Data-Driven Analysis of Dynamics and Koopman Operators

An Introduction

Christopher Sorg

Chair for Dynamics, Control, Machine Learning and Numerics
Department of Mathematics
Friedrich-Alexander-Universität Erlangen-Nürnberg
chr.sorg@fau.de

August 22, 2024

The Problem

Data-driven modelling: Starting from time dependent data, the goal is to build a model for simulation and/or prediction

- Data (e.g., position of particles, temperature, etc.) $\{(t_i, x_i)\}_{i=1}^n$ for time t_i comes from sensors
- The model has to be chosen within a certain class (ODEs, PDEs, SDEs, etc.)
- The coefficients of the model have to be optimized to explain the observed data

A Crucial Observation

Observation: Traditional approaches like Poincaré Maps to model dynamical systems often require nonlinear transformations, which can complicate analysis and predictions.

Intuition Behind Koopman Approach

The Koopman operator enables a **linear** representation of nonlinear dynamics by acting on observables, transforming the problem into an **infinite dimensional** Banach space, where functional analysis techniques can be employed.

Koopman Operators for Discrete Systems

For discrete-time systems, let $f : \mathcal{X} \to \mathbb{R}^d$ for a Banach space $\mathcal{X}$, called the state space, be a continuous map describing the system dynamics

$$x_{k+1} = f(x_k), \quad \forall k \in \mathbb{N}.$$

The Koopman operator $\mathcal{K}$ acts on a function $g \in C(\mathbb{R}^d)$ (often called observable) as

$$(\mathcal{K}g)(x) = g(f(x)).$$

This operator is linear, enabling the prediction of future values of observables.

Koopman Operators for Continuous Systems

Generalization for continuous-time systems is straight forward and the Koopman operator $\mathcal{K}_t$ is then defined as

$$(\mathcal{K}_t g)(x) = g(f_t(x))$$

where f_t is the flow map associated with the ODE.

Properties of the Continuous Koopman Operator

- **Semigroup Property:**
  The continuous Koopman operator $\mathcal{K}_t$ forms a semigroup. This semigroup property implies $\mathcal{K}_{t+s} = \mathcal{K}_t \mathcal{K}_s$ for all $t, s \geq 0$.
- **Generator:**
  The generator $\mathcal{L}$ of the Koopman semigroup is given by:

$$\mathcal{L}g = \lim_{t \to 0^+} \frac{\mathcal{K}_t g - g}{t}$$

  The generator captures the infinitesimal evolution of observables. For the dynamical system $\dot{x} = f(x)$ it is $\mathcal{L} = f \cdot \nabla$.

Predicting Observables Using Koopman Eigenfunctions

Knowledge of the spectral properties of the Koopman operator can allow prediction of the future value of an observable. Suppose ϕ is an eigenfunction of $\mathcal{K}$ with eigenvalue λ. Then:

$$\phi(x_{k+l}) = \lambda^l \phi(x_k) \quad \forall l \in \mathbb{N}$$

Given the current value of $\phi(x_k)$, the future value can be predicted easily.

(Continued)

If the observable g of interest can be expanded as:

$$g = \sum_i \lambda_i \phi_i$$

using Koopman eigenpairs (ϕ_i, λ_i), then:

$$g(x_{k+l}) = \sum_i \lambda_i^l \phi_i(x_k)$$

This allows prediction of the future value of f using the linear properties of the Koopman operator.

Example: Continuous Spectrum Issue

While this approach is appealing, it has a main drawback: eigenfunction expansions of observables may not exist.

Consider for this $R_\theta : \mathbb{S}^1 \to \mathbb{S}^1$ as an irrational rotation and $\mathcal{K}_\theta : L^2(m) \to L^2(m)$.

$\rightsquigarrow \sigma(\mathcal{K}_\theta) = \mathbb{S}^1$

Example not pathological; based on it e.g. the famous Pendulum has also a continuous spectrum.

Extended Dynamic Mode Decomposition (EDMD) (1/4)

We want to approximate $\mathcal{K}^\tau$ in a data-driven manner. We use the technique known as EDMD:

We use snapshots of data $(t_i, x_i, y_i)_{i=1}^n$, where

- x_i is a vector in the state space $\mathcal{X}$, containing all the necessary variables that describe the condition or configuration of the system at that moment and
- $y_i = f_{t_i+\tau}(x_i)$ for a time step $\tau > 0$.

