Institut für Mathematik
Fakultät II – Mathematik und Naturwissenschaften
Technische Universität Berlin

Master in Scientific Computing

Internship Report

Approximation of Koopman Operators By Randomly Chosen Basis Functions

Student:

Moussa Atwi

ID:

464241

Supervisor:

Dr. Marcus Weber

Institution:

Zuse Institute Berlin (ZIB)

Address:

Takustraße 7, 14195 Berlin

Duration:

01/09/2022 - 01/12/2022 (240 hours)

Contents

1 The Company

1.1 Description

Konrad Zuse, born in Berlin in 1910, is the namesake of Zuse Institute Berlin (ZIB). After completing school he studied civil engineering and architecture at the predecessor of today's Technical University Berlin. With the Z3, he developed the world's first freely programmable computer. In 2010, on the occasion of his 100th anniversary, numerous events all over Germany was commemorate his work. Konrad Zuse, known today as the inventor of the world's first freely programmable computer, is ZIB's namesake.

The Zuse Institute Berlin is a non-university research institution of the State of Berlin and an interdisciplinary for applied mathematics and data-intensive high-performance computing. Its research focuses on modeling, simulation, machine learning, optimization and on innovation in application areas with scientific cooperation partners from academia and industry.
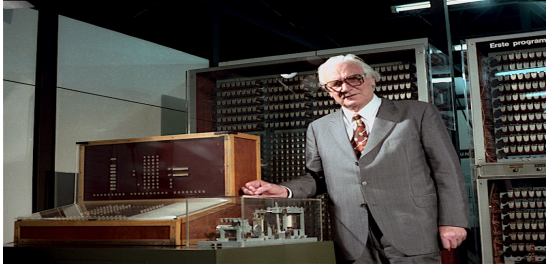

Figure 1: Konrad Zuse stands next to a replica of his Z3 computer at the Deutsches Museum in Munich.

Figure 2: The Zuse Institute Berlin located in the campus of Freie Universität Berlin in Dahlem, Berlin, Germany.

Regarding the research services and activities, ZIB operates Northern Germany's supercomputer HLRN and provides associated scientific support for users on planning and running large scale computing projects. ZIB also runs BRAIN, the high speed data net of Berlin's public institutions for science, education and culture, and provides data management and archive services. Furthermore, ZIB hosts the Cooperative Library Network Berlin-Brandenburg (KOBV), the common platform of all university libraries and public libraries and numerous research, special and government libraries in Berlin and Brandenburg, and the Research and Competence Center Digitalization Berlin (digiS), an institution to coordinate and support digitization efforts of cultural assets in the city state of Berlin. Last but not least, ZIB offers research software developed and maintained at ZIB, library services for the Campus Dahlem, and offers meeting facilities for scientific events.

1.2 Research areas and departments

Priority application areas are the life and materials sciences, logistics, infrastructure planning, and operations research. There are three research units or divisions that are divided into research departments as follows:

1. Parallel and Distributed Computing
   - Distributed Algorithms
   - Supercomputing

2. Mathematics of Complex Systems
   - Modelling and Simulation of Complex Processes
   - Visual and Data-Centric Computing

3. Mathematical Algorithmic Intelligence
   - AI in Society, Science, and Technology
   - Applied Algorithmic Intelligence Methods
   - Network Optimization

Furthermore, the departments are divided into working groups and their respective heads all you can find in the Institute part of the webpage. For instance, the department of Modeling and Simulation of Compex Processes is divided into five research groups

1. Computational Anatomy and Physiology
2. Computational Systems Biology
3. Computational Nano Optics
4. Computational Molecular Design
5. Computational Humanities

and it develops efficient modelling, simulation, and optimization tools and algorithms for challenging application problems from medicine, systems biology, and molecular dynamics as well as nano-optical systems and humanities. For these, the research group develop adaptive algorithms in the areas of finite elements, agent-based models, conformation dynamics, optimal control, and parameter estimation, exploiting multi-level structures for efficiency and robustness.

Finally, the training at ZIB has turned out to be an excellent environment for the academic growth of young researchers. Many former junior researchers of ZIB got offers for professorships and other leading academic position but for also as many others their time at ZIB has been the start for a career in industry and commerce, often as founders of their own spin-off companies.

