Project-VAMPnets

February 25, 2020

Luca Pasquali, Andreas Mardt, Tim Hempel, Frank Noé

- Using mdshare, fetch the data from the remote file '5d-folding.npy'. This is a system consisting of a single point moving in a 5-dimensional space, but whose behaviour depends only on its euclidean norm, or distance of the point from the center
- Build a network that extracts the main slow process(es) in the system, and identify the corresponding timescale value. This requires you to start from a higher number of output nodes, study the result, then go lower until all the timescales are properly estimated
- Add a bottleneck with y nodes, followed by a layer with 10 nodes. Find the minimum value of y that allows you to recover the main timescale correctly. Is the dimensionality of the bottleneck relate to the description of the system?
- Plot the correlation between the distance of your system from the center and the outputs of the bottleneck layer in a scatter plot. *(hint: you can create a model using the bottleneck output. If the bottleneck layer is already compiled and trained in another model, this model doesn't need to be compiled or fitted but can still predict and thus transform data)*
- Plot the correlation between the outputs of the bottleneck layer and the eigenfunction of the estimated Koopman operator. You can use the function `estimate_koopman_op(transformed_data, tau)` present in the file `helper_vampnets.py` located in the **notebooks** folder of the tutorial's Github repository