CS7357 HW1, Noah Gardner, 000843905

1 Assignment 1

1. **[Points 30]** We have given 5 students aptitude test marks and their statistically computed grades below.

Student	Marks	Grade
1	95	85
2	85	95
3	80	70
4	70	65
5	60	70

   (a) **[Points 20]** We are interested to compute grades from given test marks using linear regression. Can you please estimate the line equation using the normal equation method? Please show the detailed computation of each step. Also, show the final line equation and its parameters.

   Suppose $h(\theta) = \theta_0 + \theta_1 x_1$

   $$\overrightarrow{h(\theta^*)} = \begin{bmatrix} 85 \\ 95 \\ 70 \\ 65 \\ 70 \end{bmatrix}$$

   $$X = A = \begin{bmatrix} 1 & 95 \\ 1 & 85 \\ 1 & 80 \\ 1 & 70 \\ 1 & 60 \end{bmatrix}$$

   $$A^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 95 & 85 & 80 & 70 & 60 \end{bmatrix}$$

   $$A^T * A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = \begin{bmatrix} 5 & 390 \\ 390 & 31150 \end{bmatrix}$$

   where $a_1 = 1 * 1 + 1 * 1 + 1 * 1 + 1 * 1 + 1 * 1 = 5$

   $a_2 = 1 * 95 + 1 * 85 + 1 * 80 + 1 * 70 + 1 * 60 = 390$

   $a_3 = 95 * 1 + 85 * 1 + 80 * 1 + 70 * 1 + 60 * 1 = 390$

   $a_4 = 95 * 95 + 85 * 85 + 80 * 80 + 70 * 70 + 60 * 60 = 31150$

   Finally,

   $$(A^T A)^{-1} = \begin{bmatrix} 8.53425 & -0.106849 \\ -0.106849 & 0.00136986 \end{bmatrix}$$

$\overrightarrow{\theta^*} = (A^T A)^{-1} * A^T * \overrightarrow{h(\theta^*)} = \begin{bmatrix} 26.7808 \\ 0.643836 \end{bmatrix}$

Therefore, by the normal equation, the line going through the sample data is: $h(\theta) = 26.7808 + 0.643836x_1$

(b) **[Points 10]** Predict grades for the following given student marks using your computed equation in step 1. Given marks `[65, 75, 77, 83, 87]`

The output values are `[68.6301, 75.0685, 76.3562, 80.2192, 87]`