1.3 My function at ZIB

I am working as a student research assistant for eighteen months from September 01, 2022 to March 29, 2024 in Dr. Marcus Weber research group which is the "Computational Molecular Design" group related to the department of "Modeling and Simulation of Complex Processes" and "Mathematics of Complex System" area. Indeed, my work is divided into two parts: an internship between September 01, 2022 and December 01, 2022 which is equivalent to 240 working hours and a master thesis afterwards. My aim task in the internship is to do an implementation of Extended dynamic mode decomposition (EDMD) using different randomly chosen basis functions, including randomized Gaussians and randomly-initialised neural networks. This requires a training on Python, implementing a data-drive methods, Dynamic mode decomposition (DMD) and EDMD with different basis functions to approximate the Koopman operator, learn the required software engineer tools (Linux, Conda, Git and Deeptime) and understand the basics of Koopman operator theory. The numerical experiments I will perform will create statistical series that will provide valuable insight into the relations between system dimensionality, intrinsic system dimensionality, and the number and smoothness of basis functions.

2   Activities and results

In the first six weeks, I was actually supervised by Dr. Andreas Bittracher (he left ZIB lately) and we used to meet once every two weeks with Dr. Marcus Weber to discuss the working plan. Afterwards, I used to work individually and ask my colleagues for particular issues. I was working almost five days a week (at ZIB) more than the required working hours.

2.1   EDMD with monomial basis functions

In the first four weeks, I was reading about Koopman operator theory and learn the numerical methods using special basis functions.

- **First Week:** An overview of Koopman theory, A practical introduction to the Koopman operator framework and Dynamic mode decomposition [1]. Extended Dynamic Mode Decomposition, Approximating the Koopman Operator and its Eigenfunctions, Computing the Koopman Modes and the Algorithm Summary [2].
- **Second Week:** Learning new software engineer skills, for instance installing Conda and deeptime on Linux and creating library extension to deeptime. As well as, taking Python tutorial, and understand the implementation of DMD and EDMD on Deeptime.
- **Third Week:** I Learnt about Molecular dynamics and Metastable sets to get an information about the existence of dominant eigenvalues of the associated transfer operator [3]. Implement from scratch the EDMD with different monomial basis functions and with Dyadic map that is defined as a piecewise linear function in order to approximate the eigenfunctions of the Koopman operator.
- **Fourth Week:** Finally, we experienced Stochastic EDMD on deeptime by using again the monomials up to order ten by considering a one dimensional triple well potential.

Result: We experienced by using the monomials an overfitting problem that leads to a bad approximation for the eigenfunctions.

2.2   Random Dynamic Mode Decomposition (RDMD)

In this activity, I am interested in combining EDMD with randomly chosen Gaussian functions (as basis functions) to a very high dimensional system to discuss how EDMD depends on the number of basis functions to approximate the operator. I would be interested in the dominant eigenfunctions (or dominant eigenvalues close to 1) to consider metastable regions where the system is stable. For this, consider the linear dynamical system described by the Brownian motion on d-dimensional Torus $\mathbb{T} := [-\pi, \pi]^d$ and the stochastic differential equation

$$\mathrm{d}X_t = \sigma \ \mathrm{d}W_t, \quad \text{where the variance} \quad \sigma = \mathrm{diag}\left(\sigma_1, \ldots, \sigma_d\right), \quad \sigma_i > 0. \tag{1}$$

W_t and X_t respectively denote the standard Brownian motion and stochastic process with slow diffusion in the first coordinate direction.

let $\mathbb{X}$ be the state space, and $\{x_1, \ldots, x_M\} \sim \mathcal{U}(\mathbb{X})$ be uniformly randomly distributed points from $\mathbb{X}$. Let $\rho > 0$ be a kernel bandwidth parameter. Consider the basis $\Psi := \{\psi_1, \ldots, \psi_M\}$, where

$$\psi_i(x) = \exp\left(-\frac{\|x_i - x\|^2}{\rho^2}\right).$$

denote the random Gaussian functions with random chosen centers.

