Theory

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)

Using Closed Form Linear Regression

1. Standardize the data.

$$\mu = \frac{-2 - 5 - 3 + 0 - 8 - 2 + 1 + 5 - 1 + 6}{10} = -0.9$$

$$\sigma = \frac{(-2 + 0.9)^2 + (-5 + 0.9)^2 + (-3 + 0.9)^2 \ ...}{10} = 4.2282$$

$$x = \frac{data - \mu}{\sigma} = \begin{bmatrix} -0.2602 \\ -0.9697 \\ -0.4967 \\ 0.2129 \\ -1.6792 \\ -0.2602 \\ 0.4494 \\ 1.3954 \\ -0.0237 \\ 1.6319 \end{bmatrix}$$

2. Compute the weights by using $\theta = (X^T X)^{-1} X^T Y$.

$$X = \begin{bmatrix} 1 & -0.2602 \\ 1 & -0.9697 \\ 1 & -0.4967 \\ 1 & +0.2129 \\ 1 & -1.6792 \\ 1 & -0.2602 \\ 1 & +0.4494 \\ 1 & +1.3954 \\ 1 & -0.0237 \\ 1 & +1.6319 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 10 & -0.0001 \\ -0.0001 & 9.0002 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1111 \end{bmatrix}$$

$$\theta = \begin{bmatrix} 1.4 \\ -1.7449 \end{bmatrix}$$

$$y = 1.4 - 1.7449 x_{:,1}$$

Closed Form Linear Regression

The final model in the form is

$$y = 3.4256 \times 10^3 + 846.9475 x_{:,1} - 369.2202 x_{:,2}$$

The root mean squared error.

RMSE:	853.380588

S-Folds Cross-Validation

The root mean squared error.

RMSE:	634.816732