Final Report

Leopold Marx
CS 5350

April 16, 2019

I, Leopold Marx, have worked alone on the final project in CS 5350.

1 Introduction

The Lorenz System is a system of ordinary differential equations that generate a butterfly-like figure in the domain $\mathbb{R}^3$ over time. These are the sets of differential equations:

$$\frac{dx}{dt} = \sigma(y - x)$$
$$\frac{dy}{dt} = x(\rho - z) - y$$
$$\frac{dz}{dt} = xy - \beta z.$$

Iterating this differential equation with parameters of $(\sigma, \rho, \beta) = (10, 28, \frac{8}{3})$ produces the iconic butterfly figure as shown below in Figure 1[1]. This figure is defined as the Lorenz Attractor. The Lorenz Attractor is often viewed as a model for atmospheric convection and is often used for a temporary replacement for real atmospheric data. In the RISE Research paper[2], they discuss how to predict the number of orbits in each regime before switching to the other.

Often in atmospheric sciences, one only has a fraction of the true model, and a forecast needs to take place to predict the remaining portions of the graph. The importance of forecasting in the real world is influential and applies to everyone's lives. Some common applications are standard weather forecasting, predicting trajectory, intensity and time predictions of hurricanes, and verification that it is safe enough for an airplane to fly over storms.

My plan for this project is to linearly separate the data in half to produce the best prediction on z, given variables x and y. I will be using Gaussian Process Regression to predict z, as it produces confidence intervals and is a strong learner.

[1] https://matplotlib.org/examples/mplot3d/lorenz_attractor.html
[2] Evans et al. (2003)

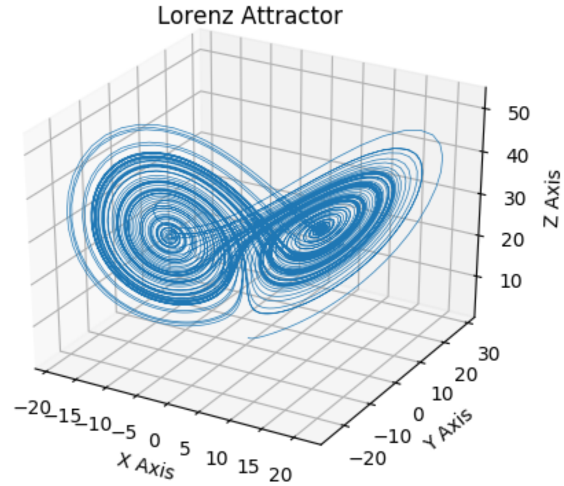


Figure 1: Lorenz system illustrated with parameters $(\sigma, \rho, \beta) = (10, 28, \frac{8}{3})$.

To further extend forecasting, I aim to find the optimal plane that separates the two butterfly halves that minimizes root mean squared error of the test data set. The plane will use the following equation with a and b as hyper-parameters for tuning:

$$ax + by = 0.$$

My motivation to use machine learning techniques for this project is mostly driven by getting a solution faster, rather than by mathematically computing a perfect solution. Moreover, the answer can simply be approximated. Regression is an extremely powerful tool that is excellent for approximating and generalizing data.

2 What have I done to reach my goal?

I have collected 4096 data points of the Lorenz Attractor. I will use this data to make my predictions. I have predicted x, given y and z, with random examples of testing and training data points and illustrated the prediction surface shown in Figure 2. I also generated a figure with the most error prone areas of this same prediction, which is illustrated in Figure 3. I created a forecasting model for predicting on z with three different splits, splitting perfectly through the two butterfly wings, splitting between the average of the minimum and maximum x values, and lastly splitting between the average of the minimum and maximum y values. Each are shown in Figures 2, 3, and 4 respectively. The perfect split of the data occurs at the plane $2y - x = 0$.