In practice, if $(t_i, x_i)_{i=1}^n$ is the measured data, we simply restructure it into $(t_i, x_i, x_{i+1})_{i=1}^{n-1}$

EDMD (2/4)

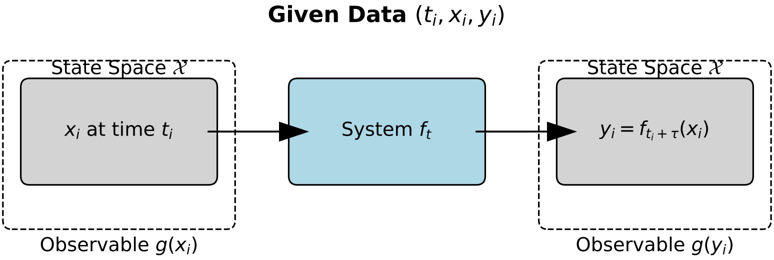

Figure: Structure of the assumed Data

EDMD (3/4)

Let $\phi_1, \ldots, \phi_\ell$ and $\psi_1, \ldots, \psi_m$ be two finite dictionaries in $C_b([0,T] \times \mathbb{R})$. Construct the matrices

$$\mathit{\Psi}_n := \begin{bmatrix} \psi_1(t_1, x_1) & \psi_1(t_2, x_2) & \ldots & \psi_1(t_n, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_m(t_1, x_1) & \psi_m(t_2, x_2) & \ldots & \psi_m(t_n, x_n) \end{bmatrix} \in \mathbb{R}^{m \times n},$$

$$\mathit{\Phi}_n^\tau := \begin{bmatrix} \phi_1(t_1 + \tau, y_1) & \phi_1(t_2 + \tau, y_2) & \ldots & \phi_1(t_n + \tau, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_l(t_1 + \tau, y_1) & \phi_l(t_2 + \tau, y_2) & \ldots & \phi_l(t_n + \tau, y_n) \end{bmatrix} \in \mathbb{R}^{l \times n}$$

Approximating $\mathcal{K}^\tau$ (4/4)

Given the matrices Ψ_n, Φ_n^τ, the approximation $\mathcal{K}_{mn}^\tau$ of the Koopman operator is computed as

$$\mathcal{K}_{mn}^\tau = \arg\min_K \|\Phi_n^\tau - K\Psi_n\|_F,$$

where F denotes the Frobenius norm

One can show that in the infinite data limit

$$\lim_{n\to\infty} \mathcal{K}_{mn}^\tau \varphi = \mathcal{K}_m^\tau \varphi,$$

a.s., where $\mathcal{K}_m^\tau$ is the restriction of $\mathcal{K}^\tau$ on $\mathrm{span}(\psi_1, \dots, \psi_m)$.

Problem Statement: Finding UPOs in Chaotic Dynamical Systems

- Unstable Periodic Orbits (UPOs) are periodic trajectories sensitive to initial conditions.
- Theoretical Importance: often UPOs are dense in chaotic attractors, providing insights into the structure of the attractor.
- Practical Importance: Critical in applications like space transportation, where periodic trajectories are vital for planning efficient satellite routes.

UPOs and Invariant Measures

- Invariant Measures: Measures that remain unchanged under the dynamics of the system. For discrete dynamical systems:

$$\mu(f^{-1}(A)) = \mu(A)$$

- UPOs are often connected to the extremal invariant measures that maximize or minimize time averages of observables.
- The Poincaré map reduces the problem to a discrete dynamical system, making it easier to work with invariant measures.

Koopman Operator in this Context

- Invariant measures are closely related to the Koopman operator, because μ is invariant if

$$\int v(f(x)) - v(x)\, d\mu(x) = \int \mathcal{L}v(x)\, d\mu(x) = 0$$

for all compactly supported test functions v. Here $\mathcal{L}$ is the generator of the discrete Koopman operator.

- Koopman allows us to approximate the dynamics in a data-driven manner, which is crucial when the system is not fully known.

Optimization Problem

- Given an observable $g : \mathbb{R}^n \to \mathbb{R}$ and a compact set $K \subset \mathbb{R}^n$, solve:

$$g^* = \min_{\mu \in Pr(K)} \int g(x)\, d\mu(x),$$

  where μ is invariant.
- Reformulation using polynomial approximations:
  - Approximate the space of continuous functions $C(K)$ by polynomials $\mathbb{R}[x]$, dense in $C(K)$ (special case of Stone-Weierstrass).
  - The problem can be equivalently written as:

$$g^* = \min_{\mu \in \mathcal{P}(K)} \left\{ \int g(x)\, d\mu(x) \,\middle|\, \int \mathcal{L}v(x)\, d\mu(x) = 0,\ \forall v \in \mathbb{R}[x] \right\}$$

Polynomial Data Assumptions

- f is a polynomial vector field, and each component of $f(x)$ is a polynomial of degree $\leq d$.
- The observable $g : \mathbb{R}^n \to \mathbb{R}$ is a polynomial of degree $\leq d$.
- The compact set $K \subset \mathbb{R}^n$ is defined by polynomial inequalities

$$K = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \ldots, h_m(x) \geq 0\},$$

where $h_i(x)$ are polynomials of degree $\leq d$.

