ML - Linear Regression

Aims

- Evaluate the mean squared error loss function

- Find the closed-form solution for a set of observations

Tasks

- For Python users, familiarize yourself with Jupyter Notebook (http://jupyter.org/) and Matplotlib (https://matplotlib.org/)

- Define the following 4 observations:
  $(x_1, y_1) = (1.3, 7.5)$
  $(x_2, y_2) = (2.4, 8.5)$
  $(x_3, y_3) = (5.9, 19.1)$
  $(x_4, y_4) = (8.1, 25.2)$
  which are assumed to be linearly dependent, i.e.

$$y_i = w_0 \cdot x_i + w_1$$

  for unknown parameters w_0 and w_1.

- Evaluate the mean squared error for:
  $(\hat{w}_0, \hat{w}_1) = (4.1, 2.7)$
  $(\hat{w}_0, \hat{w}_1) = (0.7, 3.4)$

- Plot the lines associated with the two sets of parameters along with the observed points.

- Find the closed-form solution for (w_0, w_1) given the observations. Recall that

$$w_1 = \frac{\overline{(xy)} - \overline{x} \cdot \overline{y}}{\overline{(x^2)} - (\overline{x})^2}, \qquad w_0 = \overline{y} - w_1 \overline{x}$$

- Evaluate the mean squared error for the optimal parameters and verify that it is smaller than the previous error values.

- Plot the line associated with the optimal parameters against the observed points.