CS 613 - Machine Learning

Assignment 1 - Linear Regression

Robert Thompson

April 11, 2022

1 Theory

$$X = \begin{bmatrix} -2 \\ -5 \\ -3 \\ 0 \\ -8 \\ -2 \\ 1 \\ 5 \\ -1 \\ 6 \end{bmatrix}, Y = \begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix}$$

1.1 Add Bias Feature

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -5 \\ 1 & -3 \\ 1 & 0 \\ 1 & -8 \\ 1 & -2 \\ 1 & 1 \\ 1 & 5 \\ 1 & -1 \\ 1 & 6 \end{bmatrix}, Y = \begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix}$$

1.2 Direct Solution

$w = (X^TX)^{-1}X^TY$

$$X^TX = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -5 & -3 & 0 & -8 & -2 & 1 & 5 & -1 & 6 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -5 \\ 1 & -3 \\ 1 & 0 \\ 1 & -8 \\ 1 & -2 \\ 1 & 1 \\ 1 & 5 \\ 1 & -1 \\ 1 & 6 \end{bmatrix}$$

$$X^TX = \begin{bmatrix} 10 & -9 \\ -9 & 169 \end{bmatrix}$$

$$(X^TX)^{-1} = \frac{1}{|(10*169)-(-9*-9)|} \begin{bmatrix} 169 & 9 \\ 9 & 10 \end{bmatrix} = \frac{1}{1690-81} \begin{bmatrix} 169 & 9 \\ 9 & 10 \end{bmatrix} = \frac{1}{1609} \begin{bmatrix} 169 & 9 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 169/1609 & 9/1609 \\ 9/1609 & 10/1609 \end{bmatrix} = \begin{bmatrix} 0.10503418 & 0.00559354 \\ 0.00559354 & 0.00621504 \end{bmatrix}$$

$$(X^TX)^{-1}X^T = \begin{bmatrix} 0.10503418 & 0.00559354 \\ 0.00559354 & 0.00621504 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -5 & -3 & 0 & -8 & -2 & 1 & 5 & -1 & 6 \end{bmatrix}$$

$(X^TX)^{-1}X^T =$

$$\begin{bmatrix} 0.09384711 & 0.0770665 & 0.08825357 & 0.10503418 & 0.06028589 & 0.09384711 & 0.11062772 \\ -0.00683654 & -0.02548167 & -0.01305158 & 0.00559354 & -0.04412679 & -0.006836540.01180858 & 0.03666874 \end{bmatrix}$$

$(X^TX)^-1X^TY =$

$$\begin{bmatrix} 0.09384711 & 0.0770665 & 0.08825357 & 0.10503418 & 0.06028589 & 0.09384711 & 0.11062772 \\ -0.00683654 & -0.02548167 & -0.01305158 & 0.00559354 & -0.04412679 & -0.006836540.01180858 & 0.03666874 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.0285891858297085 & -0.4126786824114355 \end{bmatrix}$$

$w = \begin{bmatrix} 1.0285891858297085 & -0.4126786824114355 \end{bmatrix}$

1.3 $\hat{Y}$ Predictions

1. Learned Model: $y = 1.0285891858297085 + -0.4126786824114355 x_{:,1}$
2. Predictions
   - $\hat{Y} = \begin{bmatrix} 1.85394655 \\ 3.0919826 \\ 2.26662523 \\ 1.02858919 \\ 4.33001865 \\ 1.85394655 \\ 0.6159105 \\ -1.03480423 \\ 1.44126787 \\ -1.44748291 \end{bmatrix}$

1.4 RMSE and MAPE

1. Root Mean Squared Error (RMSE): 3.7013259176662716
2. Mean Absolute Squared Error (MAPE): 142.73053114282442

2 Closed Form (Direct) Linear Regression

1. Final Model: $y = -131.04963658130077 + 4.159936830388848x_{:,1} + 0.030819358920040 47x_{:,2}$
2. Training Output
   - Root Mean Squared Error (RMSE): 19.86256862907285
   - Mean Absolute Percentage Error (MAPE): 21.397596050628863
3. Validation Output
   - Root Mean Squared Error (RMSE): 20.067704981328184
   - Mean Absolute Percentage Error (MAPE): 30.47810152100226

3 S-Folds Cross-Validation

1. With S-Fold = 4
   - Mean of RMSE: 21.599081781852323
   - Standard Deviation of RMSE: 2.5030965382840233
2. With S-Fold = 11
   - Mean of RMSE: 21.092988359747313
   - Standard Deviation of RMSE: 2.29872805834 30545
3. With S-Fold = 22
   - Mean of RMSE: 19.725759616288194
   - Standard Deviation of RMSE: 1.7199475688764514
4. With S-Fold = N
   - Mean of RMSE: 20.994992352914746
   - Standard Deviation of RMSE: 2.7099268309540 12

4 Locally-Weighted Linear Regression

1. Validation Root Mean Squared Error (RMSE): 26.396286898865213
2. Validation Mean Absolute Percentage Error (MAPE): 28.76650617672107