The Koopman Operator Applied to 3-State Markov Chains

Mathematics Departmental Honors Research Thesis

Jack Graham
Advisor: Dr. David Gaebler

Fall 2022 - Spring 2023

1 Introduction

This document is a write-up describing the work done for my departmental honors thesis with Dr. David Gaebler during my senior year at Hillsdale college. The math presented here is not new. Most of my time spent researching involved learning about dynamical systems, investigating related topics, and trying to develop an understanding of the Koopman operator in how it is applied in certain instances. As a result, this document will present material that has already been proven by various sources. Also, while there were certain basic results on related topics that I found with the help of Dr. Gaebler, we both believe they have been proven elsewhere even though we could not find a source for them. The positive results of this research are the working out of new mathematical technology (the Koopman Operator) in a specific situation, as well as a report of the various things I learned in the process of my departmental honors research.

2 Dynamical Systems

2.1 Hamiltonian Mechanics

I did not know what exactly what a dynamical system was until I began research. It helped to walk through an example of one. Consider a physical system such as a mass on a spring. It can be useful to describe this situation in terms of the **Hamiltonian**, an expression of the total energy of the system. For a case as simple as a mass on a spring, we can express the Hamiltonian (H) simply as the sum of potential and kinetic energies:

$$H = T + V \tag{1}$$

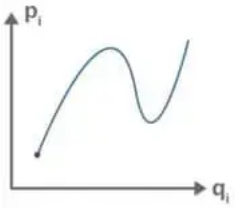


Figure 1: Parameterized Curve in Phase Space

In Hamiltonian mechanics, the system is described in terms of position (q_i and momentum (p_i). Using the following differential equations, the entire movement of the system can be described:

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{2}$$

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{3}$$

Not only are these equations sufficient to describe the evolution of the system, it allows for a new way to represent the system. By plotting position and momentum against each other, the system can be expressed as a curve parameterized by time (Figure 1). This space is called **phase space**. For n particles in k dimensions, this would be some kind of $2nk$-dimensional manifold, since each particle as k position coordinates and k momentum coordinates.

In the specific example of a mass on a spring, the phase space consists of 2 coordinates, one for the x-position and one for momentum in the x-direction. The Hamiltonian for a harmonic oscillator is

$$H = \frac{p^2}{2m} + \frac{1}{2}\kappa x^2 \tag{4}$$

where m is mass and κ is the spring constant. This gives the following system of differential equations:

$$\dot{p} = -2\kappa x \tag{5}$$

$$\dot{x} = \frac{p}{m} \tag{6}$$

In phase space, the harmonic oscillator is depicted as an ellipse (Figure 2).

2.2 State Space

For a general dynamical system, a **state space** is needed. This is the generalization of phase space. A state space is a set of states of a system. For example, a state of the harmonic oscillator is a point in time the oscillator exists, given by the position and momentum of the mass attached to the spring.

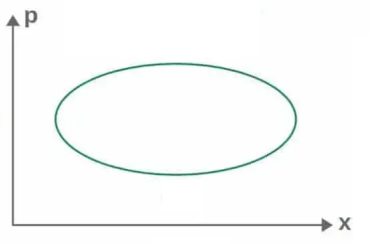


Figure 2: Harmonic Oscillator in Phase Space

2.3 Discrete Dynamical Systems

Let $S \subset \mathbb{R}^n$ be the state space of a dynamical system. Let $T : S \mapsto S$ be a dynamic map. Let $x \in S$ be a point, and let $t \in \mathbb{Z}$ be the discrete time index. Then we say

$$x^{t+1} = T(x^t). \tag{7}$$

This equation describes a **discrete-time dynamical system**. If there is some kind of additional structure on S, (like, if it is a vector space or topological space), then that structure is respected. A single map T is equivalent to a family of maps $\{T^{(n)}\}_{n \in \mathbf{N}}$ for which $T^{(0)}$ is the identity map and there exists a semigroup of maps (indexed by the natural numbers) satisfying the property

$$T^{(a)}T^{(b)} = T^{(a+b)} \tag{8}$$

where $T^{(a)}$ is just T composed with itself a times. For example, consider a game of Monopoly. Let the x be the probability vector that player one is on any given board space. Then S is the set of all possible probability vectors. Start at GO (x_0), and let the dynamic map T be the role of two dice. Then $T^{(4)}(x_0)$ is the set of probable locations after 4 turns. This is a discrete dynamical system.

