MAE263F HW3

Zhaoyi Ye 906070940

I. Assignment 3

Q1

The predicted values versus the actual data:

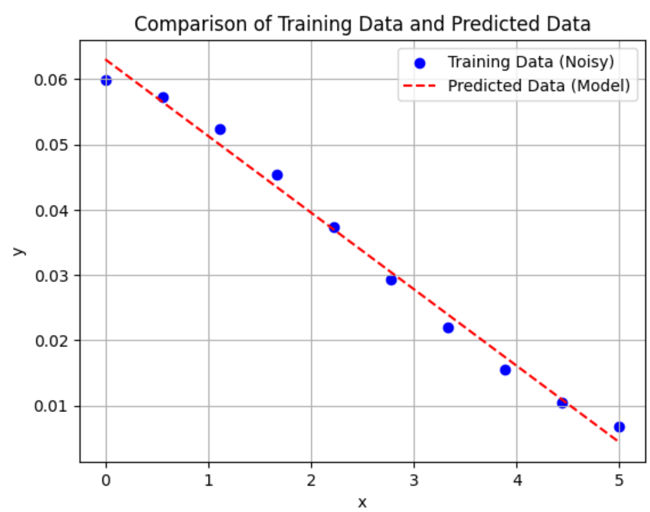

Fig.1 the predicted values versus the actual data with linear method

According to the fig.1, it shows the linear predicted data has similar shape with the trainning data. The loss when complete the prediction is 0.000003

Fig.2 the predicted values versus the actual data with linear method and smallest loss

According to the code ,the best combination that minimize the loss and converges are 0.0037 in learning rate and 13000 in number of epochs. The loss is 0.000003 for the final result. The smaller learning rate reduces losses during convergence while it requires larger number of epochs to reach convergence. Usually it will took longger time for calculation but result fit better. The larger learning rate will going faster in gradient descent and require smaller number of epochs but the result of loss could be larger compaired with smaller learning rate.

Q2

The predicted values versus the actual data:

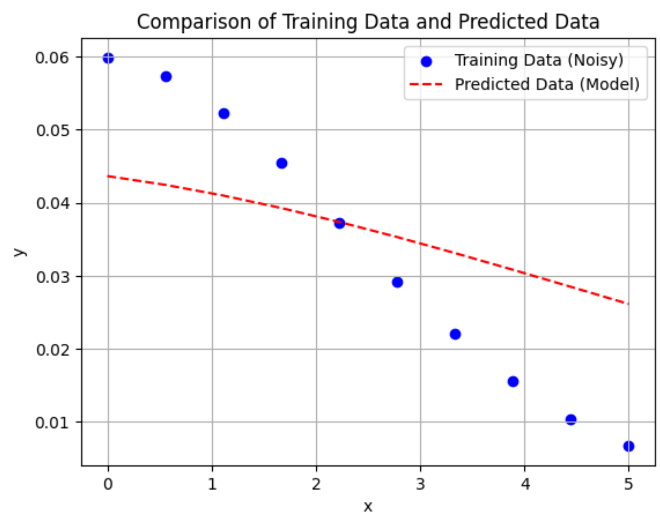

Fig.2 the predicted values versus the actual data with non linear method

According to the fig.2, it shows the nonlinear predicted data with the actual data. The figure did not show a good prediction of the data. From the epoch vs loss could find the result is converged but the loss is very large:0.00175.

Fig.4 the predicted values versus the actual data with linear method and smallest loss

According to the code ,the best combination that minimize the loss and converges are 0.0061 in learning rate and 1000 in number of epochs. The loss is 0.000002 for the final result. Similar as the linear method: A smaller learning rate results in a more gradual reduction in loss during convergence, requiring a greater number of epochs to reach the optimum. While this approach typically takes longer to compute, it often leads to a better fit for the model. On the other hand, a larger learning rate facilitates faster progress in gradient descent, requiring fewer epochs to converge. However, it may result in a higher final loss compared to using a smaller learning rate, as it can overshoot or fail to fine-tune the solution as effectively.

REFERENCES

[1] M. Khalid. Jawed and S. Lim, *Discrete simulation of slender structures*, ch 4