(a) Prediction Surface

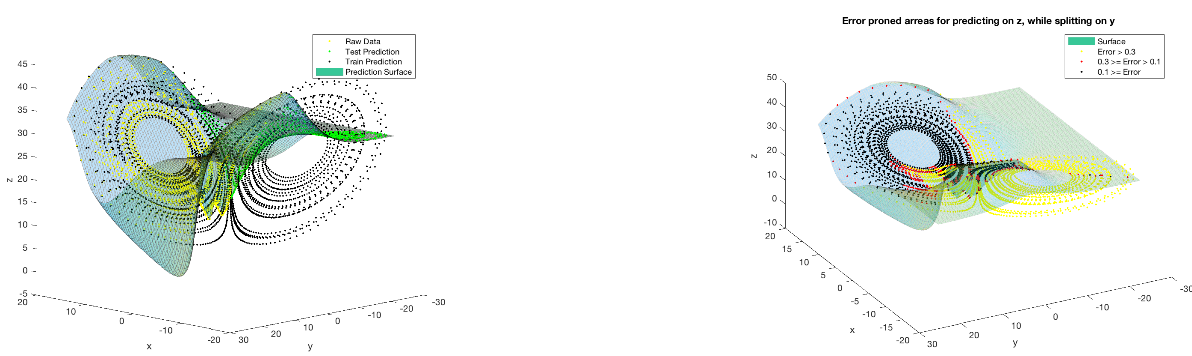
(b) Illustration of Errors

Figure 2: Perfect Split.

(a) Prediction Surface

(b) Illustration of Errors

Figure 3: $x = 0$ Split.



(a) Prediction Surface

(b) Illustration of Errors

Figure 4: $y = 0$ Split.

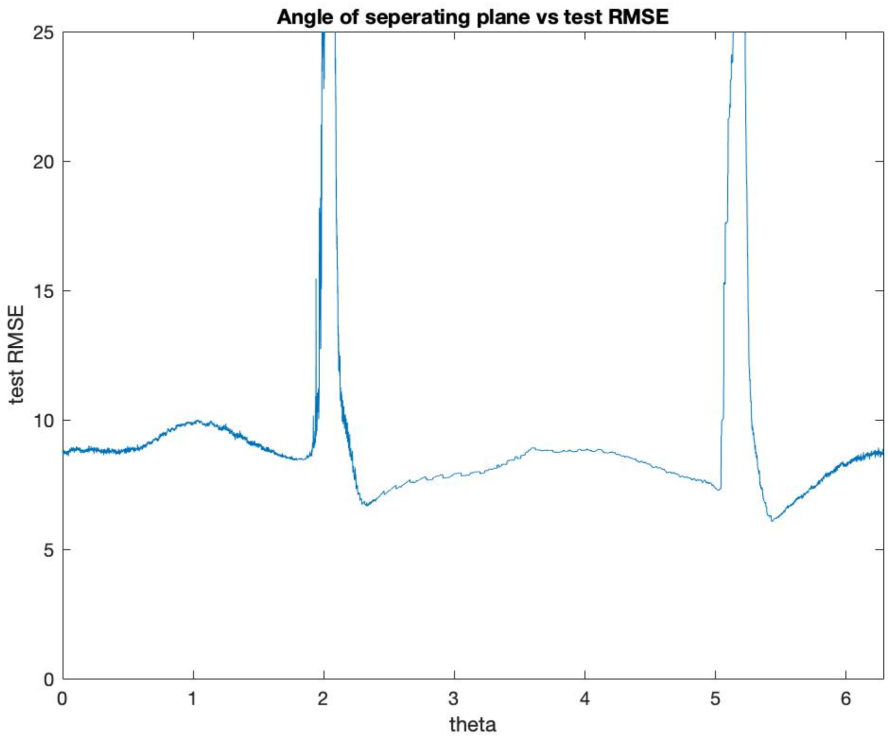


Figure 5: Test RMSE vs angle of separating plane.

3 Solution

Since values of a and b can result in the same plane with just a different scale in the equation $ax + by = 0$ (e.g. $a = 1$, $b = 2$ and $a' = 2$, $b' = 4$ result in the same plane), it is useful to consider the angle, θ, of the normal vector of the dividing hyper-plane. In fact, we can think of $a = sin(\theta)$ and $b = cos(\theta)$ in $ax + by = 0$ as the corresponding normal vector. Figure 5 shows the test RMSE values iterating over angles 0 through 2π.

