Problem 5.

For one dimensional data and label vectors, $\vec{x} = (x_1, x_2, \ldots, x_n)$, $\vec{y} = (y_1, y_2, \ldots, y_n)$
We will find $h(x) = ax+b$ that realizes the minimum MSE.

$f(a,b) = ax+b$, R^2 is defined as $\sum (y_i - f(x_i, a))^2$

$$R^2(a,b) = \sum_{i=1}^{n} [y_i - (ax_i + b)]^2$$

$$\frac{\partial R^2}{\partial b} = -2\sum_{i=1}^{n}[y_i - (ax_i + b)] = 0, \quad \frac{\partial R^2}{\partial a} = -2\sum_{i=1}^{n}[y_i - (bx_i + a)]x_i = 0$$

We can produce the following equations:

$$nb + a\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \quad b\sum_{i=1}^{n} x_i + a\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix}$$

$$\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} b \\ a \end{bmatrix} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \cdot \begin{bmatrix} \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i \\ n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \end{bmatrix}$$

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = \frac{(\sum_{i=1}^{n} x_i y_i) - n\vec{x}\vec{y}}{\sum_{i=1}^{n} x_i^2 - n\vec{x}^2}$$

$$b = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = \frac{\vec{y}(\sum_{i=1}^{n} x_i^2) - \vec{x}\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2 - n\vec{x}^2}$$