

## Лабораторная работа #2 (часть 1).

### Градиентный спуск.

1. 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,  $\boldsymbol{\mu}_0 = (1, 1, \dots, 1)^\top \in \mathbb{R}^n$ .

2. Исследуйте задачу (1) на выпуклость. Запишите необходимые условия минимума (находить минимум аналитически не требуется).
3. Для каждого значения  $n \in \{10, 20, \dots, 100\}$  сгенерируйте  $N = 100$  тестовых примеров. В каждом случае найдите глобальный минимум,  $\mathbf{x}^* \in \mathbb{R}^n$ , с помощью CVX. Проверьте, что в точке минимума выполняется условие оптимальности (т.е. вектора градиента к ограничению и антиградиента к целевой функции сонаправлены). Кроме того, для каждого тестового примера сгенерируйте случайно 100 начальных точек из единичного шара.
4. В зависимости от варианта реализуйте следующие методы (для заданной точности по норме градиента  $\varepsilon = 0.01$ ):

- Gradient descent for strongly convex and Lipschitz functions;

(Section 3.4.1, Theorem 3.9, <https://arxiv.org/pdf/1405.4980.pdf>)

- Gradient descent for smooth functions;  
(Section 3.2, Theorem 3.7, <https://arxiv.org/pdf/1405.4980.pdf>)
- Gradient descent for strongly convex and smooth functions;  
(Section 3.4.2, Theorem 3.12, <https://arxiv.org/pdf/1405.4980.pdf>)
- Conditional gradient descent, aka Frank-Wolfe;  
(Section 3.3, Theorem 3.8, <https://arxiv.org/pdf/1405.4980.pdf>)
- Exact line search;  
(Section 9.2, [https://web.stanford.edu/boyd/cvxbook/bv\\_cvxbook.pdf](https://web.stanford.edu/boyd/cvxbook/bv_cvxbook.pdf))
- Backtracking line search;  
(Section 9.2, [https://web.stanford.edu/boyd/cvxbook/bv\\_cvxbook.pdf](https://web.stanford.edu/boyd/cvxbook/bv_cvxbook.pdf))

Для вариантов 5 и 6 при использовании методов дополнительно необходимо проецировать найденные точки на допустимое множество в задаче (1). В каждом методе, если это необходимо, оцените константу Липшица градиента, константу сильной выпуклости и другие параметры.

5. Объясните принцип работы метода, опишите его преимущества и недостатки. Сколько вычислений требуется на каждой итерации?
6. В качестве результата работы метода подсчитайте:
  - Для каждого значения  $n \in \{10, 20, \dots, 100\}$  среднее время работы метода и среднее число итераций (усреднение проводится по всем начальным точкам и по всем тестовым примерам);

- Для отдельного тестового примера и  $n = 10$  постройте зависимость средней (по начальным точкам) точности от числа итераций. Сравните полученные результаты с теоретическими верхними оценками (только для вариантов 1-4). Как зависит качество решений от отношения максимального и минимального собственных чисел матрицы  $A$ ? *(Точность по значению функции может измеряться разными способами в зависимости от метода, см. верхние оценки для каждого метода)*
7. Оформите отчет с последовательным изложением пунктов 1-6 и выводами.