Critical Paper Review for ME 599: B. Lusch, J. N. Kutz, and S. L. Brunton, "Deep learning for universal linear embeddings ofnonlinear dynamics," arXiv:1712.09707, 2017

Sahil Kommalapati, 1926528
April 25, 2020

Abstract

I have read this paper multiple times over the last two months and I have come to the following conclusions: This paper is short yet communicates effectively. However, this work is targeted for an informed reader. It was tough to pierce through the technical jargon to get to the crux of the paper during my initial reads.

$$\nabla\phi.f = \lambda\phi \tag{1}$$

However, solving a few cases of the Koopman eigen functions analytically really helped me understand the big picture. In Eq.(1), ϕ refers to the Koopman eigen function and λ refers to the corresponding eigen value. f represents the non-linear dynamics of the system. Nevertheless, with the necessary pre-requisites, the paper is extremely well organized with succinct descriptions of their novel implementations.

1 Introduction

The goal of this paper is to identify the Koopman operator and its Eigen functions from data generated by a non-linear system. They can be understood as the coordinate transformation that makes the dynamics globally linear. A linearized system representation allows us to tap into the vast resources of linear stability and control theories. Hence, it is very desirable to find the Koopman eigen functions. This work takes on systems with continuous eigen spectrum to find the Koopman eigenfunctions. However, notably the framework also works for a discrete spectrum.

The contents of this paper are elegantly organized. It begins with an introduction to the relevance and historical context of the work and the tools which have been used, notably neural networks and Koopman Operators. The introduction also points out the contemporary work, notably by Mardt et al, Takeishi et al and Otto et al, who have used Deep Neural Network architectures to identify Koopman embeddings. This particular work addresses two key challenges in dynamical systems. System Identification from data and Low dimensional representation of high dimensional data by discovering latent states.

2 Auto Encoders for learning Koopman Eigen functions.

A discrete time system ($x(t)$) is assumed through out the paper. Snapshots of the data are collected along Δt increments. The appendix of the paper also shows the data collection techniques used for all the examples. The Koopman operator (K) is an infinite dimensional linear operator which advances the measurements of a system to the next time step. Identifying the Koopman eigen functions (ϕ) makes the framework interpretable which is really challenging while using deep learning models. So this work uses the interpretability of Koopman eigen functions and the capability of deep learning to handle huge data sets and avoid over fitting to achieve the goals.

The paper presents the classic example of a nonlinear pendulum to explain the ubiquitous continuous spectra observed in nature. As the energy of the pendulum increases, the frequency decreases continuously resulting in a continuous spectrum.

$$\ddot{x} = -sin(\omega x) \tag{2}$$

The crux of the paper is using an auxiliary, i.e. supplementary network to account for the parametric dependence of the Koopman operator on the continuously varying frequency ($\lambda = \pm\omega$). This auxiliary network identifies and parametrizes the continuous frequency, which then parametrizes the compact koopman operator, located between the encoder and the decoder. The auxiliary network is an addition to the auto-encoder network which is used to identify the Eigenfunctions. In general, the combined system serves for the following three purposes:

1. To identify the intrinsic coordinates (ϕs) on which the state dynamics ($x(t)$) evolve in a low dimension($y(t)$). The decoder aspect of the auto-encoder can be used to identify the inverse of these intrinsic coordinates (ϕ^{-1}) to get back to the original coordinates($x(t + \Delta t)$).

2. The dynamics of the low dimensional space are restricted to evolve linearly and thus we find the Koopman operator (K). This lets us define the coordinate transformations essentially as the eigenfunctions (ϕ) of the Koopman operator.
3. Finally, utilizing the Koopman eigen function and the operator, we can successfully predict the m time steps in the future of the time series data, from the k^{th} state. This is done by using the .(3).

$$x_{k+m} = \phi^{-1}(K^m(\phi(x_k))) \tag{3}$$

To account for the continuous spectrum, the eigen values of the K(λ) matrix are allowed to vary as a function of its dependent variable λ. Thus, this network allows the eigen values to vary across the phase space, yet providing only a small number of eigen functions.

3 Implementation of the Auto Encoder Framework

This framework is demonstrated on three different examples.

1. A simple system with two nonlinear set of equations (Eq.(4)). This system has discrete Eigen functions which can be analytically calculated by using (1). The analytical results and the neural network embedding were found to be in agreement. It validates that the framework can also handle the discrete case.

$$\dot{x_1} = \mu x_1; \dot{x_2} = \lambda(x_2 - \mu x_1^2) \tag{4}$$

2. The classical long-standing case of a non-linear pendulum with a continuous spectrum. A parametric ($\phi(\omega)$) complex conjugate pair of eigen functions is obtained. This simple result is quite significant, given the history of issues hovering around the problem of dealing with continuous spectrum. Notably, this example helped in discovering the constraint on radial symmetry in pendulum-like cases for effective convergence of the neural networks.
3. The final example is a benchmark case in physics dealing with High dimensional non-linear fluid flow. The equations for the low dimensional attractor were given by Noack et al. which were used to simulate the training data and further validation. This system is known to show a stable limit cycle corresponding to the von-Karman vortex shredding and an unstable limit cycle corresponding to the low-drag conditions. Both have been reconstructed by using the linearized Koopman operator with two eigen functions, solely from the initial conditions. The results are consistent with the previous literature.

I believe that re-simulating a benchmark is a crucial step in establishing the validity of any new framework. The initial simple example made things clear, the benchmark established the validity and the finally an age-old problem in dynamical system was finally solved by identifying the eigen functions of the Koopman operator for a non-linear pendulum. It was an interesting narrative to read.

4 Conclusions and Remarks

The shortcomings of their work were well identified by the authors, noting the general extrapolation failure in DNN implementations, the generalizability and the scaling issues. Nevertheless, this work could potentially help in making great strides in advancing nonlinear control and stability theory.

The framework of this implementation is designed to work with time series data of the system which is structured to enable Koopman eigen functions extraction from systems whose equations of motion are unknown. However, it was surprising to see that all the examples have used models with known governing equations to generate data, instead of using experimental (maybe PIV) data for training. I understand that this was executed to get a greater control over the training data. However, It would have been interesting to find some results from real experimental data and provide a comparison with Koopman Operator's reconstruction of the linearized and predicted future states.

Future directions could include work on systems with more complicated energy spectrums. This work could have a great impact in epidemiology, neuroscience and understanding turbulence which are particularly rich in data.