THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

Report of EIE2112 Lab 3

Lam Ching Yiu

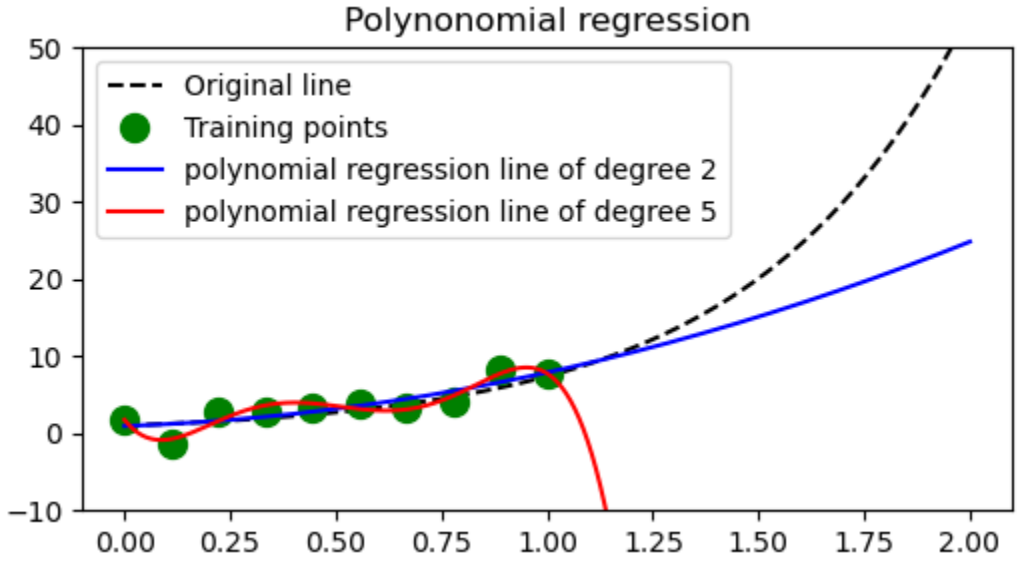
21093734D

Task 1: Polynomial regression model

In linear regression, we can train the line with training data to predict the function. However, it does not work well when predict the function, $y = e^{2x}$, because it is a polynomial function. So, we can use a polynomial representation model to learn the model and have a good predict result.

For example, we use the formula $$y_{tr,i} = \sum_{j=0}^{n} w_j * x_{tr,i}^j$$

Then using the formula, two models, polynomial regression line of degree 2 and polynomial regression line of degree 5 are created.



There is a huge difference between the two models. By observation, the model, polynomial regression of degree 5 is too sensitive to the training points and does not produce a good result of prediction.

It means that the polynomial regression line of degree 5 is overfit to the training data. It shows a higher degree of regression line, leading to a higher weight of the training data in the model.

Task 2: Ridge regression model

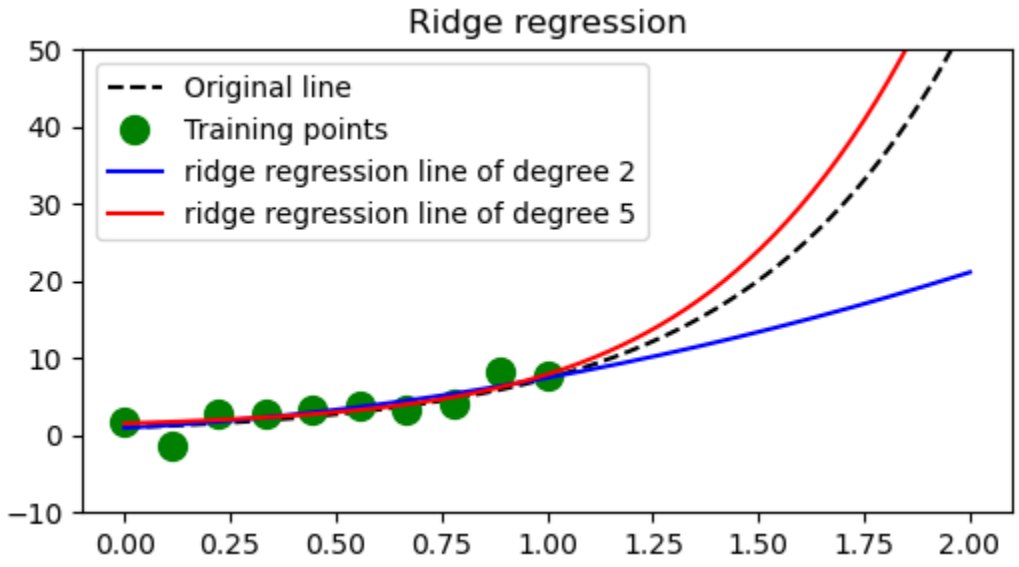
To improve the result of Task 1, we can use ridge regression. The ridge regression adds a regularization term when learning the model. So, the term will make the model less likely to overfit the training model and product the result of Task 1 polynomial regression degree 5.

In this task, we use l2 regularization. It adds a penalty term as known as regularization term to the loss function, so the total loss will be higher if the l2 norm is too high. That lead to the result that the model will learn using the training data and at the same time will not be overfit by the training data.

By using the formula ,

$$\arg \min_{w_0, w_1, \cdots, w_n} \sum_i (y_{tr,i} - \sum_{j=0}^{n} w_j * x_{tr,i}^j)^2 + \alpha * \sum_{j=0}^{n} w_j^2$$

And using the function RidgeCV to set the alphas value, which affects the weight of the penalty term. A large alpha can lead a higher weight of the penalty term and make the model less fit to the training data.



The above result shows a better model then task 1 when add a penalty term.

Conclusion

There are factors affecting the models using polynomial regression and ridge regression while training and learning.

The most important factors are below.

Degree of polynomial: The degree of polynomial means the data can be explained using a n-th order polynomial. A higher degree of polynomial leads the model more fit to the training data. However, there will be an overfit to training data if the degree of polynomial is too high and without a penalty term added.

Regularization term: As known as the penalty term, this term will be added to the loss function. It will cause the total loss to be higher. It means cause the model can learn the data but at the same time will be affected by the penalty term. The result will be the coefficient of the polynomial will not be overfit the training data and have a good result when using higher degree of polynomial.

Alpha: To be notes that alpha is used when using ridge regression. The alpha value affects the weight given to the penalty term. The penalty term will make the coefficient of the polynomial relatively small, which means that the model will not be overfit with the training data. Especially when the training data have a large difference with the others, the total loss will be higher.