{k+1}\|^2}{D^2 + \|w_k - z_{k+1}\|^2} \right)$ .
  - 6: **end for**
- 

## Оценка скорости сходимости метода

$$O \left( \inf_{\nu \in (0,1)} \left( \frac{(3-\nu)L_\nu}{\varepsilon} \right)^{\frac{2}{1+\nu}} D^2 \right).$$

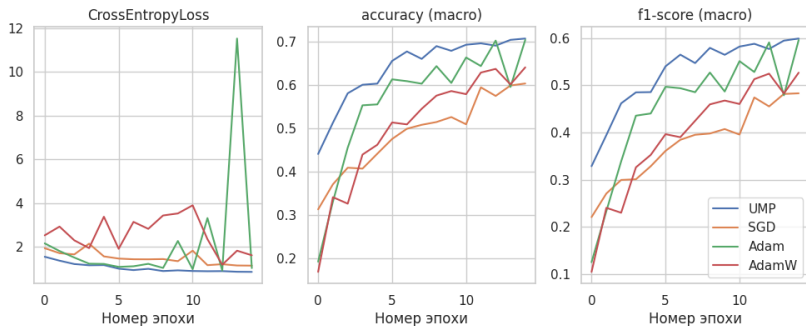
# Сравнение методов

*Сравнение метрик оптимизаторов на валидации при обучении resnet50 датасете MNIST*



# Сравнение методов

*Сравнение метрик оптимизаторов на валидации при обучении resnet50 датасете CIFAR10*





## Результат

- ▶ Получен универсальный проксимальный зеркальный метод и доказана оценка его скорости сходимости.
- ▶ Проведён эксперимент, который показывает, что метод может быть применён для задачи минимизации функции наравне с популярными оптимизаторами.

## План на будущее

- ▶ Сравнение нашего метода с другими методами на задаче обучения генеративно-состязательных сетей.