2.4 Continuous Dynamical Systems

A **continuous-time dynamical system** is a family $\{F^t\}$ of functions from $S \mapsto S$ where F^0 is the identity and the semigroup property holds:

$$F^s F^t = F^{s+t} \tag{9}$$

We can make one of these families by first letting f be the function for some dynamical system, such as $\dot{x} = f(x)$. Then the **flow map** $F^t : S \mapsto S$ is the map from the initial state, x_0, to the state at time $t \in \mathbb{R}$, i.e.

$$F^t(x_0) = x_0 + \int_{x_0, t'=0}^{t} f(x(t'))dt'. \tag{10}$$

An example of a continuous dynamical system is the x-position of a harmonic oscillator.

2.5 Points of Interest

For any initial state x_0, the set $\{T^{(n)}(x_0)|n \in \mathbf{N}\}$ or $\{F^t(x_0)|t \in \mathbf{R}\}$ is called the **orbit** or **trajectory** of x_0. **Fixed points** are orbits that consist only of themselves. **Limit cycles** are closed curves in state space that correspond to periodic orbits. **Invariant sets** are sets whose points have trajectories that always remain in that set. **Attractors** are invariant sets with a dense orbit that many initial conditions converge to. Basically, identifying key characteristics of certain orbits is extremely useful for understanding the evolution of a dynamical system.

3 The Koopman Operator

3.1 Observables

An **observable** of a state space is a function $g \mapsto \mathbf{C}$ (or sometimes just $\mathbf{R}$). The name *observable* comes from quantum mechanics. An observable is essentially looking at one piece of data from a dynamical system instead of the whole system. For example, an observable of the harmonic oscillator could be the total potential energy of the system. For another example, an observable of a fluid inside a box could be the pressure or velocity at a certain specified point. The set of all observables form a linear vector space (linearity coming from the Koopman Operator as shown later).

3.2 What is the Koopman Operator?

Definition 3.1 (Koopman Operator - Discrete Time System)**.** Let $g : S \mapsto \mathbb{C}$ be a real-valued observable of some discrete-time dynamical system (7). The collection of all such observables forms a linear vector space. The **Koopman Operator**, denoted by U, is a linear transformation on this vector space given by

$$Ug(x) = g \circ T(x), \tag{11}$$

where $\circ$ denotes the composition operator.

The linearity of the Koopman Operator follows from the linearity of the composition operator:

$$U[g_1(x) + g_2(x)] = [g_1(x) + g_2(x)] \circ T(x) \tag{12}$$

$$= g_1(x) \circ T(x) + g_2 \circ T(x) = Ug_1(x) + Ug_2(x) \tag{13}$$

Notice this definition applies to a discrete-time dynamical system. In practice, this can refer to a discrete time sampling of a continuous-time dynamical system given by a set of differential equations. However, an analogue to the Koopman operator can also be defined for continuous-time systems.

