Санкт–Петербургский государственный университет

Математическое обеспечение и адмиистрирование информационных систем

Гусев Егор Игоревич

Вычислительный практикум

Отчет по заданию №14

Преподователь:
Т.О. Евдокимова

Санкт-Петербург
2021 г.

Содержание

1. Ссылка на код

Код доступен по ссылке на github.

2. Постановка задачи

Необходимо реализовать алгоритм градиентного спуска и метод Нестерова решения задачи выпуклой оптимизации.

3. Теория

Рассматривается f : H $\to R$, где H — Гильбертово, f — сильно выпуклая, обладающая свойством $\forall x, y \in H \exists L > 0 : \|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|$ Необходимо найти точку минимума заданной функции.

3.1 Градиентный спуск

1. Выбирается начальное приближение $\theta_0 \in H$, коэффициент скорости обучения λ и параметр ϵ
2. $\theta_{n+1} = \theta_n - \lambda \nabla f(\theta_n)$
3. Шаг 2 повторяется, пока $\|\theta_{n+1} \theta_n\| > \epsilon$ или до достижения максимального числа итераций

3.2 Метод Нестерова

1. Выбирается начальное приближение $y_0 \in H, z \in H : z \neq y_0, \nabla f(y_0) \neq \nabla f(z)$
2. Полагаем $a_0 = 1, x_1 = y_0, a_1 = \frac{\|y_0 - z\|}{\|\nabla f(y_0) - \nabla f(z)\|}$
3. На k-й итерации метода:

   Выбирается наименьший индекс i:

$$f(y_k) - f(y_k - 2^{-i} a_{k-1} \nabla f(y_k)) \ge 2^{-1-i} a_{k-1} \|\nabla f(y_k)\|^2$$

Полагаем:

$$a_k = 2^{-i} a_{k-1}$$

$$x_k = y_k a_k \nabla f(y_k)$$

$$a_{k+1} = \frac{1 + \sqrt{4a_k^2 + 1}}{2}$$

$$y_{k+1} = x_k + \frac{(a_k 1)(x_k x_{k1})}{a_{k+1}}$$

4. Шаг 3 повторяется, пока $\|y_{k+1} y_k\| > \epsilon$ или до достижения максимального числа итераций.

4. Численный эксперимент

4.1 Описание

Рассмотрм функцию g(x) = 4x + 11. Возьмем N = 100 равноотстоящих точек $x_1, \dots x_{100}$ на отрезке [10; 10] и положим $y_1 = g(x_1), \dots y_{100} = g(x_{100})$.

Рассмотрим квадратичную функцию потерь

$$L(a, b) = \frac{1}{2} \sum_{n=1}^{N} (ax_i + b - y_i)^2$$

$$L(a, b), ..N = 100g(x).$$

4.2 Результаты

В таблицах представлены результаты работы алгоритма градиентного спуска и метода Нестерова для данной функции. Таблица содержит сведения о числе итераций, потребовавшихся для методов, о значениях параметров λ, ϵ, точности приближения коэффициентов функции g(x) и значении функции потерь в получившейся точке минимума.

	lambda	eps	iters	|a_eps-a|	|b_eps-b|	loss
0	0.010000	0.010000	6.0	4.380675e+09	4.103879e+08	3.198069e+22
1	0.010000	0.000100	7.0	5.567618e+09	8.388616e+11	3.518941e+25
2	0.010000	0.000001	6.0	3.580076e+09	9.843211e+03	2.136584e+22
3	0.000100	0.010000	250.0	3.071887e-03	9.934416e-01	4.933152e+01
4	0.000100	0.000100	698.0	3.075252e-05	9.929292e-03	4.928065e-03
5	0.000100	0.000001	1162.0	3.571931e-07	9.934884e-05	4.933674e-07
6	0.000001	0.010000	0.0	2.934287e+00	1.203571e+01	2.124269e+04
7	0.000001	0.000100	2000.0	3.819711e-02	1.032261e+01	5.326302e+03
8	0.000001	0.000001	2000.0	3.880933e-02	1.043288e+01	5.440710e+03

Рис. 1: Градиентный спуск

	eps	iters	|a_eps-a|	|b_eps-b|	loss
0	1.000000e-02	0.0	2.189094e+00	10.030417	1.279937e+04
1	1.000000e-04	490.0	4.150901e-04	0.134225	9.005510e-01
2	1.000000e-06	678.0	1.147943e-06	0.000355	6.302209e-06
3	1.000000e-08	1418.0	6.134034e-08	0.000004	6.771544e-10

Рис. 2: Метод Нестерова