Problem 3

Given data points $(x_1, y_1), \ldots, (x_m, y_m)$, we define the design matrix:

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix}, \quad \theta = \begin{bmatrix} a \\ b \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

The normal equation for linear regression is given by:

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

Computing $X^T X$:

$$X^T X = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & m \end{bmatrix}$$

Now, we compute $X^T y$:

$$X^T y = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m y_i \end{bmatrix}$$

Thus, the normal equation becomes:

$$\theta = \begin{bmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & m \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m y_i \end{bmatrix}$$

Thus,

$$a = \frac{m \sum_{i=1}^m x_i y_i - \left(\sum_{i=1}^m x_i\right)\left(\sum_{i=1}^m y_i\right)}{m \sum_{i=1}^m x_i^2 - \left(\sum_{i=1}^m x_i\right)^2}$$

$$b = \frac{\left(\sum_{i=1}^m x_i^2\right)\left(\sum_{i=1}^m y_i\right) - \left(\sum_{i=1}^m x_i\right)\left(\sum_{i=1}^m x_i y_i\right)}{m \sum_{i=1}^m x_i^2 - \left(\sum_{i=1}^m x_i\right)^2}$$