ROBOT LEARNING
EXCERCISE- MACHINE LEARNING

SUMMER SEMESTER 2021

Problem 1

In this excercise you will implement the linear regression algorithm in a computer language of your choice. A simple dataset is provided with two variables, X denoting the Population of a city and Y denoting the Profit of a Coffee Shop in that city. Dataset can be downloaded here: https://github.com/sa32953/lero

A. Segregate the data into Input and Output Variable and plot the dataset to visualize.

B. Implement the Gradient Descent algorithm with Square Loss function and report the results.

(Consider error allowance in weight vector as 10^{-4} for convergency; Learning Rate = 0.01).

C. Implement Gradient Descent with Absolution Loss function. Plot both regression lines in one plot.

Hint: Use subgradient for Absolute Loss func with parameters in part B.

D. Vary the learning rate the observe the number of iterations algorithm takes to converge. Plot the graph of Rate vs Iterations in Sqaured Loss func.

[THIS PROBLEM IS NOT FULLY SOLVED BY ME, HAVING A PROBLEM WITH THE CODE].

Problem 2

A. Prove that with Squared Loss function, analytical solution can be denoted as :

$$w^\star = \left(\sum_{i=1}^{m} x_i x_i^T\right)^{-1} \left(\sum_{i=1}^{m} x_i y_i\right)$$

B. Prove that the same can be represented compactly in ‚Normal Equations‘ format:

$$w^\star = (X^T X)^{-1} X^T y$$

C. Is is true for all types of loss functions ? Why not ?

Problem 3

In which cases the Sqaured Loss can lead to a improper result ? Is there a better Loss function in that case?

Problem 4

A. Suppose that each output y is equal to hypotheses function $h_w(x)$ plus some Gaussian Noise e.

$$y = h_w(x) + e$$

With the probability density function $p(e) = \frac{1}{2\pi\sigma^2} exp(\frac{-e^2}{2\sigma^2})$. Prove that this approach is similar to that of Least-Sqaured Regression (minimizing Squared Loss func).

B. In Problem 3 we saw that results depend on type of loss function we choose. Is this problem really solved with using Probabilistic Representation?