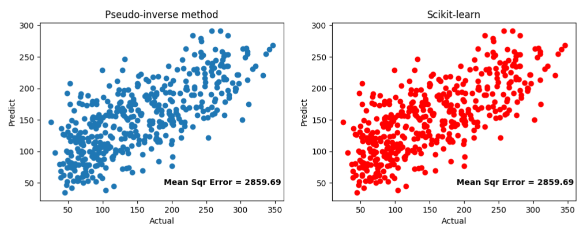
COMP6245 : Lab 4 Report

Thanakorn Panyapiang(tp2n19@soton.ac.uk)

1 Linear Least Square Regrssion

The result of the linear predictors using pseudo-inverse method and using scikit-learn is shown below.



Both methods give exactly the same result as can be seen from the identical mean square error.

2 Regularization

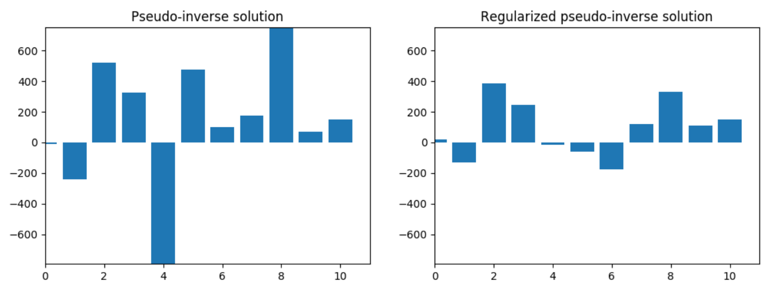
The gradient of L_2 regularization can be derived as follow:

$$\begin{aligned} \nabla_a E_{L2} &= \frac{\partial ||f - Ya||_2^2 + \gamma ||a||_2^2}{\partial a} \\ &= 2Y^T(f - Ya) + 2\gamma a \end{aligned} \tag{1}$$

Equating the gradient to 0 gives the closed form of a as follow

$$\begin{aligned} 0 &= 2Y^T(f - Ya) + 2\gamma a \\ Y^T Ya - \gamma a &= Y^T f \\ a(Y^T Y - \gamma) &= Y^T f \\ a &= (Y^t Y - \gamma)^{-1} Y^T f \end{aligned} \tag{2}$$

It can be observed that the coefficients of the predictor using L_2 regularization are significantly lower than the model without regularization. This is caused by adding a quadratic penalty of the weights to the error function. The coefficients of the model with and without L_2 regularization is shown below.



3 Sparse Regression

4 Solbility Prediction