CS584 – Machine Learning
Homework #1 Writing
Submitted by: Abdurakhmon Urazboev (aurazboev@hawk.iit.edu)

1. To solve this linear regression in closed form, we need to calculate the weight by the formula we that is used in the lecture $w = (X^T X)^{-1} X^T Y$.
   To do that we add an intercept (column of ones) to construct 2 by 4 matrix X and perform our calculations.
   $X = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ so, now we transpose the matrix to get $X^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix}$.
   Now we will calculate the multiplication:
   $$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} (1+1+1+1) & (-1+0+1+2) \\ (-1+0+1+2) & (1+0+1+4) \end{bmatrix}$$
   $$= \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}$$
   Now we have 2x2 matrix which we need to inverse. By the formula:
   $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, the invers of the matrix will be $(X^T X)^{-1} = \frac{1}{24-4} \begin{bmatrix} 6 & -2 \\ -2 & 4 \end{bmatrix} =$ $0.05 \times \begin{bmatrix} 6 & -2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 0.3 & -0.1 \\ -0.1 & 0.2 \end{bmatrix}$
   Now we need to calculate the second half of the formula.
   $$X^T Y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} (0+2+4+5) \\ (0+0+4+10) \end{bmatrix} = \begin{bmatrix} 11 \\ 14 \end{bmatrix}.$$
   As we have both parts now, we can multiply them together to get the weight:
   $$w = \begin{bmatrix} 0.3 & -0.1 \\ -0.1 & 0.2 \end{bmatrix} \begin{bmatrix} 11 \\ 14 \end{bmatrix} = \begin{bmatrix} (3.3-1.4) \\ (-1.1+2.8) \end{bmatrix} = \begin{bmatrix} 1.9 \\ 1.7 \end{bmatrix}$$
   So, the liner model will have $w_1 = 1.7, b = 1.9$ parameters giving us $y = 1.7x + 1.9$.
2. To calculate MSE and MAE we use following formulas respectively:
   $$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \; and \; MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$
   Let's get started with calculation:
   For $x = -1, y = 0, \hat{y} = 1.7 * (-1) + 1.9 = 0.2$
   For $x = 0, y = 2, \hat{y} = 1.7 * 0 + 1.9 = 1.9$
   For $x = 1, y = 4, \hat{y} = 1.7 * 1 + 1.9 = 3.6$
   For $x = 2, y = 5, \hat{y} = 1.7 * 2 + 1.9 = 5.3$
   Now, as we have all the values, we can calculate MSE and MAE.

$$MSE = \frac{1}{4}((0-0.2)^2 + (2-1.9)^2 + (4-3.6)^2 + (5-5.3)^2)$$
$$= \frac{1}{4}(0.04 + 0.01 + 0.16 + 0.09) = \frac{1}{4}(0.3) = \frac{3}{40} = 0.075$$
$$MAE = \frac{1}{4}(|0-0.2| + |2-1.9| + |4-3.6| + |5-5.3|) = \frac{1}{4}(0.2 + 0.1 + 0.4 + 0.3)$$
$$= \frac{1}{4} = 0.25$$

3. In order to estimate bias, we calculate the mean of differences of actual values from predicted values. $bias = \frac{1}{4}\big((0.2-0) + (1.9-2) + (3.6-4) + (5.3-5)\big) = \frac{1}{4}(0.2 - 0.1 - 0.4 + 0.3 = 0$. As we can see Average difference in this case is 0.

   To calculate variance, we need to calculate mean of all predicted values $\bar{y}' = \frac{1}{4}(0.2 + 1.9 + 3.6 + 5.3) = 2.75$. Now we can calculate variance by this formula: $Variance = \frac{1}{n}\sum_{i=0}^{n}(\hat{y} - \bar{y}')^2 = \frac{1}{4}((0.2-2.75)^2 + (1.9-2.75)^2 + (3.6-2.75)^2 + (5.3-2.75)^2) = \frac{1}{4}(6.5025 + 0.7225 + 0.7225 + 6.5025) = \frac{14.45}{4} = 3.612$