Exercose 3

Arnas Sniokaitis (mat. nr. 3252523)

November 8, 2025

1 Excercise 1

1. Write down the loss function of the mean squared error over a mini-batch of B samples.

$$L_{MSE} = \frac{1}{2B} \sum_{i=1}^{B} (y_i - \hat{y}_i)^2$$

NOTE: in theory we can choose to not divide by $2B$. Since it is a relative constant, we can include it in the growth-rate, however writing it like this will allow us to write the derivative for gradient descent in a slightly more elegant way.

2. Briefly describe the idea of a mini-batch.

The idea of a mini-batch is to not use all of the available testing data to train the model at once. This allows us to not use too many resources for calculations regarding gradient descent as well as making over-fitting less likely.

3. Compute the partial derivatives of the Loss Function w. r. t x, θ_1, θ_2