

## Лабораторная работа #2.

### Стохастический градиентный спуск.

Рассмотрим задачу оптимизации

$$\min \left\{ (\mathbf{x} - \boldsymbol{\mu}_0)^\top \mathbf{A}(\mathbf{x} - \boldsymbol{\mu}_0) : \|\mathbf{x}\|_2^2 \leq 1 \right\}, \quad (1)$$

где  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{A}$  - симметричная, положительноопределенная матрица,  $\boldsymbol{\mu}_0 = (1, 1, \dots, 1)^\top \in \mathbb{R}^n$ .

1. Запишите двойственную задачу к (1) и примените метод из Лабораторной работы №1 к двойственной задаче.
2. Для метода из Лабораторной работы №1 замените градиент на стохастический градиент<sup>1</sup>

$$\tilde{\nabla} f(x) = \frac{1}{m} \sum_{j=1}^m (0, 0, \dots, n f'_{x_i}, \dots, 0)^\top$$

и примените метод стохастического градиентного спуска для различных значений параметра  $m \in \{1, \frac{n}{8}, \frac{n}{4}, \frac{n}{2}, n\}$ .

3. Представьте следующие результаты:

- Для заданной точности  $\varepsilon > 0$  по значению целевой функции и для каждого значения  $n \in \{10, 20, \dots, 100\}$  и  $m \in \{1, \frac{n}{8}, \frac{n}{4}, \frac{n}{2}, n\}$  подсчитайте среднее число арифметических операций во всех 3 рассмотренных методах

---

<sup>1</sup>Шаг нужно выбирать тем же способом, как и в обычном градиентном спуске.

(усреднение проводится по всем начальным точкам и по всем тестовым примерам). Если число операций подсчитать не удастся, то укажите среднее время работы метода. Какой метод работает лучше и почему?

- Для одного тестового примера при  $n = 10$  и нескольких различных начальных точек постройте зависимость точности от числа арифметических операций в методах градиентного и стох. градиентного спуска, примененных для задачи (1). Если число операций подсчитать не получается, то постройте зависимость средней точности от времени работы метода.

4. Оформите отчет с последовательным изложением пунктов 1-3 и выводами.