Advanced deep learning - Assignment 5: Rössler attractor

Philibert de Broglie - Quentin Nativel - Marine Sobas

12 March 2020

1 Introduction

In this exercise, we aimed at modelling the data generated by the Rössler attractor. We developped the discrete approach. So given the linearized system: $w_{t+t} = Jw_t$, our objective is to estimate the Jacobian J. The quality of the model will then be assessed thanks to statistics and physical quantities.

2 Model

We decided to use the full state and to model the problem as a discrete system, with a $\delta_t = 10^{-2}s$.

$$W_{k+1} = NN(W_k), \forall k \in N \tag{1}$$

2.1 Architecture of the neural network

For the neural network, we decided to take a model with two hidden layers of 10 neurons and to apply ReLU activation functions. This relatively straightforward model was chosen because we aim at estimating J, which is a 3x3 matrix. As the amount of data generated with a $\delta_t = 10^{-2}s$ was limited, we preferred to keep the model simple. Thus, we prevent the model from excessively overfitting the data.

2.2 Loss function

To get relevant results, we first used the Mean Square Error Loss. The nature of the problem, namely a regression, accounts for the choice of this function. Given $\hat{w}$, the predicted trajectory and t_f, the ending time ($t_f = 100,000$), we calculate :

$$MSE(w, \hat{w}) = \sum_{t=1}^{t_f} (w_t - \hat{w}_t)^2$$

More compelling though, a custom loss function was computed to add a derivative term into the loss function. Thanks to finite difference coefficients, we estimated the derivatives of the trajectory. The formula is as follow :

$$\frac{\delta}{\delta t} w_t \approx \frac{1}{12} \times w_{t-2} - \frac{2}{3} w_{t-1} + \frac{2}{3} w_{t+1} - \frac{1}{12} w_{t+2} = D_t$$

Similarly, we approximated the derivative of the predicted trajectory :

$$\frac{\delta}{\delta t} \hat{w}_t \approx \frac{1}{12} \times \hat{w}_{t-2} - \frac{2}{3} \hat{w}_{t-1} + \frac{2}{3} \hat{w}_{t+1} - \frac{1}{12} \hat{w}_{t+2} = \hat{D}_t$$

Thus, we determined a derivative term to improve the loss function :

$$DVT(w, \hat{w}) = MSE(D, \hat{D}) = \sum_{t=1}^{t_f} (D_t - \hat{D}_t)^2$$

Eventually, the selected loss function is the following, for $\alpha \in [0, 1]$:

$$L(w, \hat{w}) = \alpha MSE(w, \hat{w}) + (1 - \alpha) \times DVT(w, \hat{w}).$$

Based on the quality of the predictions and the statistics which will be introduced later on, we decided to take an $\alpha = 0.5$.

2.3 Training

The model was trained on a trajectory computed with the real Jacobian using the *RosslerMap* class. The training trajectory lasts for 1,000s and is therefore composed of 100,000 points. It is plotted in figure 1.

3 Results and statistics

To assess the quality of the prediction, not only did we compare the trajectories, but we also computed relevant statistics. It is necessary to evaluate whether the prediction behaves in the same way as the real trajectory.

3.1 Predicted trajectory

With the model which was previously described, we obtained the trajectory plotted in figure 2.

We notice that even if the amplitudes of the trajectories are different, they seem to follow similar behaviours. The equilibrium point, which is the center of the pseudo-circle, is also observed in the predicted trajectory. At first sight, the predicted trajectory is akin to the one which was generated by Rössler map.

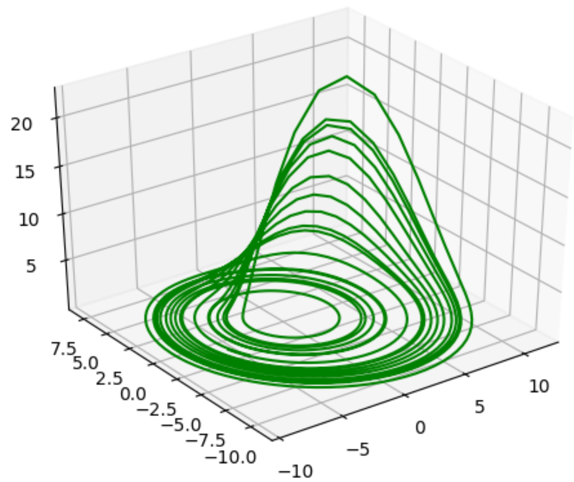


Figure 1: Real trajectory, $\delta_t = 10^{-2}s$

3.2 Probability density functions

Are the probabilities to reach a given location comparable? To compare the dynamics of the true trajectory to our predictions, we plotted histograms as a function of the x axis. The results are shown in figure 3.

