Greg Matthews
CS 383
February 8th , 2018

Assignment 3 – Linear Regression

Part 1: Theory

1. Consider the following data:

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

a) Compute the coefficients for the linear regression using global least squares estimate (LSE) where the second value is the dependent variable (the value to be predicted). Show your work and remember to add a bias feature and to standardize the features (10pts). Compute this model using all of the data (don't worry about separating into training and testing sets).

Answer:

Step 1: Standardize Data

$Feature\ 1$: $\mu = -0.9,\ \sigma = 4.23$ $\quad$ $where\ \mu = mean,\ \sigma = standard\ deviation,$

$$X_{new} = \frac{X_{i,j} - \mu}{\sigma} \quad \forall i, j \in X$$

Data after standardizing becomes:

$$\begin{bmatrix} -0.260 & 1 \\ -0.969 & -4 \\ -0.497 & 1 \\ 0.212 & 3 \\ -1.68 & 11 \\ -0.260 & 5 \\ 0.449 & 0 \\ 1.39 & -1 \\ -0.024 & -3 \\ 1.63 & 1 \end{bmatrix}$$

Step 2: Adding bias feature:

$$\begin{bmatrix} 1 & -0.260 & 1 \\ 1 & -0.969 & -4 \\ 1 & -0.497 & 1 \\ 1 & 0.212 & 3 \\ 1 & -1.68 & 11 \\ 1 & -0.260 & 5 \\ 1 & 0.449 & 0 \\ 1 & 1.39 & -1 \\ 1 & -0.024 & -3 \\ 1 & 1.63 & 1 \end{bmatrix}$$

Step 3: Computing weight matrix θ:

From least square estimate:

$$J(\theta) = \sum_{i=1}^{N} (Y_i - X_i\theta)^2$$

In vectorized form:

$$J(\theta) = (Y - X\theta)^{\mathrm{T}}(Y - X\theta)$$

Taking Derivative w.r.t. θ and solving for θ:

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

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -0.260 & -0.969 & -0.497 & 0.212 & -1.68 & -0.260 & 0.449 & 1.39 & -0.024 & 1.63 \end{bmatrix} \begin{bmatrix} 1 & -0.260 \\ 1 & -0.969 \\ 1 & -0.497 \\ 1 & 0.212 \\ 1 & -1.68 \\ 1 & -0.260 \\ 1 & 0.449 \\ 1 & 1.39 \\ 1 & -0.024 \\ 1 & 1.63 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 10 & 0 \\ 0 & 8.98 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 10 & 0 \\ 0 & 8.98 \end{bmatrix}^{-1} = \frac{1}{10 * 8.98 - 0 * 0} * \begin{bmatrix} 8.98 & 0 \\ 0 & 10 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.11 \end{bmatrix}$$

$$(X^T X)^{-1} X^T = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.11 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -0.260 & -0.969 & -0.497 & 0.212 & -1.68 & -0.260 & 0.449 & 1.39 & -0.024 & 1.63 \end{bmatrix}$$

$$(X^T X)^{-1} X^T = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ -0.0289 & -0.108 & -0.0551 & 0.0237 & -0.187 & -0.0289 & 0.0449 & 0.155 & -0.00263 & 0.181 \end{bmatrix}$$

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

$$\begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ -0.0289 & -0.108 & -0.0551 & 0.0237 & -0.187 & -0.0289 & 0.0449 & 0.155 & -0.00263 & 0.181 \end{bmatrix} \begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix}$$

$$(X^T X)^{-1} X^T Y = \boldsymbol{\theta} = \begin{bmatrix} \mathbf{1.4} \\ \mathbf{-1.74} \end{bmatrix}$$

Coefficients are therefore: **{1.4, -1.74}** Equating to: $\boldsymbol{\hat{y} = 1.4 - 1.74x}$

Part 2: Closed Form Linear Regression:

(a) Final Model

$$\boldsymbol{\hat{y} = 3343 + 1037x_1 - 295.7x_2}$$

(b) RMSE

$$\boldsymbol{RMSE = 653.7}$$

Part 3: S-Folds Cross-Validation:

1. The root mean squared error for S = 3: $\boldsymbol{RMSE = 696.6}$
2. The root mean squared error for S = 5: $\boldsymbol{RMSE = 675.6}$
3. The root mean squared error for S = 20: $\boldsymbol{RMSE = 608.1}$
4. The root mean squared error for S = N: $\boldsymbol{RMSE = 599.8}$