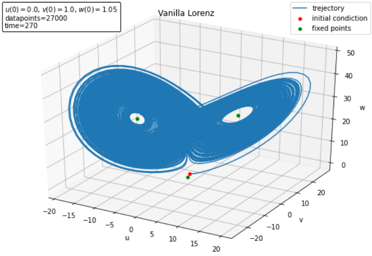
Lorenz Model and Information

February 9, 2022

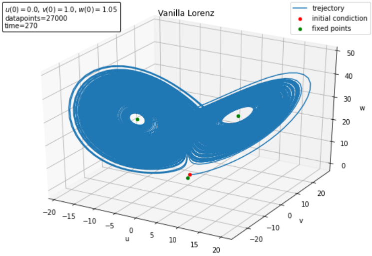
Lorenz Model 63

- goal: learn its trajectories with an LSTM.



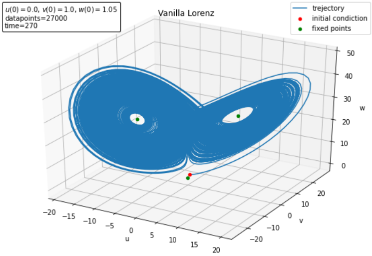
Lorenz Model 63

- goal: learn its trajectories with an LSTM.
- make a dataset of different type of trajectories.



Lorenz Model 63

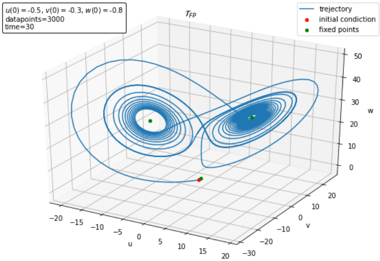
- goal: learn its trajectories with an LSTM.
- make a dataset of different type of trajectories.
- evaluate model's performance.
- different type of trajectories $\rightarrow$ different performances : why ?

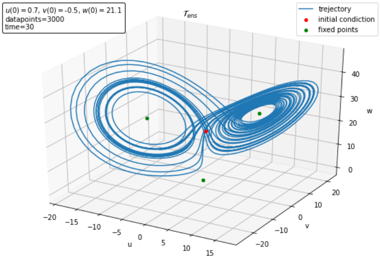


Lorenz Dataset

The dataset is divided in mainly two parts.

- "manifold" $\rightarrow$ trajectories close to fixed points

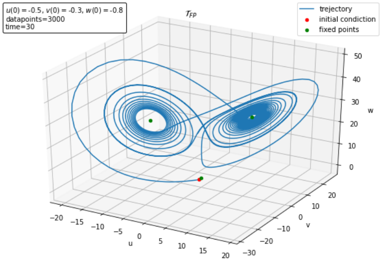


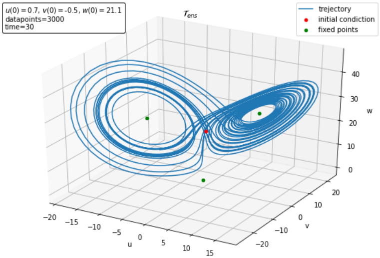


Lorenz Dataset

The dataset is divided in mainly two parts.

- "manifold" $\rightarrow$ trajectories close to fixed points
- "part" $\rightarrow$ trajectories sampled from a longer one

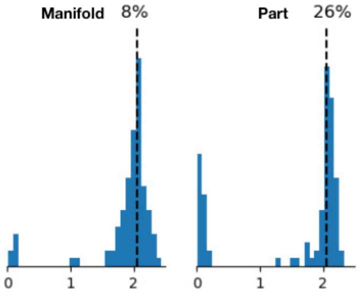




Performance of different Dataset

The performance of the LSTM trained on these datasets is here evaluated.

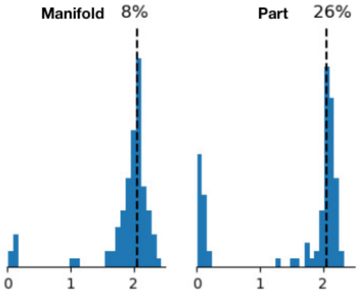
- d_2-dimension performance metric on the predicted model.



Performance of different Dataset

The performance of the LSTM trained on these datasets is here evaluated.

- d_2-dimension performance metric on the predicted model.
- The model trained with the "manifold" data has higher performance than "part".



Why this difference of performance ?

The goal now is to understand why we find this difference in performance.

- Find a measure of the **information** that each trajectory contains.

$$
\begin{aligned}
\mathcal{H}_{svd} &= -\sum_{i=1}^{N} \bar{\sigma}_i \log \bar{\sigma}_i \\
\mathcal{I}_f &= \sum_{i=1}^{N-1} \frac{(\bar{\sigma}_{i+1} - \bar{\sigma}_i)^2}{\bar{\sigma}_{i-1}}
\end{aligned}
\tag{1}
$$

Where σ_i are eigenvalues of the embedded trajectory.

Why this difference of performance ?

The goal now is to understand why we find this difference in performance.

