MAE 263F: Homework3_He

Jiahua He

***Abstract*— This electronic document is a report of Homework1 for 263F.**

I. Problem Statement

This study applies gradient descent and backpropagation to fit noisy data to linear and nonlinear models.

II. Linear model

A. *Problem 1*

After fitting the model, the predicted values y_{pred} are compared with the actual data y, as shown in the following below

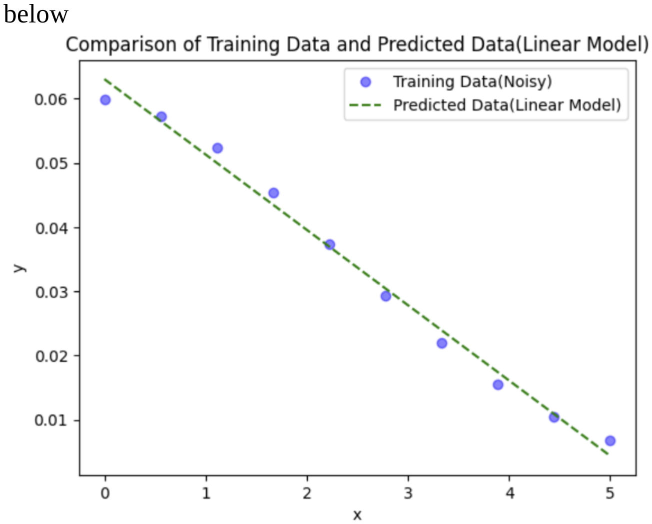

Figure. 1 y_{pred} vs. y

B. *Problem2*

Change the number of epochs and keep the learning rate. The results are shown below:

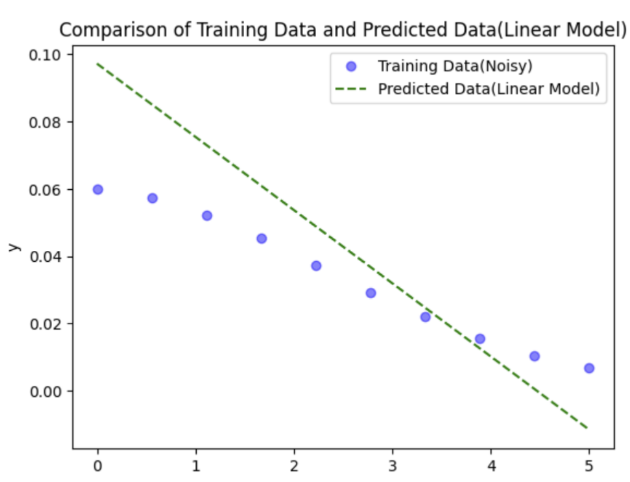

Figure. 2 y_{pred} vs. y (Epochs = 100, lr = 0.001)

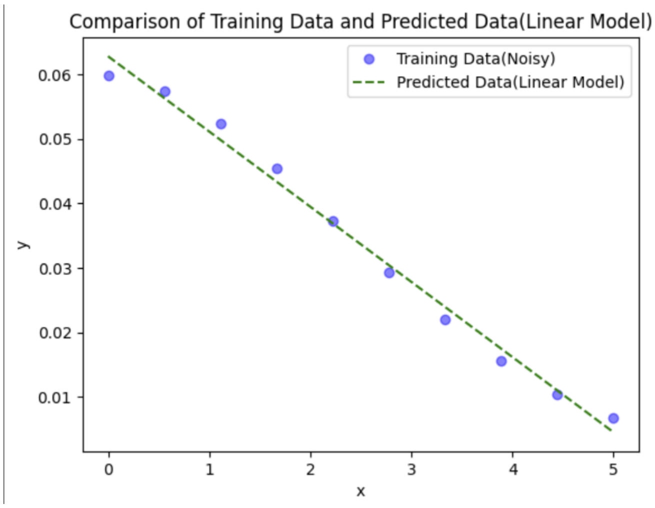

Figure. 3 y_{pred} vs. y (Epochs = 100000, lr = 0.001)