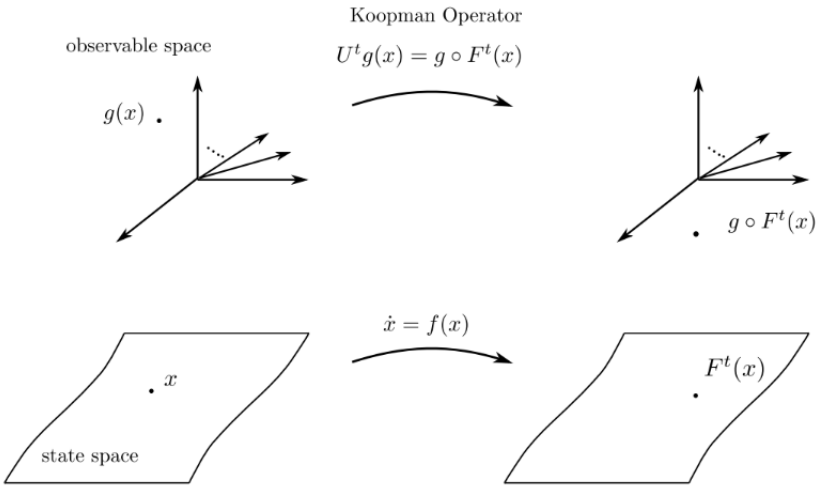


Figure 3: Lifting Dynamics from State to Observable Space

Definition 3.2 ("Koopman Operator" - Continuous Time System)**.** Then $\{U^t\}_{t\geq 0}$ is a one-parameter **semigroup of Koopman Operators**, where each element is given by

$$U^t g(x) = g \circ F^t(x). \tag{14}$$

U^t is also linear, similarly to the discrete-time case. All these proofs and definitions can be found in *Introduction to Koopman operator theory of dynamical systems*. The Koopman operator is to observable space what T or F^t is for state space; it steps time forward (Figure 3). The Koopman operator, however, works under a space with different properties, i.e., observable is linear, but also infinite-dimensional. As a result, there is a trade-off when using it. This trade-off is similar to the trade-off in Hamiltonian mechanics when going from physical space to phase space. Given the positive benefits from linearity, this could be a worthwhile trade.

3.3 Eigenstuffs of the Koopman Operator

Definition 3.3 (Eigenfunction-Eigenvalue Pair)**.** Let $\phi_j : S \mapsto \mathbb{C}$ be a complex-valued observable of (1), and λ_j a complex number. Then $(\phi_j,\ \lambda_j)$ is an **eigenfunction-eigenvalue pair** of the Koopman operator if it satisfies

$$U^t \phi_j = e^{\lambda_j t} \phi_j \tag{15}$$

Since the Koopman Operator is linear, we can look at its eigenvalues and eigenvectors to determine long term behavior. In particular, it might be useful to look at the *spectrum* or the appropriate analogue of the Koopman Operator.

Definition 3.4 (Spectrum)**.** The *resolvent set* $\rho(U)$ of U is the set of all scalars λ in $\mathbb{C}$ such that the linear transformation $\lambda I - U$ has a densley defined continuous inverse. Then the **spectrum** $\sigma(U)$ of U is the complement of $\rho(U)$ in $\mathbb{C}$. That is, $\sigma(U) = \{\lambda \in \mathbb{C} : \lambda I - U \text{ is not invertible}\}$.

In line with this, looking at the *spectral properties* of the Koopman Operator or the *spectral theorem* could give useful insights.

3.4 Related Ideas

For good record keeping, I will mention several things that I learned about and that we discussed pursuing.

3.4.1 Relationship between Discrete and Continuous

Given a continuous-time system, one can derive a corresponding discrete-time system by sampling at regular intervals. That is, for some F^t, with $t \in \mathbf{R}$, define $T^{(n)} = F^n$ for $n \in \mathbf{N}$. We examined the convergence of this discrete sample to its corresponding continuous system. If you keep sampling in smaller and more frequent intervals (i.e. $T^{(\frac{n}{\tau})}$ for $\tau \to \infty$), then it is trivially true that the discrete sample converges to its corresponding continuous system. A more interesting question would be the convergence of the discrete sample if you sampled at arbitrary times.

