

# Тятя! Тятя! Наши сети притащили мертвеца!

Листочек с задачками №2: 50 оттенков градиентного спуска

[https://github.com/FUlyankin/neural\\_nets\\_prob](https://github.com/FUlyankin/neural_nets_prob)

РАНХ  
осень 2020

Производная это просто  
Скорость роста, это скорость роста.  
Возьми предел  $\frac{\Delta y}{\Delta x}$  и получишь.  
Чем выше она — тем круче.  
Научно-технический рэп

## Упражнение 1 (регрессия)

Маша Нестерова, хозяйка машин лёрнинга<sup>1</sup>, собрала два наблюдения:  $x_1 = 1, x_2 = 2, y_1 = 2, y_2 = 3$  и собирается обучить линейную регрессию  $y = \beta \cdot x$ . Маша очень хрупкая девушка, и ей не помешает помощь.

- Получите теоретическую оценку методом наименьших квадратов.
- Сделайте два шага градиентного спуска. В качестве стартовой точки используйте  $\beta_0 = 0$ . В качестве скорости обучения возьмите  $\eta = 0.1$ .
- Сделайте два шага стохастического градиентного спуска. Пусть в SGD сначала попадает первое наблюдение, затем второе.
- Если вы добрались до этого пункта, вы поняли градиентный спуск. Маша довольна. Начинаем заниматься тупой технической бессмыслицей. Сделайте два шага Momentum SGD. Возьмите  $\alpha = 0.9, \eta = 0.1$
- Сделайте два шага Momentum SGD с коррекцией Нестерова.
- Сделайте два шага RMSprop. Возьмите  $\alpha = 0.9, \eta = 0.1$
- Шоб ещё такого сделать? Придумал! Давайте сделаем два шага Adam. Возьмём  $\beta_1 = \beta_2 = 0.9, \eta = 0.1$

---

<sup>1</sup>Лёрнинг ей папа подарил

## Упражнение 2 (логрегрессия)

Маша решила, что нет смысла останавливаться на обычной регрессии, когда она знает, что есть ещё и логистическая:

$$z = \beta \cdot x \quad p = P(y = 1) = \frac{1}{1 + e^{-z}}$$
$$\text{logloss} = -[y \cdot \ln p + (1 - y) \cdot \ln(1 - p)]$$

Запишите формулу, по которой можно пересчитывать веса в ходе градиентного спуска для логистической регрессии.

Оказалось, что  $x = -5$ , а  $y = 1$ . Сделайте один шаг градиентного спуска, если  $\beta_0 = 1$ , а скорость обучения  $\gamma = 0.01$ .