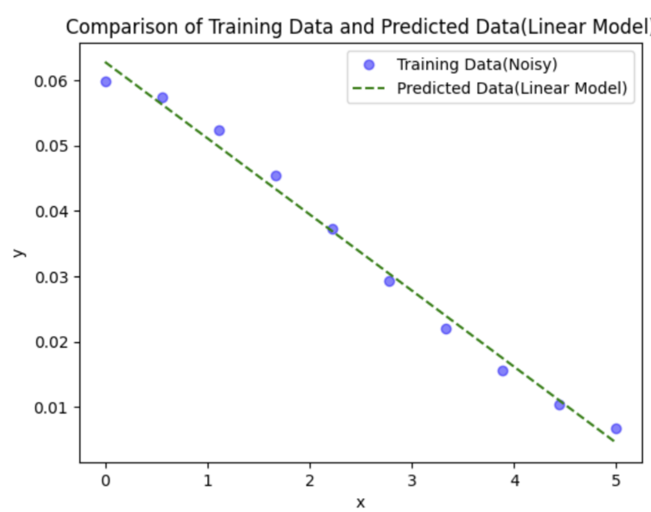

Figure. 4 y_{pred} vs. y (Epochs = 1000000, lr = 0.001)

As shown in the figure above, when the learning rate is fixed and the number of epochs is too small, the simulation fails to converge, leading to poor fitting performance. Additionally, from Figure 2&, it can be seen that increasing the number of epochs does not significantly enhance the actual fitting performance.

When keeping Epochs = 10000 and changing the learning rate. The results are shown below:

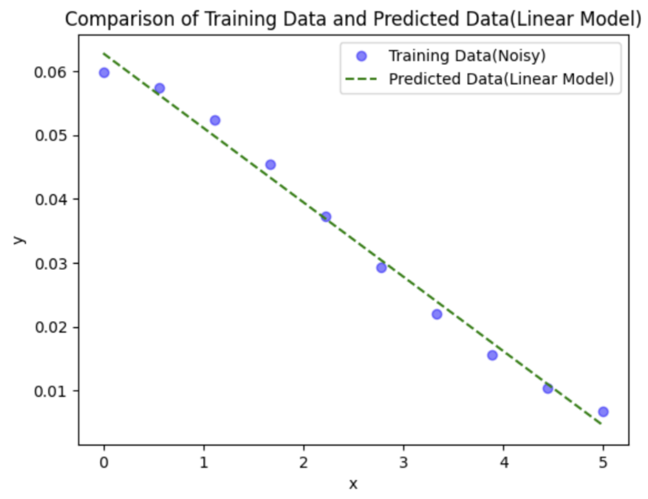

Figure. 5 y_{pred} vs. y (Epochs = 10000, lr = 0. 1)

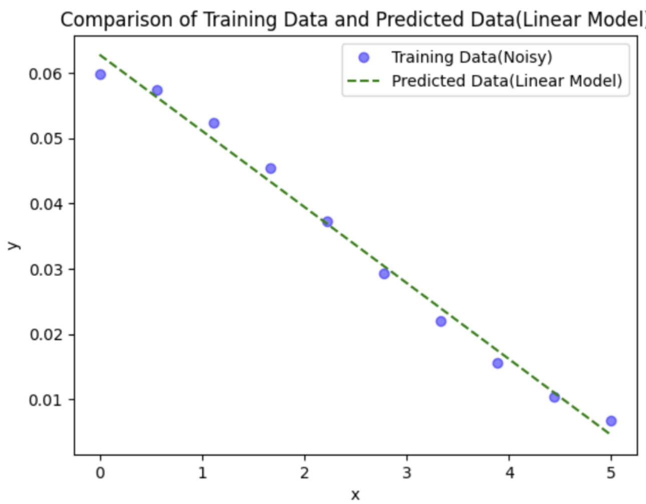

Figure. 6 y_{pred} vs. y (Epochs = 10000, lr = 0.01)

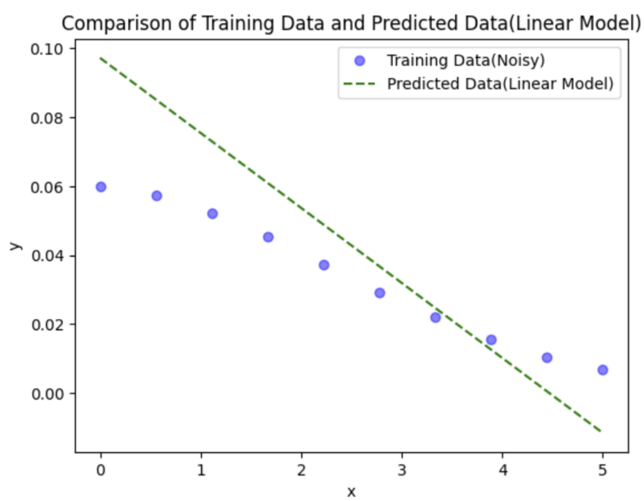

Figure. 7 y_{pred} vs. y (Epochs = 10000, lr = 0.0001)