- **First Week:** Create a new class on deeptime called Gaussian which is a new set of basis functions and I run EDMD using the Gaussian functions with random chosen centers to a one dimensional model.
- **Second Week:** Generate multi-dimensional data and run RDMD to analyze linear stochastic 2D models using Euler–Maruyama.
- **Third Week:** I tried EDMD in deeptime with monomials up to order 10 considering the one-dimensional Ornstein-Ulenbeck Process, given by an *Itô* stochastic differential equation and using Euler-Maruyama method that leads to a good approximation of the Koopman eigenvalues. Also, I tried the same process with the Gaussian basis functions (RDMD) in order to figure out the required number of basis functions for the best result [4].
- **Fourth Week:** For (1), I considered different values of σ_1 and σ_2 and change the number of Gaussians M in 1D and 2D.

First result: For $d = 2$, choosing $\sigma_1 = 0.01, \sigma_2 = 0.1, M = 58, \rho = 1.09$ and number of data points 25000, the diffusion in the second coordinate direction is much faster than in the first coordinate direction, eigenfunctions will be associated to the slowest direction and the first four dominant eigenvalues were almost 1. In the figure below, we observe that the first dominant eigenfunction is constant (≈ 1) and the second and third change signs, this shows the existence of two metastable regions that are attractive for the dynamics in the sense that the trajectories remain with such regions for long periods of time before exiting towards other metastable sets i.e., metastable regions required to study the long term behavior of the system. This result is due to many reasons, first the number of required basis fucntions (M = 58) and by the difference between σ_1 and σ_2, where in this case $\sigma_1 = 10 * \sigma_2$
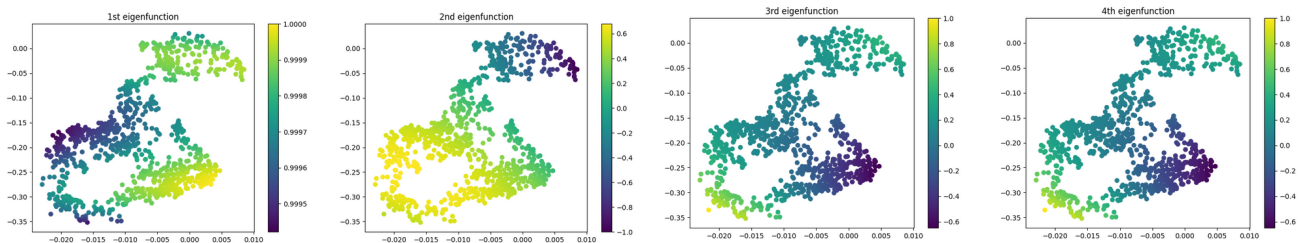

Also, using Euler Maruyama, I considered 2D triple well system for $\sigma = 1.09$ described by the stochastic differential equation (see link)

$$dX_t = \nabla V(X_t)dt + \sigma(t, X_t)dW_t.$$

For Koopman approximation, we selected $N = 25000$ random points inside the domain and integrate them for 1000 steps under an integration step $h = 10^{-5}$. It is already done on Deeptime using EDMD with the monomials up to degree 10. Now, I used EDMD with 45 Gaussians and $\rho = 1.038$ to obtain good approximation for the dominant eigenvalues associated to the processes in the system, close to the result below.

Second result: For $\rho = 1.05$ and $M = 56$, I obtained a very good result with eight real (positive) dominant eigenvalues: $\lambda_1 = 1.0, \lambda_2 = 0.99, \lambda_3 = 0.93$ and $\lambda_4 = 0.68$ and visualize the first four eigenfunctions where the first one is almost constant as we expect (≈ -1) and the second and third show clearly existence of two metastable regions (usually the third region is hard to see).
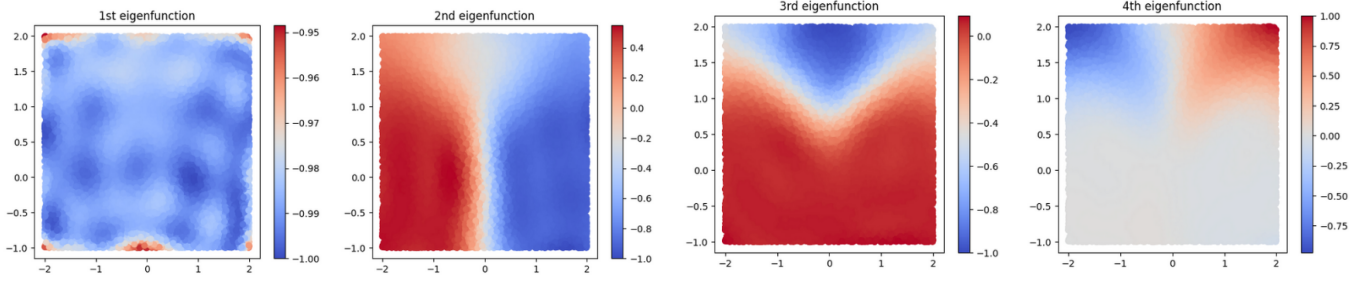
2.3 Randomized neural networks

