Homework 1 – Linear Regression & Gradient Descent

Exercise 1: You have a small training dataset (one attribute x, one class y and 5 datapoints) and you want to learn a linear regression model using Gradient Descent algorithm to predict/explain the concept y using the attribute x. The following are the details for you to do your work.

Dataset:

x	y
1	3
3	5
4	7
6	8
7	9

Hypothesis: $f_\theta(x) = \theta_0 + \theta_1 x$

Parameters: θ_0 and θ_1

Loss/Error: $L(\theta_0, \theta_1) = \frac{1}{M}\sum_m (f_\theta(\boldsymbol{x}^{(m)}) - \boldsymbol{y}^{(m)})^2$

Goal: $\min_{\theta_0,\theta_1} L(\theta_0, \theta_1)$

Consider $\gamma = 0.015$, initial guess $\theta_0 = \theta_1 = 0$ for running the Gradient Descent algorithm.

Please compute what will be the final parameters (θ_0 and θ_1) after 4 iterations, showing all your calculations, including gradients and losses along all the 4 iterations.

Exercise 2: Do the same procedure executed in Exercise 1, but now you will use other values for learning rate γ:

(a) Use $\gamma = 0.04$. Do you see any difference when increasing the learning step γ from 0.015 to 0.04? Maybe plotting the lines and the points may help you to have an idea about how well the regressors are predicting $\boldsymbol{y}$.

(b) Use $\gamma = 0.001$. Again, do you see any difference when decreasing the learning step γ from 0.015 to 0.001?

Exercise 3: Do the same procedure executed in Exercise 1, but now you will use a slightly different loss function given below.

Loss/Error: $L(\theta_0, \theta_1) = \frac{1}{2M}\sum_m (f_\theta(\boldsymbol{x}^{(m)}) - \boldsymbol{y}^{(m)})^2$