Moment Method

- The moment method is used to encode the invariance condition in terms of moments of the measure.
- The moment method is crucial because it allows the optimization problem to be formulated in a finite-dimensional space.
- The moment method also ensures that the optimization problem is linear, which is critical for solvability using standard numerical techniques.
- The use of polynomials is justified because they provide a natural basis for the moment method, allowing for representation of the optimization problem.

Our Final Goal

- We aim to apply these techniques to the classical three-body problem, a chaotic system with applications in aerospace.
- The objective is to find UPOs that could be used in mission planning and space transportation.
- We are developing a method inspired by Bramburger-Fantuzzi, focusing on two-dimensional Poincaré maps.

Demonstration: Henon Map

- We have tested our code on the Henon Map, successfully approximating the invariant measure.
- The following slides show visual results, confirming the code's correctness and robustness.

The Henon Map

- Dynamics governed by:

$$f(x) = \begin{pmatrix} 1 - ax_1^2 + x_2 \\ bx_1 \end{pmatrix}.$$

- Parameters: $a = 1.4$, $b = 0.3$ (chaotic dynamics).

Illustrations

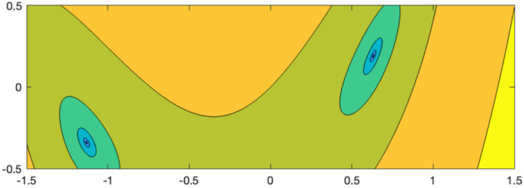

Figure: Fixed points of the Henon Map (explicit)

Henon Attractor

Stochastic Differential Equations

In concrete situations, there is often some intrinsic randomness in the measurements and observations, for example:

- Molecular dynamics
- Evolution of stock prices
- Data coming from biological processes (e.g. tumor growth, cellular growth, spreading of genes)
- Motion of pollutants in fluids
- etc.

This motivates the study of data-driven modeling through Stochastic Differential Equations (SDEs)

The SDE model

We look at models in the following class of Itô SDEs

$$dX_t = f(X_t)dt + \sqrt{2\epsilon}dB_t,$$

where:

- $X_t \in \mathbb{R}^d$ is an Itô process.
- $f : \mathbb{R}^d \to \mathbb{R}^d$ is the drift term.
- $\epsilon > 0$ is the diffusion coefficient.
- B_t is a d-dimensional Brownian motion.

This Itô SDE is known as the Langevin dynamics, which models the velocity of particles subject to both deterministic forces and random noise

Data and SDEs

The data have the following structure:

- Data snapshots $(x_i, y_i, t_i)_{i=1}^n$
- $x_i = X_{t_i}$, $y_i = X_{t_i+\tau}$ for a $\tau > 0$
- These are the observed values of the stochastic process at times t_i and $t_i + \tau$

Goal: Suppose we have some data, realization of the SDE

$$dX_t = f(X_t)dt + \sqrt{2\epsilon}dB_t.$$

Can we identify the drift f and the diffusion ε in a data-driven manner?

Comparing data and synthetic SDE

To fit f and ε we need to minimize the "distance" between our synthetic SDE model and the observed data

Due to the intrinsic randomness, it is not possible to compare directly the data with the solutions of the synthetic model

We need a deterministic quantity. Given an observable $g = g(x)$, we try to learn

$$\mathbb{E}[g(X_{t+\tau})|X_t = x],$$

that is, the expected value of the random variable $g(X_t)$ at time $t + \tau$, given the observation $X_t = x$ at time t

Koopman Operator for Stochastic Dynamics

Extending Koopman theory to stochastic differential equations (SDEs) involves regularizing dynamics with noise. Consider the Itô SDE:

$$dX_t = f(X_t)dt + \sqrt{2\epsilon}dB_t$$

where $X_t \in \mathbb{R}^d$, $f : \mathbb{R}^d \to \mathbb{R}^d$, and B_t is a d-dimensional Brownian motion. The stochastic Koopman operator $\mathcal{K}_t$ acts on a function g as:

$$(\mathcal{K}_t g)(x) = \mathbb{E}[g(X_{t+\tau}) \mid X_t = x]$$

The Koopman generator

Even though I can approximate the Koopman operator, I still have no information on f and ε. This information is hidden in the infinitesimal generator of the Koopman operator

$$\mathcal{L}g = \lim_{\tau \to 0^+} \frac{\mathbb{E}[g(X_{t+\tau})|X_t = x] - g(x)}{\tau}$$

is the operator that describes the instantaneous evolution of observables on the trajectories of the SDE. Indeed, in this case

$$\mathcal{L} = f \cdot \nabla + \varepsilon \Delta$$

Approximating the Koopman generator

In order to fit f and ε we need the Koopman generator we also need to approximate. This can be done as

$$\mathcal{L}_{mn}^{\tau} g = \frac{\mathbb{E}[\mathcal{K}_{mn}^{\tau} g(X_t) | X_t = x] - g(x)}{\tau}$$

