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

Matrices below are rounded to four decimals places due to page space limitations but the resulting weights (coefficients) and $\hat{Y}$ predictions contain all decimal places.

$$w = (X^T X)^{-1} X^T Y$$

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

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

$$(X^T X)^{-1} = \frac{1}{|(10*169)-(-9*-9)|} \begin{bmatrix} 169 & 9 \\ 9 & 10 \end{bmatrix} = \frac{1}{1690-81} \begin{bmatrix} 169 & 9 \\ 9 & 10 \end{bmatrix} =$$
$$\frac{1}{1609} \begin{bmatrix} 169 & 9 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 169/1609 & 9/1609 \\ 9/1609 & 10/1609 \end{bmatrix} = \begin{bmatrix} 0.1050 & 0.0056 \\ 0.0056 & 0.0062 \end{bmatrix}$$

$$(X^T X)^{-1} X^T = \begin{bmatrix} 0.1050 & 0.0056 \\ 0.0056 & 0.0062 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -5 & -3 & 0 & -8 & -2 & 1 & 5 & -1 & 6 \end{bmatrix}$$

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

$$\begin{bmatrix} 0.0938 & 0.0770 & 0.0883 & 0.1050 & 0.0603 & 0.0938 & 0.1106 & 0.1330 & 0.0994 & 0.1386 \\ -0.0068 & -0.0255 & -0.0130 & 0.0056 & -0.0441 & -0.0068 & 0.0118 & 0.0367 & -0.0006 & 0.0429 \end{bmatrix}$$

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

$$\begin{bmatrix} 0.0938 & 0.0770 & 0.0883 & 0.1050 & 0.0603 & 0.0938 & 0.1106 & 0.1330 & 0.0994 & 0.1386 \\ -0.0068 & -0.0255 & -0.0130 & 0.0056 & -0.0441 & -0.0068 & 0.0118 & 0.0367 & -0.0006 & 0.0429 \end{bmatrix}$$

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

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

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 Percentage Error (MAPE) as Percent: 142.73053114282442
3. Mean Absolute Percentage Error (MAPE) as Decimal: 1.42730531143

2 Closed Form (Direct) Linear Regression

1. Final Model: $y = -131.04963658130077 + 4.159936830388848x_{:,1} + 0.03081935892004047x_{:,2}$
2. Training Output
    - Root Mean Squared Error (RMSE): 19.86256862907285
    - Mean Absolute Percentage Error (MAPE) as Percent: 21.397596050628863
    - Mean Absolute Percentage Error (MAPE) as Decimal: 0.2139759605
3. Validation Output
    - Root Mean Squared Error (RMSE): 20.067704981328184
    - Mean Absolute Percentage Error (MAPE) as Percent: 30.47810152100226
    - Mean Absolute Percentage Error (MAPE) as Decimal: 0.30478101521

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: 2.2987280583430545
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: 2.709926830954012

4 Locally-Weighted Linear Regression

1. Validation Root Mean Squared Error (RMSE): 26.396286898865213
2. Validation Mean Absolute Percentage Error (MAPE) Percent: 28.76650617672107
3. Validation Mean Absolute Percentage Error (MAPE) as Decimal: 0.28766506176