

Лабораторная работа #1

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

1. Исследуйте задачу (1) на выпуклость.
2. 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).
3. Для каждого значения $n \in \{10, 20, \dots, 100\}$ и для каждого тестового примера сгенерируйте 100 начальных точек. В зависимости от варианта реализуйте следующие методы решения задачи (1) для заданной точности $\varepsilon = 0.01$ ¹:

- Gradient descent for strongly convex and Lipschitz functions;

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

¹Используйте в качестве точности либо разность между текущим значением функции и оптимальным (из солвера), либо некоторую меру выполнения условий оптимальности

- 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)
- Backtracking line search;
(Section 9.2, https://web.stanford.edu/boyd/cvxbook/bv_cvxbook.pdf)

4. Объясните принцип работы метода, опишите его преимущества и недостатки.

5. В качестве результата работы метода:

- Для каждого значения $n \in \{10, 20, \dots, 100\}$ подсчитайте среднее время работы метода и среднее число итераций (усреднение проводится по всем начальным точкам и по всем тестовым примерам);
- Для одного тестового примера при $n = 10$ и нескольких различных начальных точек постройте зависимость точности от числа итераций. Зависит ли скорость сходимости метода от отношения максимального и минимального собственных чисел матрицы A ? Сравните полученные

результаты с теоретическими верхними оценками (только для вариантов 1-4)².

6. Оформите отчет с последовательным изложением пунктов 1-5 и выводами.

²Обратите внимание, что для сравнения с теоретическими верхними оценками необходимо считать точность тем же способом, как это делается в соответствующей оценке