From figure 3, we observe that the distributions of the two trajectories are globally similar. The real trajectory in blue is more spread than the predicted one in orange. It is consistent with our previous observations : the amplitudes of the true trajectory are larger in figure 2. The two maximums which represent the most likely x-coordinates, are almost identical. The equilibrium point, which seems to be reached for x=0 corresponds to a probability of 0.05 for the true trajectory while it is 0.06 for the prediction. They both correspond to a local minimum.

3.3 Time correlations

To have a more accurate comparison between the dynamics of the trajectories, we computed time correlations between (w_t, w_{t+T}). For a global view, we computed it for $T = 1000 \times \delta_t$. Results are shown in figure 4. It appears that time correlations are quite different, but some trends are grasped by the prediction. For instance, the correlations from (t=5000,T=60000) to (t=8000,T=90000) seem to be pretty consistent. From (t=2000,T=3000) to (t=5000,T=60000), the global trend is predicted but the local maximum at (t=4000,T=5000) was

Comparison between the predicted and the real trajectories

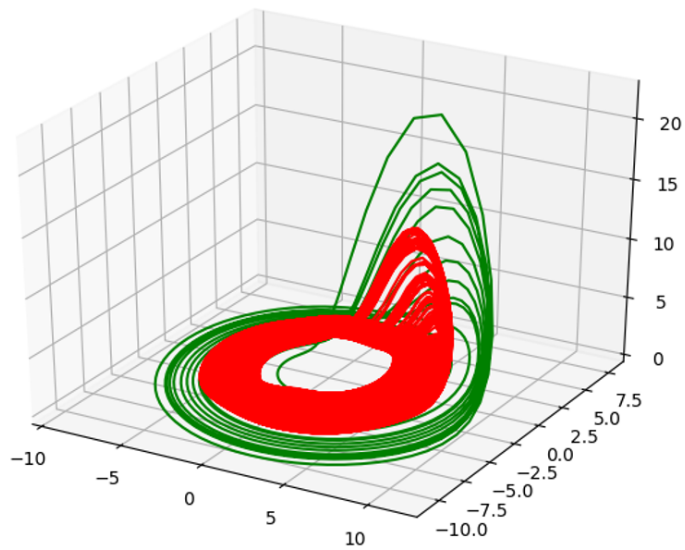


Figure 2: The predicted trajectory (red) is alike to the real trajectory (green) in terms of behaviour. Similar patterns are observed yet it is noteworthy to mention that their amplitudes are different.

not correctly reproduced.

3.4 Frequencies

To compare the patterns followed by the trajectories, we computed their frequencies using discrete Fourier transforms. As a matter of fact, the system globally follows a loop so it is rather relevant to compare frequencies. In figure 5 we compared those frequencies for same-size trajectories. The first one was extracted from the beginning of the global trajectory : [0,1000] while the second one takes place halfway through : [4000,5500]. Frequencies from the beginning are closer than they are in the second phase. So the dynamics are more similar in the beginning, which is consistent with the Lyapanov exponent logic : it is more complicated to predict a trajectory after a while. For a shorter time range, $t \in [0, 5]$, we compared the top 20 frequencies of the two trajectories. It appears that 18 out 20 top frequencies were the same. So when short trajectories are considered, trajectories seem to be quite coherent.

Probability density function : how likely is it to reach a specific x-coordinate?

