# Домашняя работа №6
## Студент: Павел Хоменко

**Задача 1 (2 балла).** Бывший технический директор Open AI Мира выписала на листочке функцию $f(x) = x \cdot \ln x$ и просит вас найти её минимум. Помогите ей.

Решение

$$f'(x) = (x \ln x)' = x' \ln x + x \cdot \ln x' = \ln x + 1$$

$$f'(x) = 0 : \quad \ln x + 1 = 0 \to \ln x = -1 \to e^{\ln x} = e^{-1} \to x_0 = \frac{1}{e}$$

$$f''(x) = (\ln x + 1)' = \frac{1}{x} \qquad f''\left(\frac{1}{e}\right) = \frac{1}{e^{-1}} = e > 0 \Rightarrow$$

$x_0$ – точка минимума

$$f(x_0) = \frac{1}{e} \ln \frac{1}{e} = -\frac{1}{e}$$

Ответ: $\left(\frac{1}{e}; -\frac{1}{e}\right)$

**Задача 2 (2 балла).** Бывший президент гугла Сергей подзабыл матанализ. Исследуйте за него на экстремум функцию

$$f(x, y) = y^2 + 2xy - 4x - 2y - 3.$$

Решение:

$$\frac{\partial f}{\partial x} = 2y - 4 \qquad \frac{\partial f}{\partial y} = 2y + 2x - 2$$

$$\nabla_{x,y} f = \begin{pmatrix} 2y-4 \\ 2y+2x-2 \end{pmatrix}$$

$$\nabla_{x,y} f = 0: \begin{cases} 2y-4=0 \\ 2y+2x-2=0 \end{cases} \to \begin{cases} y=2 \\ 4+2x-2=0 \end{cases} \to \begin{cases} y=2 \\ x=-1 \end{cases}$$

$$\frac{\partial^2 f}{\partial x^2} = 0 \quad \frac{\partial^2 f}{\partial x \partial y} = 2 \quad \frac{\partial^2 f}{\partial y \partial x} = 2 \quad \frac{\partial^2 f}{\partial y^2} = 2$$

$$H = \begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix} \qquad \Delta_1 = 0 \qquad \Delta_2 = -4$$

Точка $(-1;2)$ – седловая

**Задача 3 (2 балла).** Бывший директор Яндекса Аркадий выписал на листочек функцию $f(x,y) = (1+2x+3y)^{2023}$ и просит вас найти её разложение в ряд Тэйлора до второго члена в окрестности точки $x=0, y=0$. Сделайте это.

Решение:

$$f(x,y) = f(x_0,y_0) + \frac{1}{1!}\left(\frac{\partial f(x_0,y_0)}{\partial x}(x-x_0) + \frac{\partial f(x_0,y_0)}{\partial y}(y-y_0)\right) +$$

$$+ \frac{1}{2!}\left(\frac{\partial^2 f(x_0,y_0)}{\partial x^2}(x-x_0)^2 + 2\frac{\partial^2 f(x_0,y_0)}{\partial x \partial y}(x-x_0)(y-y_0) + \frac{\partial^2 f(x_0,y_0)}{\partial y^2}(y-y_0)^2\right) +$$

$$+ o(\rho^2)$$

$$\frac{\partial f}{\partial x} = 2023(1+2x+3y)^{2022}\cdot 2 \qquad \frac{\partial f(x_0, y_0)}{\partial x} = 4046$$

$$\frac{\partial f}{\partial y} = 2023(1+2x+3y)^{2022}\cdot 3 \qquad \frac{\partial f(x_0, y_0)}{\partial y} = 6069$$

$$\frac{\partial^2 f}{\partial^2 x} = 2023\cdot 2022\,(1+2x+3y)^{2021}\cdot 2\cdot 2 \qquad \frac{\partial^2 f(x_0, y_0)}{\partial^2 x} = 16\,362\,024$$

$$\frac{\partial^2 f}{\partial x \partial y} = 2023\cdot 2022\cdot 2\cdot 3\cdot (1+2x+3y)^{2021} \qquad \frac{\partial^2 f(x_0, y_0)}{\partial x \partial y} = 24\,543\,036$$

$$\frac{\partial^2 f}{\partial^2 y} = 2023\cdot 2022\cdot 3\cdot 3\,(1+2x+3y)^{2021} \qquad \frac{\partial^2 f(x_0, y_0)}{\partial^2 y} = 36\,814\,554$$

$$f(x_0 y_0) = (1+2\cdot 0+3\cdot 0)^{2023} = 1^{2023} = 1$$

$$f(x,y) = 1 + 4046x + 6069y + 8181012x^2 + 24543036xy +$$
$$+ 18407277y^2 + o(\rho^2)$$

**Задача 4 (2 балла).** Бывший директор отдела по искусственному интеллекту Теслы Андрей ищет минимум функции $f(x) = ax^2 + bx + c$, где $a > 0$, методом градиентного спуска. Андрей стартует из точки $x_0$ и настолько ленив, что не хочет делать больше одного шага. При каком значении длины шага $\gamma$ Андрей за один шаг окажется точно в точке минимума?

Решение:

$$f'(x) = 2ax + b$$

$$f'(x) = 0: \quad x_{min} = -\frac{b}{2a}$$

$$x_{min} = x_0 - \eta \cdot \nabla f(x_0) \qquad \eta = \gamma$$

$$-\frac{b}{2a} = x_0 - \gamma(2ax_0 + b)$$

$$\gamma(2ax_0 + b) = x_0 + \frac{b}{2a}$$

$$\gamma(2ax_0 + b) = \frac{2ax_0 + b}{2a} \quad \Big/ : (2ax_0 + b)$$

$$\gamma = \frac{1}{2a}$$

Ответ: размер шага $\frac{1}{2a}$

**Задача 5 (2 балла).** Бывший директор исследовательского отдела Microsoft Research Asia, Ли не доверяет пакетным реализациям алгоритмов машинного обучения. Поэтому он написала свой собственный градиентный спуск. Для того, чтобы делать шаг градиентного спуска, он использовал следующие формулы.

$$w_t = w_{t-1} + (\nabla Q(w_t))^2$$



Какие ошибки вы тут видите? Для каждой объясните, к каким последствиям и почему она приведёт, а также как это исправить.

Решение:

1) Не задана скорость обучения $\eta$ → шаги могут быть либо слишком большими, либо слишком маленькими → $w_t = w_{t-1} + \eta (\nabla Q(w_t))^2$

2) Знак перед градиентом „+" → минимум не будет найден. Нужно двигаться в направлении, противоположном градиенту, поскольку градиент – направление макс. роста функции

→ $w_t = w_{t-1} - \eta \cdot (\nabla Q(w_t))^2$

Нахождение градиента в точке $w_t$ и использование квадрата градиента не является препятствием для нахождения минимума, однако может затруднить сходимость и ухудшить стабильность поиска

Исправленная формула: $w_t = w_{t-1} - \eta \cdot \nabla Q(w_{t-1})$

**Задача 6 (1 балл).** Запишите фамилии Миры, Сергея, Аркадия, Андрея и Ли.

Мира Мурати

Сергей Брин

Аркадий Волож

Андрей Карпатый

Ли Кайфу