Assignment 1

Student: Mascarenhas Alexandre | Professor: Keisuke KAMEYAMA | University of Tsukuba

1. Given a set of samples $\{(x_i, y_i)\}_{i=1}^{N}$ derive a minimum squared error regression of the data using a n-th order polynomial

$$y = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n \tag{1}$$

Solution:

This equation will be used to find the minimum values of the coefficients:

$$\hat{a} = (A^T A)^{-1} A^T b \tag{2}$$

where:

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_i & x_i^2 & \cdots & x_i^n \end{bmatrix} \textit{and } b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

The equation (2) comes when we derive the equation which represents the error and set it equal to zero, it will give us a column matrix with the minimized coefficients. Evaluating the transpose matrix of A, we have:

$$A^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_i \\ x_1^2 & x_2^2 & \cdots & x_i^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & \cdots & x_i^n \end{bmatrix}$$

So

$$A^T A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_i \\ x_1^2 & x_2^2 & \cdots & x_i^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & \cdots & x_i^n \end{bmatrix} \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_i & x_i^2 & \cdots & x_i^n \end{bmatrix} = \begin{bmatrix} I & \sum x_i & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{n+1} \\ \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_i^n & \sum x_i^{n+1} & \cdots & \sum x_i^{2n} \end{bmatrix}$$

And similarly

$$A^T b = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \\ \vdots \\ \sum x_i^n y_i \end{bmatrix}$$

So, finally, we just have to computing the multiplication $(A^T A)^{-1} A^T b$ in order to find the minimum values of the coefficients. This values will depends of the nature of the problem and for the given input data x and y.

$$\hat{a} = \begin{bmatrix} I & \sum x_i & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{n+1} \\ \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_i^n & \sum x_i^{n+1} & \cdots & \sum x_i^{2n} \end{bmatrix}^{-1} \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \\ \vdots \\ \sum x_i^n y_i \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_i \end{bmatrix} \tag{3}$$