In this activity, I would like to use randomly initialised neural networks as a basis functions for the approximation of the eigenfunctions using EDMD.

Let $\eta : \mathbb{R} \to \mathbb{R}$ be an activation function, e.g. the sigmoid or ReLU function and consider a fully-connected multi-layer neural network, e.g., a two-layer neural network of the form

$$\psi_i(x) = \eta\Big(\sum_{j=1}^{M_1} w_i^2(j)\eta\big(\sum_{k=1}^{d} w_i^1(j,k)x_k + b_i^1(j)\big) + b_i^2(1)\Big). \tag{2}$$

ψ_i is fully characterised by the hyperparameter $h_i = \big(w_i^1, w_i^2, b_i^1, b_i^2\big)$, where

$$w_i^1 \in \mathbb{R}^{M_1 \times d}, \quad w_i^2 \in \mathbb{R}^{M_1}, \quad b_i^1 \in \mathbb{R}^{M_1}, \quad b_i^2 \in \mathbb{R}.$$

Now choose M hyperparameters randomly (from some distribution to be specified), and consider the corresponding basis $\Psi := \{\psi_1, \ldots, \psi_M\}$.

- **First Week:** I tried to optimize my implementations, writing the Gaussian class in a better way and learn about neural network.
- **Second Week:** I Learnt about Tensorflow and Keras and how to create the two-layer neural network (2) by taking TensorFlow course on Udemy. In addition to that, I have obtained new results for RDMD by just using few number of Gaussians.
- **Third Week:** I successfully optimized my implementations and I created two more classes "Genbasis" and "Rinn" related to general basis and randomly initialised neural network respectively. Indeed, Rinn class is a deep neural network, where I can add more hidden layers (not only two) and to finally use Genbasis to obtain the dictionary Ψ from Gaussian or Rinn classes.
- **Fourth Week:** For $N = 10000$ and $M = 8$, I chose the hyperparameters randomly from a uniform distribution. Also, I considered again the two dimensional triple-well potential and two-layer neural network (i.e., input, one hidden layer and output).

First result: I considered a two-layer neural network where the input layer has two neurons (input dimension is 2), the hidden layer has eight neurons and ReLU activation function and one neuron in the output layer with Sigmoid activation function, i.e., units $= [8, 1]$ and activations $= [\text{ReLU, Sigmoid}]$. I obtained a good result where the dominant eigenvalues $\lambda_1 = 1.0, \lambda_2 = 0.6, \lambda_3 = 0.6$ and $\lambda_4 = 0.3$ and the dominant eigenfunctions where the first one is almost 1.0 (constant) and the second one changes sign.
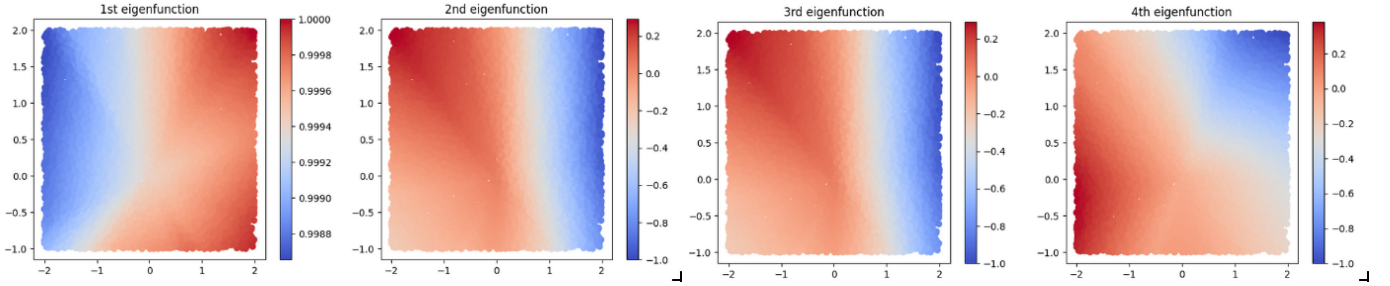
In addition to that, for (1), I considered two-layer neural network with randomly and uniformly distributed hyperparameters, where $M = 8, N = 10000, \sigma_1 = 0.01, \sigma_2 = 0.1$, units $= [2, 1]$ and activations $= [\text{ReLU, Sigmoid}]$.