In Figure 5, we can see the minimum test RMSE is approximately 6.09 at $\theta = 5.44$. This results in an equivalent value of $a = -0.7468$ and $b = 0.6651$. Figure 6 illustrates the linear split with these specific values of a and b.

4 Experimental Evaluation

Firstly, lets further look at Figure 5, as how the angle relates to RMSE. We see there are two periods in the figure. This is due to the hyperplane flipping around by π radians. In fact, this makes the training data now the test data and vise versa. The figure also illustrates the RMSE values of one of the periods is slightly larger than the other. This is simply explained by not having symmetric data. One regime has more data points than the other causing a stronger fit towards that half of the Lorenz Attractor, which causes worse test RMSE values for the regime with less data.

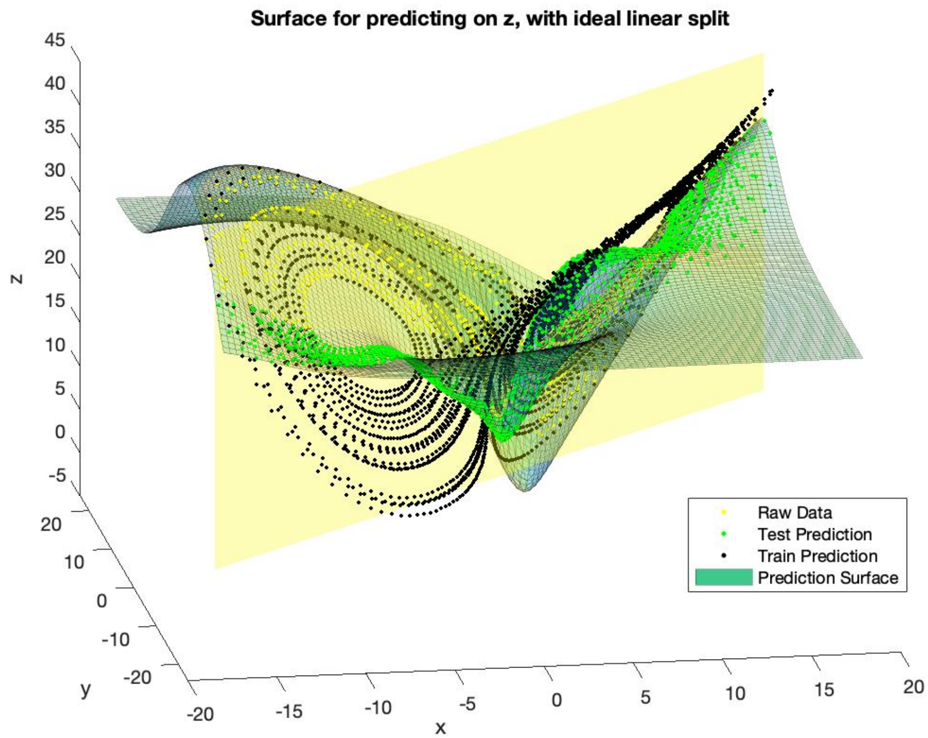


Figure 6: Illustration of the best fitting hyperplane for forecasting.

In Figure 6, we can see how the hyperplane does not perfectly split the regimes in half, but instead bisects each regime, resulting in the training data to have data in both halves.

This compromise of a split can be applied to Occam's Razor by having a simple linear split with data from each regime can result in the lowest RMSE values. This is a simple solution to the optimization problem that is "good enough" for a practical solution.

5 Github

See https://github.com/leopoldmarx/CS5350_Final_Report

6 References

1. https://matplotlib.org/examples/mplot3d/lorenz_attractor.html

2. Evans, Erin Bhatti, Nadia Kinney, Jacki Pann, Lisa Peña, Malaquias Yang, Shu-Chih Kalnay, Eugenia. (2003). RISE: Undergraduates Find That Regime Changes in Lorenz's Model are Predictable. *Bulletin of The American Meteorological Society - BULL AMER METEOROL SOC.* **85**, 10.1175/BAMS-85-4-520.