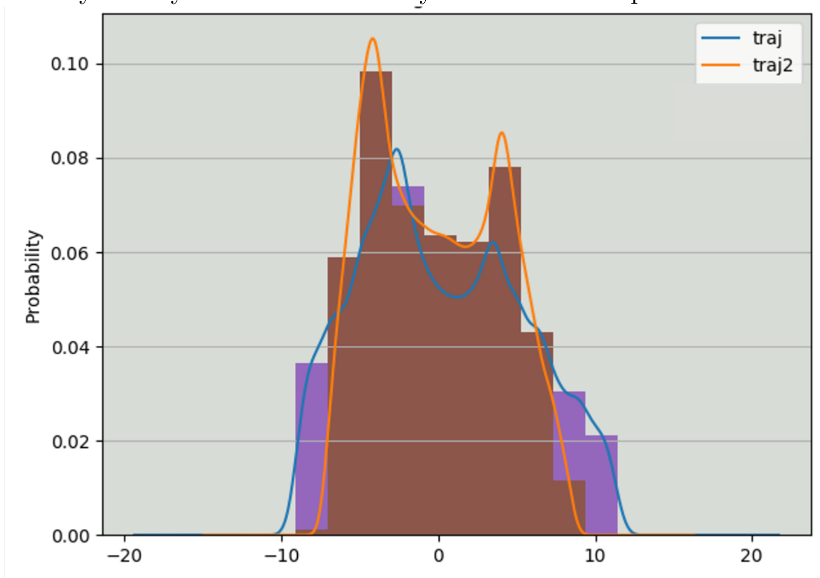


Figure 3: The real (blue) and the predicted (orange) trajectories follow analogous distributions, even if for the prediction, some x-coordinates are too frequent.

4 Conclusion

All in all, we managed to predict a chaotic system using a rather straightforward model. Statistics reveal that the obtained model was quite relevant. By using a more complex model and integrating physical properties into the model, we may get even more accurate results.

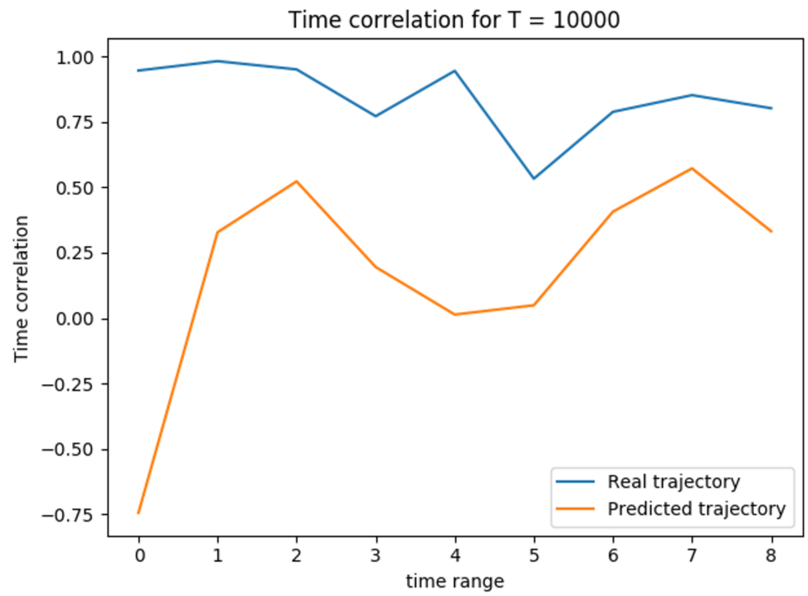
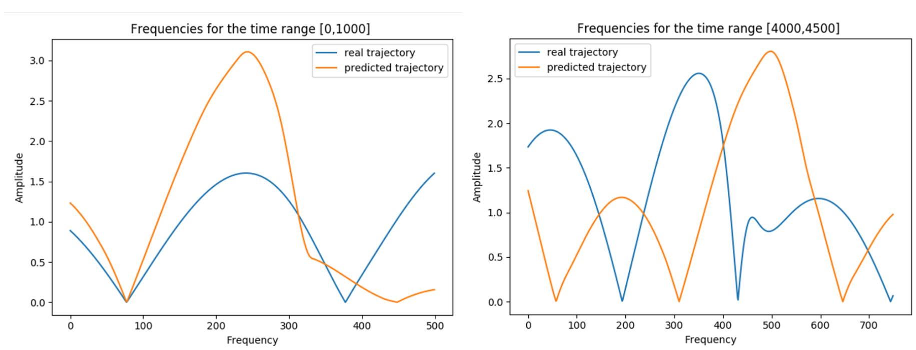


Figure 4: Time correlation for the real (blue) and predicted (orange) trajectories



(a) Frequencies between t=0 and t=1,000

(b) Between t=4,500 and t=5,500

Figure 5: Comparison between frequencies for different time ranges