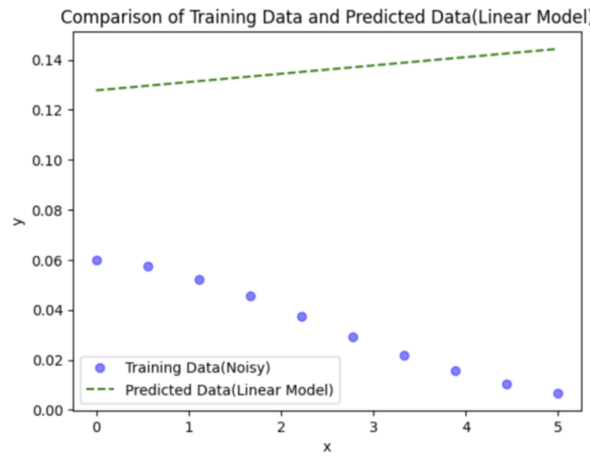

Figure. 8 y_{pred} vs. y (Epochs = 10000, lr = 0.00001)

From the figure above, as the learning rate becomes smaller, accuracy generally decreases, leading to less convergence and higher loss. When the learning rate is 0.01 or 0.1, results remain convergent.

Considering both performance and runtime, the optimal combination is epochs = 10000 and learning rate = 0.01.

III. NONLINEAR MODEL

A. *Problem 1*

By using the given parameter and code, the difference between actual number and predicted number (Epochs = 10000, lr = 0.001) is shown below:

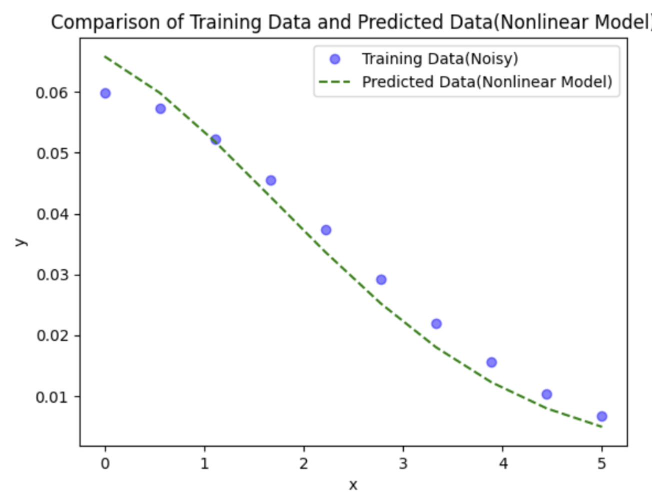

Figure. 9 y_{pred} vs. y (Epochs = 10000, lr = 0.001)

B. *Problem2*

Change the number of epochs and keep the learning rate. The results are shown below:

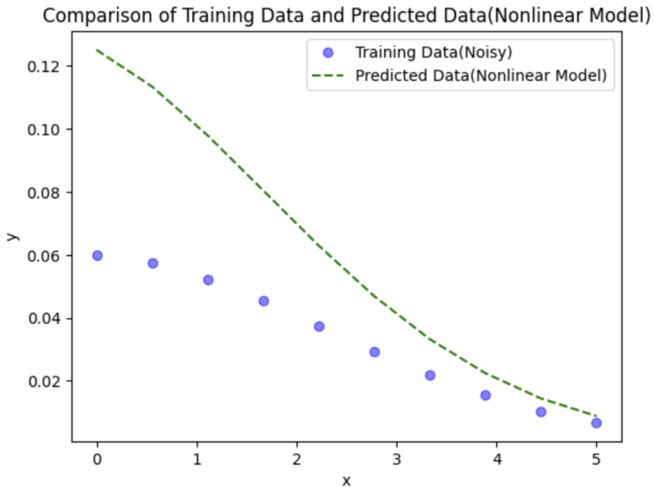

Figure. 10 y_{pred} vs. y (Epochs = 1000, lr = 0.001)

