Introduction to the Koopman Operator, the VAMP Score, and VAMPNets

Phillip Bement

January 16, 2024

1 The Koopman Operator

1.1 defining the Koopman operator

Consider any Markov chain. Examples of Markov chains:

- A Markov chain with a finite number of states. The probabilities for the transitions between states are given by the stochastic transition matrix $\mathbf{S}$.
- Brownian motion. Here the number of states is infinite. Time is continuous, so to get an actual Markov chain, we pick a discrete time step Δt and only look at the particle in evenly spaced intervals of Δt to see how it has moved.
- Classical dynamics on a continuous phase space. Again, we have to pick a time spacing Δt. Deterministic dynamics are a special case of stochastic dynamics.
- Dynamics of a polymer interacting with its environment. This is like classical dynamics, but with randomness introduced because of the interaction with the environment. Again, we still need to introduce a Δt. Often, the polymer is modelled as occupying a finite volume, often a box with cyclic boundary conditions.

From the last 3 examples, we can see that most systems we encounter in physics can be modelled as Markov chains.

If we are lucky, the dynamics of the system we are studying will be linear. This makes the system easier (or in many cases, *possible at all*) to solve analytically. But some systems have non-linear dynamics. The non-linear dynamics make them quite difficult to study.

The Koopman operator approach gives us a way around this issue: We can linearize the problem by considering functions on the state space rather than individual states in the state space.

Suppose that states of the system are denoted by points $\mathbf{x} \in X$ where X is the set of all possible system states (in classical physics, it might represent the phase space of the system). Then the Markov chain dynamics are defined by a probability (density) $p(\mathbf{x}_{t+1}|\mathbf{x}_t)$. To define the Koopman operator, we should then start considering the vector space $\mathbb{C}^X$, the space of functions on X.

Starting from a probability distribution q_0 at time t, the distribution q_1 at time $t+1$ is given by:

$$q_1(\mathbf{x}) = \int_X q_0(\mathbf{y}) d\mathbf{y} p(\mathbf{x}|\mathbf{y})$$

By inspecting this formula, we can see that q_1 considered as a function/vector is linear in q_0. The Koopman operator approach then is to name this linear operator $\mathcal{K}$. We can then write the above equation as $q_1 = \mathcal{K}q_0$, with $\mathcal{K}$ being an operator on the vector space of functions. (By which we mean reasonably nice functions; as physicists we won't worry about what exactly that means in mathematical terms.) Since $\mathcal{K}$ is an operator on the full vector space, we can use it on functions that aren't proper probability distributions. I.e. they can be unnormalized, or have negative or even complex values. We'll write the result of multiplying $\mathcal{K}$ with a general function g as follows:

$$(\mathcal{K}g)(x) = \int_X d\mathbf{y} g(\mathbf{y}) p(\mathbf{x}|\mathbf{y})$$

Our first example (the Markov chain on a finite state space).

1.2 equilibrium and inner products

Some Markov chains approach an equilibrium if you let them evolve for long enough. Some do not. For example, in classical mechanics, the dynamics are reversible, so that even after a large amount of time, the starting conditions can still be deduced by just looking at the current state of the system. So classical physics is not a case where equilibrium is approached. The fact that the dynamics is deterministic and reversible prevents it. Brownian motion is not deterministic, but it also does not give an equilibrium distribution. The variance of the particle position increases without bound.

A polymer interacting with its environment (our 4th example above) will approach an equilibrium distribution, q. This is defined by:

$$q(\mathbf{x}) = \int_X d\mathbf{y} q(\mathbf{y}) p(\mathbf{x}|\mathbf{y})$$

$$\text{i.e.} \qquad \mathcal{K}q = q$$

For systems that do indeed approach equilibrium, this should uniquely define q once we impose positivity and normalization.

Once we have an equilibrium distribution, we get a natural inner product on the function space $\mathbb{C}^X$:

$$\langle g_1 | g_2 \rangle = \int_X d\mathbf{x} g_1(\mathbf{x}) g_2(\mathbf{x}) q(\mathbf{x}) = \langle g_1 g_2 \rangle_q$$

This gives us a notion of orthogonality of functions, which will be helpful when we move ahead towards defining the VAMP score. But this depended on having access to an equilibrium distribution.

1.3 example: the Ornstein-Uhlenbeck process

NOTE: Here should present the singular-value decomposition for the O-U process and its solution in terms of Hermite polynomials.

NOTE: Should also write several paragraphs about the fact that coordinates are all polynomials in x and so the coordinates are redundant (they form a 1d curve in terms of which coordinate vectors are actually attainable). Should also maybe note the effect this has on what the best prediction of future coordinates should be given the current coordinates.

2 The VAMP Score

3 VAMPNets

4 Extra Notes

Ornstein Uhlenbeck in 2D ($k_x = k_y$). First two Koopman coordinates (i.e. Koopman singular functions) are x and y.

Restricting to just these coordinates, is this the "density matrix"?

$$\rho = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$$

with coordinate operators:

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \mathbf{Y} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

so that $\langle \mathbf{A} \rangle = \mathrm{Tr}\,[\mathbf{A}\rho]$