3.4.2 Miscellaneous Notes

Here I transcribe some notes about operator convergence (and how they involve e, which I asked about). One thing Dr. Gaebler considered is some sort of convergence between discrete and continuous Koopman Operators. Then, we asked the question, for an operator T in a Banach space: is somehow $T\phi^{(t)} = \lim_{?\to ?} T\phi$? Well, we know $T_n \to T$ in the operator norm if $\lim_{n\to\infty} ||T - T_n||_{\text{op}} = 0$. This is sort of like the operator version of uniform convergence, and it is often difficult to obtain. We also know of a *strong operator* convergence, in which we say $T_n \to T$ if, for every vector v in that space, $T_n v = Tv$. We considered these convergence definitions Also, just to note how e pops up, you can define $e^T = I + T + \frac{T^2}{2!} + \frac{T^3}{3!} + ...$ by infinite series. Then, you can do all sorts of things with operators and matrices by equating $T^t = e^{tA}$, that follow directly from rules of exponentiation and the infinite series. This led us to consider pursuing questions relating to finding some sort of exponential generator of the semigroup. Finally, we mentioned looking up *Stone's Theorem* or the *Hille-Yosida Theorem*.

4 3-State Markov Chains

4.1 What Has Markov to Do with Koopman?

The first thing I tried to do, in an attempt to get a better grasp of Koopman Operators, was to find and work out an example of one. This was hard. Maybe I just needed to practice more with differential equations and dynamical systems, or maybe it really was difficult. My best attempt was to use the harmonic oscillator, but I found it hard to grasp what observable I should use on the system, and then what to do with the Koopman Operator when I got there. So, Dr. Gaebler suggested I moved to something more discrete, and I ended up looking at Markov Chains.

Now, what about Markov Chains in 3 dimensions specifically was interesting? Broadly speaking, we know a lot of analytical properties of the complex plane because we can map it to the unit disk. Furthermore, since Markov Chains have the probability requirement that every entry of a probability vector must add up to 1, we can project 3-State Markov Chains, which exist in 3 dimensions, down to 2 dimensions. Then we can map that to the complex plane using a Schwarz-Cristoffel mapping and carry over analytical properties (and apply them to the Koopman Operator). That was the idea that started me off on Markov chains, even if what kept me on Markov chains was separate.

4.2 Markov Chains and Transition Matrices

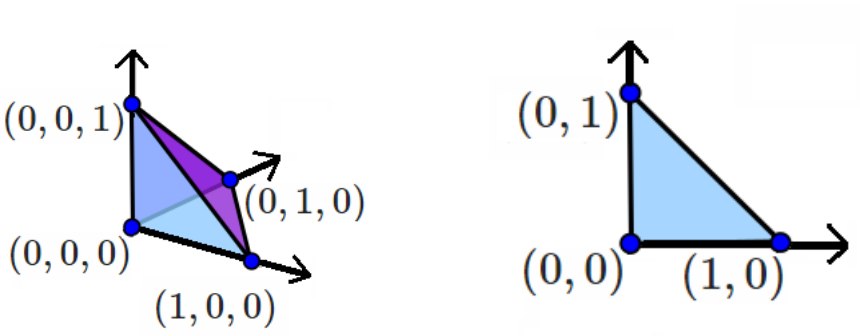
Definition 4.1. A **discrete-time Markov chain** is a sequence of random variables $x_1, x_2, x_3, ...$ with the *Markov property*, namely that the probability of moving to the next state depends only on the present state and not on the previous states. We will represent transitions in this sequence using a *stochastic matrix* as a transition matrix.

Important Note: A Markov *state* is different than a dynamical system *state*. A Markov state is a just a certain way the the object of the Markov chain exists. Often Markov chains are expressed as graphs. In that case, Markov states are the nodes that can be transitioned to. The dynamical system state for a Markov chain, in contrast, is a probability vector of being at any of the given Markov states. So, the state (of the Markov chain in terms of a dynamical system) will be the probability vector at time t denoted x_t.

Thus, for 3-states, we say $x_{t+1} = T_3 x_t$, for the timestep operator T_3 where $T_3 x_t = A x_t$, with

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ 1-(a+d) & 1-(b+e) & 1-(c+f) \end{bmatrix}, \tag{16}$$

for x_t being the probability vector at time t, and $0 \leq a, b, c, d, e, f \leq 1$. Note that the final entry of each column is determined by the previous two, since each column must add up to 1.



