MVA: Deep learning in practice (2020/2021) **Homework 5**

TP5 : Rossler System

Autors: *Charbel-Raphaël Segerie, Hugo Laurençon* *(March 2021)*

Solution by Charbel-Raphaël Segerie, Hugo Laurençon

Approach Selected
We chose the third approach, first with the Net ODE, then finally without the Net ODE.

1 Design

1.1 Design of the loss function

After building a trajectory, we take a batchsize of 50,000 points at random on the trajectory and look for each of these 50,000 points 1 step into the future. We then calculate by finite difference the velocity vector. Then we compare the simulation of the velocity vector with the groundtruth with a function of the distance L1 in R3. The absolute value allows to train the network with the same efficiency for all scales of different sizes, while the L2 norm is numerically unstable because of the difference of order of magnitude between very small distances and large distances. To this norm, we also add a regularization norm: the Jacobian Frobenius norm, taking care that the Frobenius norm is of the same order of magnitude as the speed error norm.

1.2 Design of the network

We chose a three-layer network. We use an encoding-decoding type structure with 3 dimensions in input and output and 10 dimensions in embedding. So we have the following structure:

- nn.Linear(3, 10)
- nn.Tanh(),
- nn.Linear(10, 10),
- nn.Tanh(),
- nn.Linear(10, 3)

We used hyperbolic Tanh activation to have a continuous and differentiable derivative output at any point. The network is not very deep and therefore no need to worry about the vanishing gradient problem. The ReLu activation would have made the singularity points appear apparent, which would have been undesirable for a system that wanted to be smooth like Rossler's one. We didn't need batch normalization because the network inputs and outputs are already relatively well normalized.

2 Estimating the quality of the dynamical system learned

2.1 By statistics

Figure 1 shows the different system statistics over a 10,000 second run with a delta t of 0.01 seconds. On figure 1 from left to right and from top to bottom we can see :

- The first 5 seconds of the evolution of the system according to the coordinates x, y, and z. We can see that the curves overlap quite well until the divergence appears visible around the 30th second.

- The phase portrait $y = f(x)$: We can see that the two systems overlap almost perfectly: the general shape of the system is respected. Only the 3D reconstruction of the model shows a slight flaw: the original Rossler system in green climbs slightly higher on the z scale.
- On the three figures of the second line, we can observe the histograms of the distributions of the variables x, y, and z: We can see that the distributions overlap well if we disregard the slight statistical noise.
- On the third row of figures, we can observe the temporal correlations between the variables x(t), y(t), and z(t), and their temporal translation shifted by T = 17.08 seconds, which corresponds on the first figure to a period of the system x. The value of 17.08 seconds was obtained from the Fourier analysis of the time series. It can be seen that the variables x, and, y are highly correlated. We can also see that variable z is not highly correlated, which is due to the fact that if variables x and y are highly periodic, variable z "jumps" from one loop to the other stochastically.
- On the last row of the figure, we can see the different Fourier transforms. After extraction of the main components we find the following main frequencies (we do not count the zero frequency due to the average of the system):
  - Principal frequency true x: 1708
  - Principal frequency true y: 1708
  - Principal frequency true z: 1708
  - Principal frequency pred x: 1693
  - Principal frequency pred y: 1694
  - Principal frequency pred z: 1693

2.2 Estimating the quality (bis), with physical quantities

Since we chose approach three, it is easy to get the Jacobian from our network. We can therefore calculate the equilibrium point and its error with the equilibrium point of the original system. We find :
Equilibrium state : [-0.08080456 -0.05568678 0.04885566] , error : 0.0912

Lyapunov exponents can also be calculated. We find:
Lyapunov Exponents : [0.07367473 0.00338464 -3.14234273] with delta t = 0.01

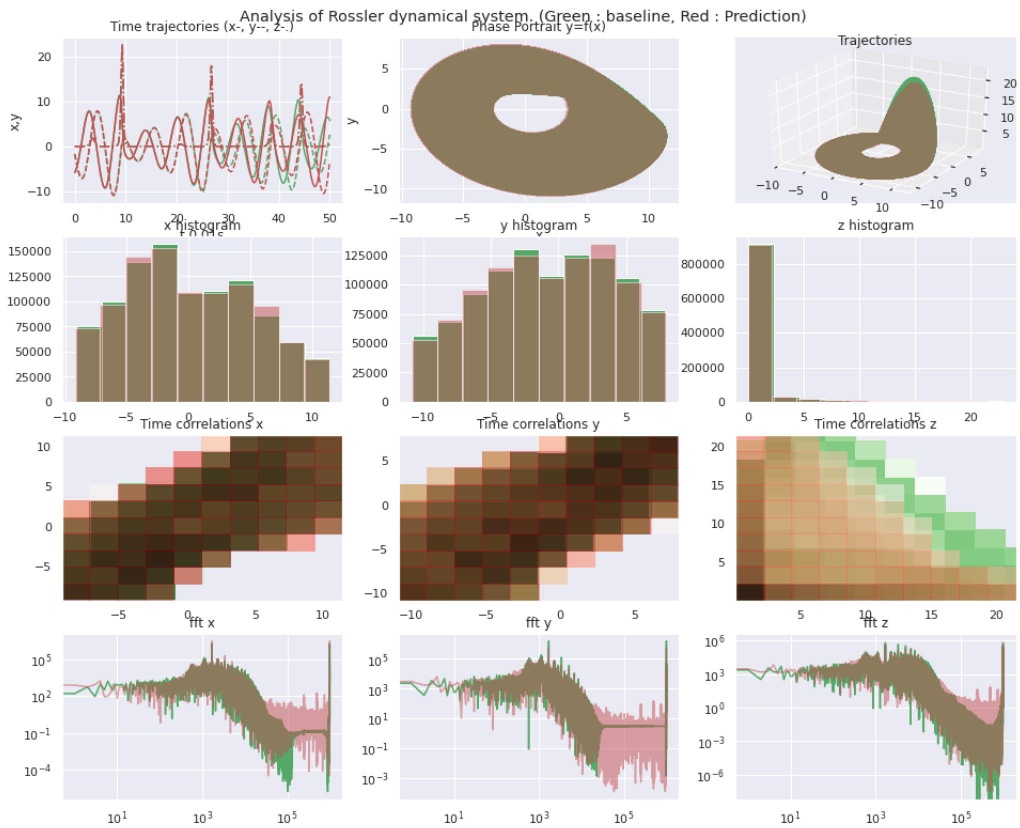

Figure 1: Results of the long term statistics