Fitting f and ε

At this point, since we know that the real generator of the SDE has the form

$$\mathcal{L} = f \cdot \nabla + \varepsilon \Delta,$$

we find the approximation of f, ε solving a minimization problem

$$(\hat{f}, \hat{\varepsilon}) = \arg \min_{f \in \mathcal{H}, \varepsilon \in \mathbb{R}} \|\mathcal{L}^{\tau}_{mn} \Phi - f \cdot \nabla \Phi - \varepsilon \Delta \Phi\|_{\mathcal{H}} + \nu \|(f, \varepsilon)\|_{\mathcal{H} \times \mathbb{R}},$$

where $\mathcal{H}$ is a suitable function space and Φ is the dictionary of function we used for EDMD

Important Results

Problem 1: We know that (Bramburger, Fantuzzi, 2024)

$$\lim_{\tau \to 0} \lim_{m \to +\infty} \lim_{n \to +\infty} \mathcal{L}^{\tau}_{mn} g = \lim_{m \to +\infty} \lim_{\tau \to 0} \lim_{n \to +\infty} \mathcal{L}^{\tau}_{mn} g = \mathcal{L} g$$

In other words, $\mathcal{L}^{\tau}_{mn}$ converges strongly to $\mathcal{L}$

Problem 2: Needs work! In particular, we need to quantify the distance

$$\|\hat{f} - f\|_{\mathcal{H}} + |\hat{\varepsilon} - \varepsilon|$$

in the limits $\tau \to 0$, $m \to +\infty$ and $n \to +\infty$. Is it true that $\hat{f} \to f$ and $\hat{\varepsilon} \to \varepsilon$ when $\mathcal{L}^{\tau}_{mm} \to \mathcal{L}$?

The Basis Pursuit Algorithm

Problem: In practice, we do not have infinite data or dictionaries!

We look for ways to improve the convergence: After having computed $\hat{f}$ and $\hat{\varepsilon}$, we can compute an approximate spectrum and eigenfunctions of $\hat{\mathcal{L}} = \hat{f} \cdot \nabla + \hat{\varepsilon} \cdot \Delta$.

If $\mathcal{H}$ is chosen well (i.e. $\hat{f}$ is regular enough and satisfies certain growth assumptions), the eigenfunctions will generate a basis, which can be used to perform EDMD and repeat the whole process (step 2 and step 3)

The Basis Pursuit Algorithm

Note that the algorithm works only if we have a basis of eigenfunctions. But this is fine if we work with SDEs and we make an additional assumption on f

Proposition

Let $f = \nabla U$ and let U be such that

$$\liminf_{|x| \to +\infty} \frac{|U(x)|}{x^2} > 0.$$

Then the spectrum $\sigma(\mathcal{L}) = \sigma(f \cdot \nabla + \varepsilon \Delta)$ is discrete and there exists a basis of eigenfunctions.

Steps of the Algorithm

0. Preparation. Obtain $\mathcal{L}^{\tau}_{mn}$ with EDMD.
1. First identification. Use the information $f \cdot \nabla + \varepsilon \varDelta$ and fit $f^{(1)}$ and $\varepsilon^{(1)}$ as

$$(f^{(1)}, \varepsilon^{(1)}) = \mathrm{argmin}_{f,\varepsilon} \|\mathcal{L}^{\tau}_{mn} \varPhi - (f \cdot \nabla \varPhi + \varepsilon \varDelta \varPhi)\|,$$

2. Update. From step 2:

$$\mathcal{L}^{(1)} = f^{(1)} \cdot \nabla + \varepsilon^{(1)} \varDelta.$$

For this operator, we can compute numerically eigenvalues λ_i and eigenfunctions $\phi_i^{(1)}$. Update the basis $\varPhi$ used in step 0.
3. Iterate. Iterate M times $\rightsquigarrow$ $f^{(M)}$ and $\varepsilon^{(M)}$ that, hopefully, converge to the right values f and ε.

Possible Advantages

Improved Approximation:

- Each iteration improves the approximation of the generator $\mathcal{L}$ by getting closer to the true eigenfunctions.
- This leads to a more accurate finite-dimensional representation of the infinite-dimensional Koopman operator.

Data Efficiency:

- The iterative basis update converges towards the correct eigenfunctions, requiring potentially less data to achieve accurate models.
- This enhances the efficiency and accuracy of data-driven modeling, especially in high-dimensional systems.

Thank You for Your Attention! ☺

If you have any questions, feel free to ask/discuss.