Exercise 1: A basic supervised problem

I. Goal of the exercise

In this exercise you will practice the basic pipeline of the supervised learning task. Implement a simple regressor. And will try to solve several hinderances found in the process.

II. Deliverables

As you progress in this exercise, you will find several questions you are expected to answer them properly with adequate figures when required and deliver a document with all these evidences in due time. A file or files with the working code used for generating and discussing the results must be also delivered.

III. Our fist regressor.

We are given the data in `'reg_data_set1.mat'`(univariate basic dataset) our goal is to predict the y value. Our first model to try is linear regression as explained in "A gentle introduction to supervised learning".

A. Data set analysis

Load the dataset and describe the basic properties of the data,

Question block 1:

1) Plot the training samples and their corresponding label.

B. Analytical solution for the linear regression method.

In this part, you are expected to implement the analytical version of linear regression.

Question block 2:

1) Which is the optimal value of the linear regression weights?
2) Plot the data set and the line learned by the model. Does it looks like a good linear approximation?

C. Linear regression and the descent method.

In this part, you are expected to implement the basic gradient descent version for optimizing linear regression with fixed learning rate.

Question block 3:

1) Which is the optimal value of the linear regression weights using the descent method?
2) Which are the parameters of the descent method used to obtain the optimal value? i.e. learning rate value, number of iterations.
3) Plot the convergence curve of the method.
4) Zoom in the flat convergence part. Does it oscillate? Why?
5) Change the learning rate to 0.1. Plot the convergence curve of the method.

Modify the descent algorithm so that your descent direction is

$$\Delta x = -\frac{\nabla f(x)}{\|\nabla f(x)\|_2}$$

Question block 4:

1) Which is the optimal value of the linear regression weights using the modified descent method?
2) Which are the parameters of the descent method used to obtain the optimal value? i.e. learning rate value, number of iterations.
3) Plot the convergence curve of the method.
4) Zoom in the flat convergence part. Does it oscillate? Why?

D. A second model.

Consider now a polynomial model of degree p for the univariate case, i.e. $f(x;w) = w_0 + w_1 x + w_2 x^2 + \cdots + w_p x^p$. Observe that it can be expressed as a linear combination of transformed data points $\mathbf{z} = \{1, x, x^2, \ldots, x^p\}$,

$$f(x;w) = \sum_{i=0}^{p} w_i x^i = \sum_{i=0}^{p} w_i z_i = \mathbf{w}^T \mathbf{z}.$$

Question block 5:

1) Transform the training set into the set with examples described by z considering $p = 3$. Apply, the analytic solution code (if properly coded it should work without modifications). Which is the optimal value of the weights?
2) Plot the data set and the curve just found. Does it fit better the data? Why?

E. Evaluating a model.

The data set `'reg_data_set2.mat'` is a small set for regression. In this exercise we just want to select the best model among the following $p = 1, 2, \ldots, 6$. The best fit is given by the polynomial displaying smaller RMS (root mean square) value. RMS is defined as the squared loss for the data considered

$$RMS = \sqrt{\sum_{i=1}^{n}(f(x_i; w) - y_i)^2}$$

Question block 7:

1) Use the **first** half of the data set for training and the **second** half for validation.
2) Optimize the models (you can use any of the methods implemented before) and plot the validation set and the 6 models plots.
3) Plot the training set and the 6 models plots.
4) Compute the RMS error on the training set and on the validation set. Plot both error curves and describe their behavior.
5) Does the selected model agrees with the model that performs the best on the training set? Which one do you think is the optimal choice? Why? (You are not supposed to give an exhaustive answer to this question but your impressions and ideas. We will see the reasons why this effect happens shortly.)