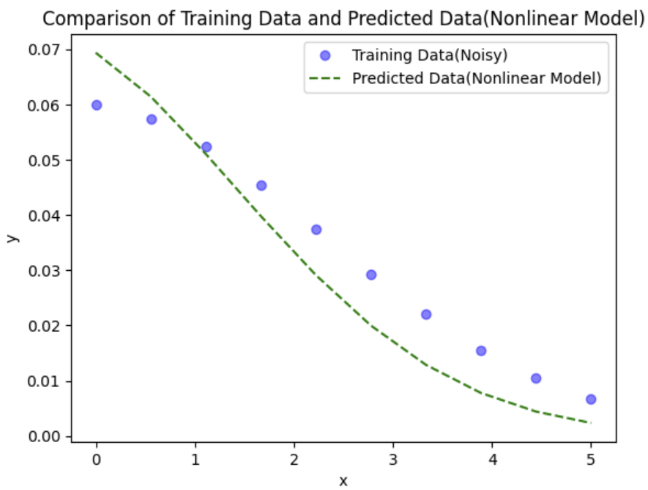

Figure. 11 y_{pred} vs. y (Epochs = 100000, lr = 0.001)

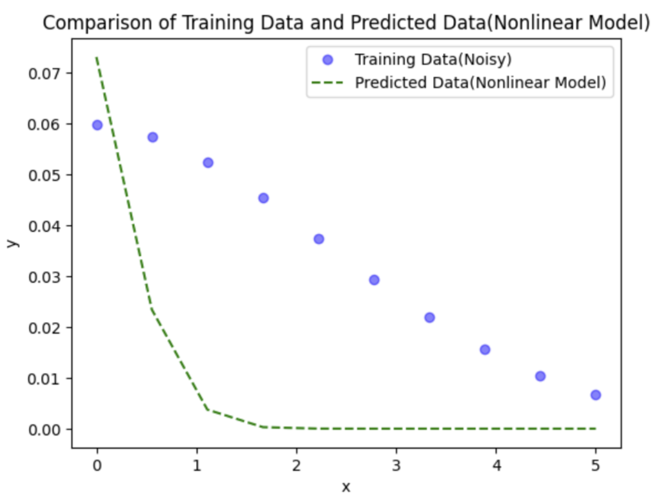

Figure. 12 y_{pred} vs. y (Epochs = 1000000, lr = 0.001)

With a small number of epochs, the results fail to converge, resulting in higher loss. Unlike linear models, increasing the number of epochs does not always improve accuracy. When the number of epochs is too large, the results may also lose convergence, and in extreme cases, the predictions can overshoot, leading to significant loss.

When keeping Epochs = 10000 and changing the learning rate. The results are shown below:

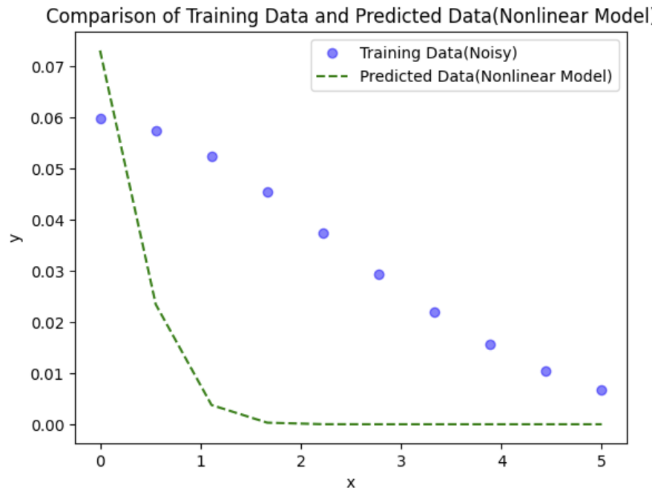

Figure. 13 y_{pred} vs. y (Epochs = 10000, lr = 0.1)

