

Машинное обучение, ФКН ВШЭ

Линейная регрессия

Задача 1. Докажите, что оценка градиента, получаемая методом mini-batch gradient descent, является несмещенной.

Задача 2. Допустим, что обучающая выборка состоит из 2ℓ объектов и верно следующее: $y_i = -y_{\ell+i} \neq 0, i = 1 \dots \ell$. Рассмотрим константное предсказание $a(x) = C$. Найдите оптимальное значение C с точки зрения функционала MAPE.

Задача 3. Рассмотрим две задачи линейной регрессии с L_1 - и L_2 -регуляризацией и одинаковыми коэффициентами $\lambda > 0$:

$$\|Xw - y\|_2^2 + \lambda \|w\|_1 \rightarrow \min_w \quad (1)$$

$$\|Xw - y\|_2^2 + \lambda \|w\|_2 \rightarrow \min_w \quad (2)$$

Пусть их решения равны, соответственно, w_1^* и w_2^* . Можно ли утверждать, что $\|w_1^*\|_1 < \|w_2^*\|_1$? А что $\|w_2^*\|_2 < \|w_1^*\|_2$?

Задача 4. Убедитесь, что вы знаете ответы на следующие вопросы:

- Почему L_1 -регуляризация производит отбор признаков?
- Почему коэффициент регуляризации нельзя подбирать по обучающей выборке?
- Почему накладывать регуляризатор на свободный коэффициент w_0 может быть плохой идеей?
- Что такое кросс-валидация, чем она лучше использования отложенной выборки?
- Почему категориальные признаки нельзя закодировать натуральными числами? Что такое one-hot encoding?
- Для чего нужно масштабировать матрицу объекты-признаки перед обучением моделей машинного обучения?
- Почему MSE чувствительно к выбросам?