Second result: I obtained the dominant eigenvalues $\lambda_1 = 1.0, \lambda_2 = 0.28, \lambda_3 = 0.17$ and $\lambda_4 = 0.23$ and the dominant eigenfucntions where the first ≈ 1.0 and the second and third change the sign.
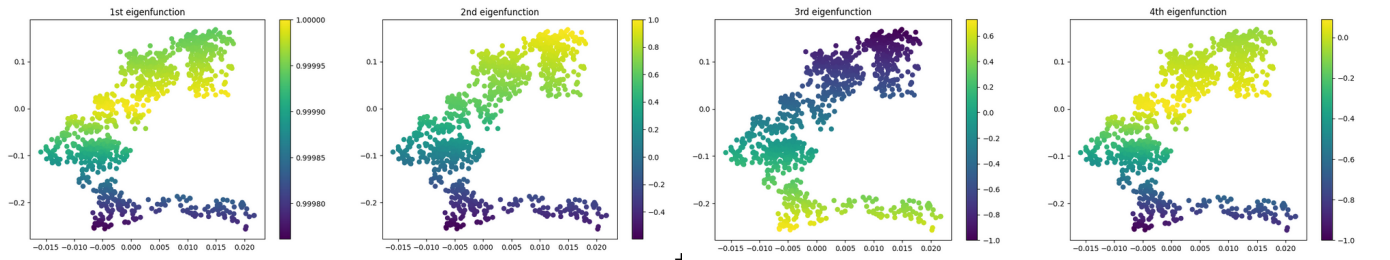
3   Conclusion and outlook

I learnt about the theory of Koopman operator and its approximation using EDMD with different basis functions. My contribution in this internship, is by creating three new classes (Gaussian, Rinn and Genbasis) using python and tensorflow to combine EDMD with

1. randomly chosen gaussian basis functions (by randomly chosen centers using Gaussian class) to approximate the eigenfunctions of the Koopman operator and I have obtained a very good results by just using 45 and 59 gaussians for two different problems.

2. randomly initialized neural network using Rinn class where I can randomly choose the hyperparameters (weight, bias, units, activation functions, ...) and consider more hidden layers (deep neural networks). For this, I obtained good results by just using 8 basis functions and two-layer neural network.

In conclusion, I would love to say, that I have learnt a lot theoretically and numerically during my internship. I made progress improving my practical and coding skills that has been my desire. Besides, the environment at ZIB is very helpful, where we can work freely and remotely as well. Also, working at ZIB allowing me to communicate with a large group of Ph.D. students and postdocs from different departments and groups.

Consequently, Dr. Marcus and I, agreed to continue by this topic for my master thesis at ZIB and try to go deeply through the theory of Koopman operator and also numerically by experimenting different basis functions and methods for better approximation.

Finally, I would like to thank Zuse institute Berlin and in particular Dr. Marcus Weber and Dr. Andreas Bittracher for the interesting topic and support throughout this internship.

References

[1] Brunton, Steven L. and Budišić, Marko and Kaiser, Eurika and Kutz, J. Nathan. Modern Koopman Theory for Dynamical Systems. arXiv, 2021

[2] Williams, M.O., Kevrekidis, I.G. & Rowley, C.W. A Data–Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition. J Nonlinear Sci 25, 1307–1346 (2015).

[3] Schütte, Christof and Sarich, Marco, Metastability and Markov state models in molecular dynamics, 24, 2013, American Mathematical Soc.

[4] Klus, S., Nüske, F., Koltai, P. et al. Data-Driven Model Reduction and Transfer Operator Approximation. J Nonlinear Sci 28, 985–1010 (2018).