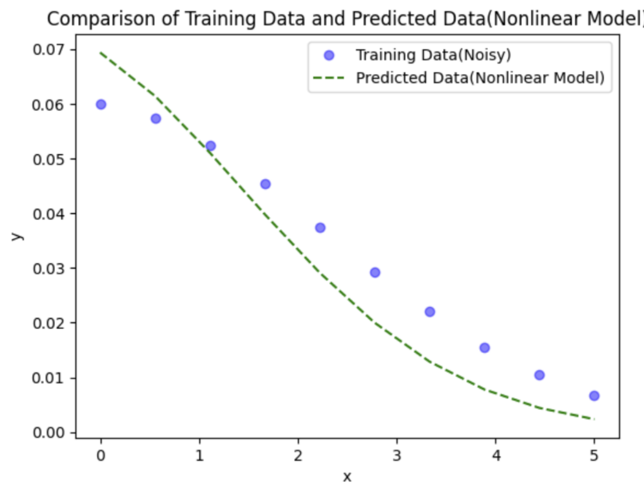

Figure. 13 y_{pred} vs. y (Epochs = 10000, lr = 0.01)

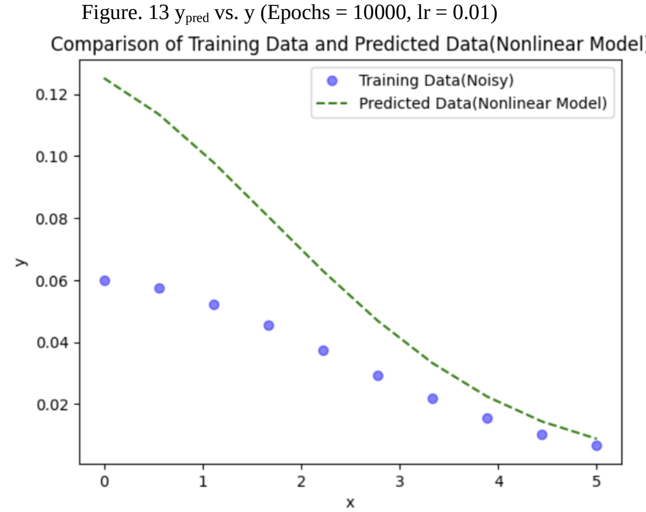

Figure. 14 y_{pred} vs. y (Epochs = 10000, lr = 0.0001)

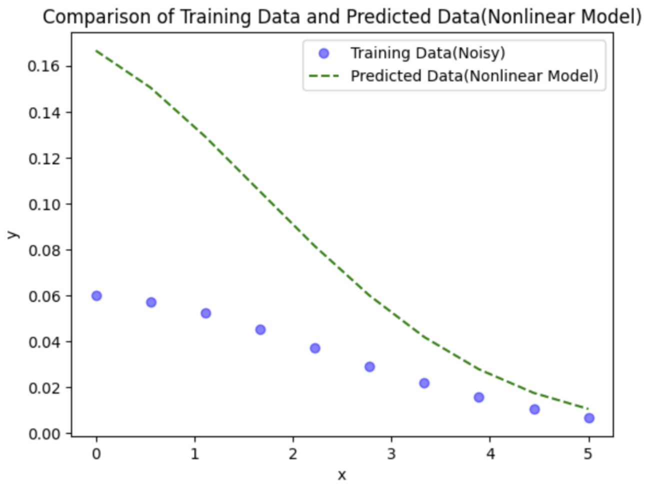

Figure. 15 y_{pred} vs. y (Epochs = 10000, lr = 0.00001)

So as the learning rate decreases, the results fail to converge, leading to higher loss. Unlike linear models, increasing the learning rate also causes the results to lose convergence. In extreme cases, a high learning rate can result in overshooting predictions and significant loss.

Therefore, the best combination that minimizes the loss for the nonlinear model are epochs = 10000 and a learning rate of 0.001.

IV. Conclusion

The performance of the two models differs significantly. In the linear model, convergence improves with more epochs, while in the nonlinear model, excessive epochs can lead to overshooting and higher loss. Similarly, reducing the learning rate decreases convergence in both models. However, in the nonlinear model, a learning rate that is too high also causes overshooting, further hindering convergence. Through these cases, the best settings that minimize the loss were determined to be Epochs = 1000 and Learning Rate = 0.01.

References

[1] M. K. Jawed and S. Lim, “Discrete simulation of slender structures,” *BruinLearn*, 2024.