

# Задача 4

Гладких Андрей 131

## Постановка задачи

Построить регрессор на плоскости. По выбору осуществляется линейная, полиномиальная или экспоненциальная регрессия.

## Решение

Для линейной регрессии получаем коэффициенты с помощью формул

$$\omega = (A^T A)^{-1} A^T y, \text{ где } A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{pmatrix}, \text{ а } y = (y_1, y_2, \dots, y_n).$$

Для полиномиальной используем библиотеку `sklearn`. Будем увеличивать степень полинома либо пока коэффициент детерминации не будет достаточно высок, либо пока не начнет уменьшаться.

Для экспоненциальной используем библиотеку `scipy`. Используем формулу  $ae^{xk} + b$

### Пример работы и анализ результатов

Точность модели оцениваем с помощью среднеквадратической ошибки

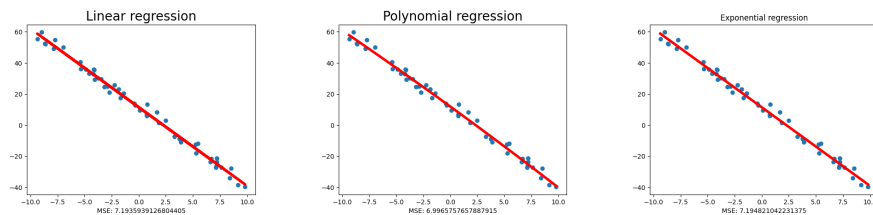


Рис. 1:

$MSE = 7.913/6.996/7.194$ . Результаты отличаются незначительно. Все модели справились с задачей примерно одинаково. Но лучше всего показывает себя полиномиальная.

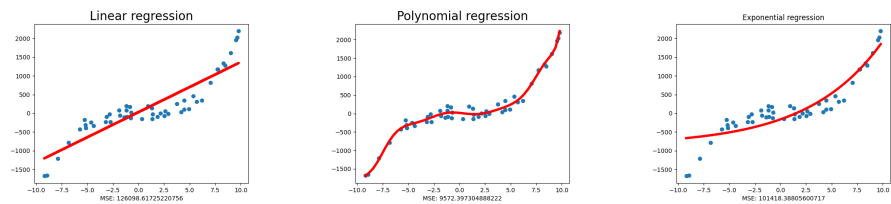


Рис. 2:

$MSE = 126098/9572/101418$ . Очевидное преимущество полиномиальной модели.

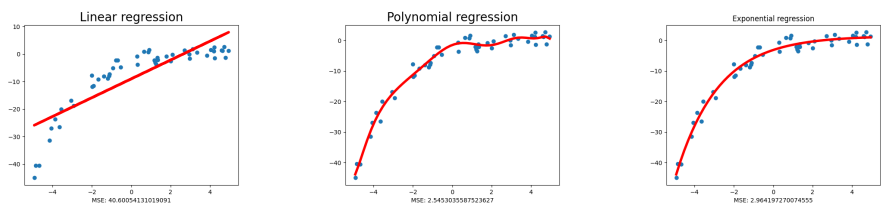


Рис. 3:

$MSE = 40.6/2.54/2.96$ . Линейная не справляется с подобным разбросом данных должным образом. Полиномиальная и экспоненциальная примерно одинаково хорошо.

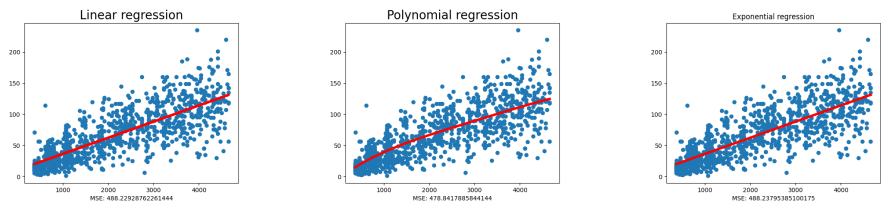


Рис. 4:

$MSE = 488.229/478.841/488.237$ . На подобном наборе данных все модели справились одинаково плохо. Высокий  $MSE$  говорит о невозможности построить вменяемую модель.