- Find a measure of the **information** that each trajectory contains.
- Fisher Information and svd entropy.

$$\begin{aligned} \mathcal{H}_{svd} &= -\sum_{i=1}^{N} \bar{\sigma}_i \log \bar{\sigma}_i \\ \mathcal{I}_f &= \sum_{i=1}^{N-1} \frac{(\bar{\sigma}_{i+1} - \bar{\sigma}_i)^2}{\bar{\sigma}_{i-1}} \end{aligned} \tag{1}$$

Where σ_i are eigenvalues of the embedded trajectory.

Why this difference of performance ?

The goal now is to understand why we find this difference in performance.

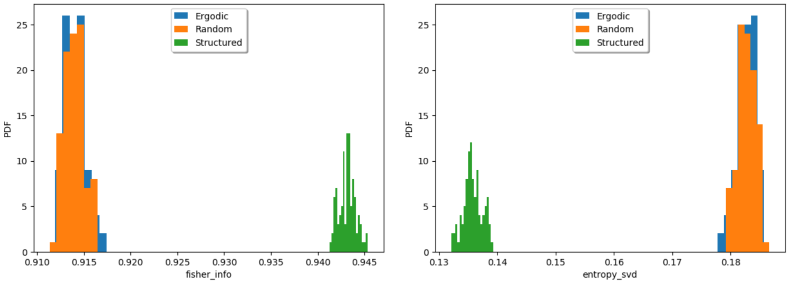
- Find a measure of the **information** that each trajectory contains.
- Fisher Information and svd entropy.
- Fisher Matrix

$$\begin{aligned} \mathcal{H}_{svd} &= -\sum_{i=1}^{N} \bar{\sigma}_i \log \bar{\sigma}_i \\ \mathcal{I}_f &= \sum_{i=1}^{N-1} \frac{(\bar{\sigma}_{i+1} - \bar{\sigma}_i)^2}{\bar{\sigma}_{i-1}} \end{aligned} \tag{1}$$

Where σ_i are eigenvalues of the embedded trajectory.

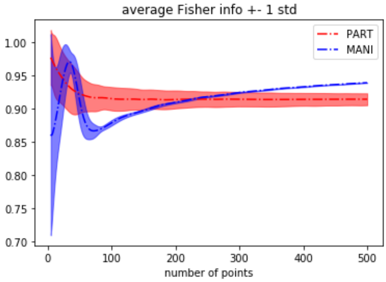
Results

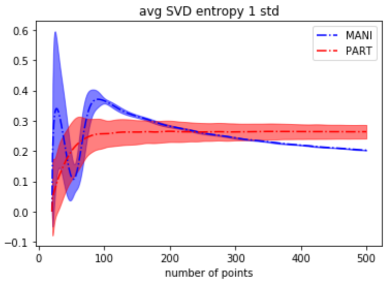
The histogram in green represents the dataset "mani". As we can see they show to have more information (less entropy).



Results

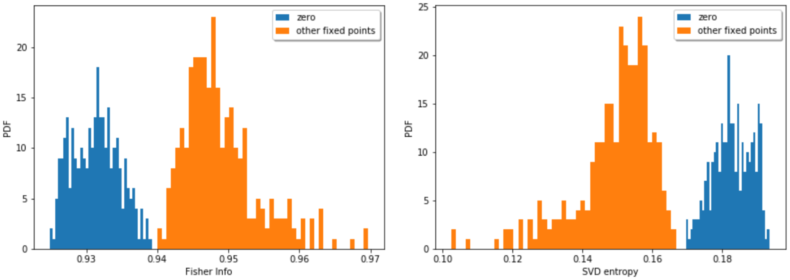
We can also analyze how the svd entropy, and the information, grow as functions of datapoints.





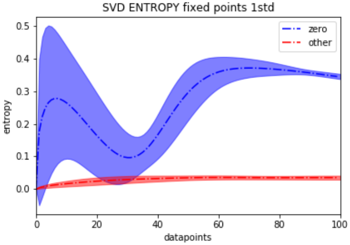
Is zero the most informative point ?

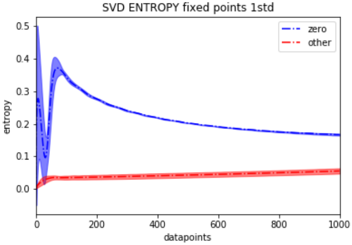
Now that we have measured that the trajectories that start close to the fixed points are the most informative, we can use the same approach to know what is the most informative point.



Is zero the most informative point ?

Now that we have measured that the trajectories that start close to the fixed points are the most informative, we can use the same approach to know what is the most informative point.





Fisher Matrix Approach

To understand better how much information our model extracts from the data we can use the Fisher Matrix. It describes the shape of likelihood function.

$$I(\boldsymbol{\theta}) := \mathbb{E}_{\widehat{Y}}\left[\left(\boldsymbol{\nabla}_{\boldsymbol{\theta}} \log f(\widehat{\mathbf{y}};\boldsymbol{\theta})\right)^2\right] \quad (2)$$

Where $f(\hat{\mathbf{y}};\theta)$ the density function of a random variable $\hat{\mathbf{Y}}$ parametrized by θ.