

Assignment 1

Devansh Jain, 190100044

29 Aug 2021

Contents

1	Ordinary Least Squares (OLS) Regression in one variable	1
1.1	CS337: Theory	1
1.2	CS335: Lab	1
	(a)	1
	(b)	1
	(c)	1
	(d)	2

1 Ordinary Least Squares (OLS) Regression in one variable

1.1 CS337: Theory

$$\begin{aligned}
 mse(w, b) &= \frac{1}{N} \sum_{i=1}^N ((wx_i + b) - y_i)^2 \\
 \frac{\partial mse(w, b)}{\partial w} &= \frac{1}{N} \sum_{i=1}^N 2 \cdot ((wx_i + b) - y_i) \cdot x_i \\
 \frac{\partial mse(w, b)}{\partial b} &= \frac{1}{N} \sum_{i=1}^N 2 \cdot ((wx_i + b) - y_i) \\
 \nabla mse(w, b) &= \begin{pmatrix} \frac{\partial mse(w, b)}{\partial w} \\ \frac{\partial mse(w, b)}{\partial b} \end{pmatrix} \\
 &= \begin{pmatrix} \frac{1}{N} \sum_{i=1}^N 2 \cdot ((wx_i + b) - y_i) \cdot x_i \\ \frac{1}{N} \sum_{i=1}^N 2 \cdot ((wx_i + b) - y_i) \end{pmatrix} \\
 &= \frac{2}{N} \begin{pmatrix} \sum_{i=1}^N ((wx_i + b) - y_i) \cdot x_i \\ \sum_{i=1}^N ((wx_i + b) - y_i) \end{pmatrix}
 \end{aligned}$$

1.2 CS335: Lab

(a)

Present in notebook

(b)

Present in notebook

(c)

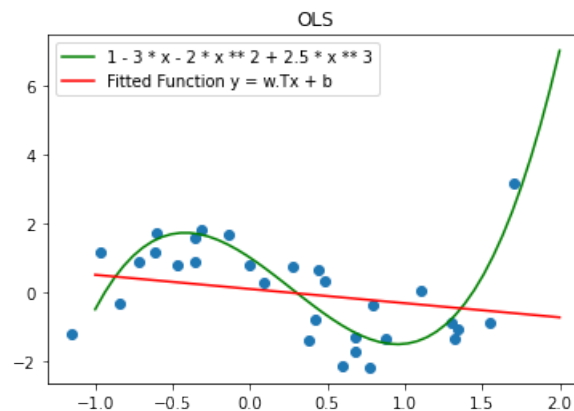


Figure 1: Single Variable - Gradient Descent (epochs=1000)

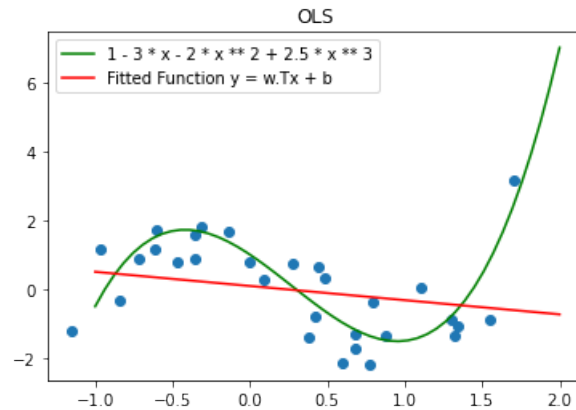


Figure 2: Single Variable - Closed Form

(d)

As mentioned in description of the question (point 5), the error function (mse) is convex and has just one minimum.

By using the closed form, we get the point corresponding to minimum mse error. As we compute the parameters based on training data, the training loss is minimum at this point.

Therefore, training loss for solution of `singlevar_grad()` can never be strictly less than that of the solution obtained by `singlevar_closedform()`.