1 Theory

1. Consider the following supervised *training* dataset:

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

(a) $(\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -5 & -3 & 0 & -8 & -2 & 1 & 5 & -1 & 6 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -5 \\ 1 & -3 \\ 1 & 0 \\ 1 & -8 \\ 1 & -2 \\ 1 & 1 \\ 1 & 5 \\ 1 & -1 \\ 1 & 6 \end{bmatrix})^{-1} *$

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

$\begin{bmatrix} 0.10503418 & 0.00559354 \\ 0.00559354 & 0.00621504 \end{bmatrix} * \begin{bmatrix} 14 & -79 \end{bmatrix}$

$w = \begin{bmatrix} 1.02858919 & -0.41267868 \end{bmatrix}$

(b) $\hat{Y} = \begin{bmatrix} 1 & -2 \\ 1 & -5 \\ 1 & -3 \\ 1 & 0 \\ 1 & -8 \\ 1 & -2 \\ 1 & 1 \\ 1 & 5 \\ 1 & -1 \\ 1 & 6 \end{bmatrix} * \begin{bmatrix} 1.02858919 & -0.41267868 \end{bmatrix}$

$\hat{Y} = \begin{bmatrix} 1.85394655 \\ 3.0919826 \\ 2.26662523 \\ 1.02858919 \\ 4.33001865 \\ 1.85394655 \\ 0.6159105 \\ -1.03480423 \\ 1.44126787 \\ -1.44748291 \end{bmatrix}$

(c) RMSE: 3.7013259176662716
SMAPE: 0.608752922166376

2 Closed Form Linear Regression

Training:
RMSE: 5757.954440690524
SMAPE: 0.18070534061034704

Validation:
RMSE: 6606.0300959684855
SMAPE: 0.18320284506099385

For pre-processing I found all unique values of non-integer observations, enumerated them based on when that unique value was seen and added a bias feature.

3   Cross-Validation

S = 3
mean: 6091.90452681748
standard deviation: 15.5622307451187

S = 223
mean: 6088.154335521822
standard deviation: 29.751920914863703

S = 1338
mean: 12145.095537090177
standard deviation: 4.866568299330092e-12

For pre-processing I found all unique values of non-integer observations, enumerated them based on when that unique value was seen and added a bias feature.