(a) 3-Dimensional Probabilities (b) 2-Dimensional Probabilities

Figure 4: Projection from 3 to 2 Dimensions

Now, we want to project this 3-dimensional transition matrix into 2 dimensions (Figure 1). We will denote the 2-dimensional timestep operator with T. Let the *projection* of x (which is essentially just erasing a dimension) be denoted $\pi(x)$. To find what T does, we must satisfy the following property:

$$\pi(T_3 x_t) = T\pi(x_t). \tag{17}$$

That is, the projection of the timestep of x_t is the projected timestep of the projected x_t (which we will also just refer to as x_t). Since (17) is also equal to $\pi(Ax_t)$, we find that it is only satisfied when T is defined as follows:

Definition 4.2 (Projected Timestep T)**.** The **projected timestep** T is given by

$$Tx_t = \begin{bmatrix} a-c & b-c \\ d-f & e-f \end{bmatrix} x_t + \begin{bmatrix} c \\ f \end{bmatrix}. \tag{18}$$

We also might refer to the matrix B and vector ζ such that $Tx_t = Bx_t + \zeta$.

4.3 Eigenvalues

We can examine the eigenvalues of A to understand the long-term behavior of T_3, especially if we consider the *Perron-Frobenius theorem*. We set $\det(A - \lambda I) = 0$ and solve for the characteristic polynomial of A. Let

$$\alpha = a - c + e - f \tag{19}$$

$$\beta = ae - af - bd + bf + cd - ce \tag{20}$$

Then, the characteristic polynomial of A is

$$-\lambda^3 + (1+\alpha)\lambda^2 - (\alpha+\beta)\lambda + \beta = (\lambda - 1)(-\lambda^2 + \alpha\lambda - \beta) = 0 \tag{21}$$

Thus the eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = \frac{\alpha+\sqrt{\alpha^2-4\beta}}{2}$, and $\lambda_3 = \frac{\alpha-\sqrt{\alpha^2-4\beta}}{2}$. The steady state eignevector corresponds to λ_1. But what are the eigenvalues of our projected timestep? Well, as it turns out, $\alpha = \text{tr}(B)$ and $\beta = \det(B)$, so the eigenvalues of B are the same λ_2 and λ_3 as A. This is good because it means that T_3 and T have the same long-term effects, as desired.

4.4 Convex Geometry

With the 3-state Markov Chain projected down into 2 dimensions, it can be plotted as a probability simplex. Intuitively, we saw that any set of all probability vectors would have to be a subset of the previous set of all probability vectors. Geometrically, this meant that each triangle had to remain within the one for the previous iteration, allowing us to use convex geometry and describe the simplex in terms of **barrycentric coordinates**. In addition to learning about convex geometry, I also proved that the limit point is constant under barrycentric coordinates (likely already proven somewhere).

Proof. Let $\vec{x}$ be the limit point. The triangular points in the first iteration are (0,0), (1,0), and (0,1). Let p, q, and r be the weights for a barrycentric coordinate. Then

$$\vec{x} = p\begin{bmatrix}1\\0\end{bmatrix} + q\begin{bmatrix}0\\1\end{bmatrix} + r\begin{bmatrix}0\\0\end{bmatrix} = p\begin{bmatrix}1\\0\end{bmatrix} + q\begin{bmatrix}0\\1\end{bmatrix} \tag{22}$$

since the corner is the zero vector. Then

$$\vec{x} = \begin{bmatrix}p\\q\end{bmatrix} \tag{23}$$

and the barrycentric coordinate weights are simply $p = x_1$, $q = x_2$, and $r = 0$. This also satisfies $x_1 + x_2 + 0 = 1$. Consider

$$T\vec{x} = B\vec{x} + \zeta = \begin{bmatrix}(a-c)x_1 + (b-c)x_2 + c\\(d-f)x_1 + (e-f)x_2 + f\end{bmatrix} \tag{24}$$

The barrycentric coordinates are weights of the iterations of the vertices of the simplex:

$$T\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}c\\f\end{bmatrix} \tag{25}$$

$$T\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}a\\d\end{bmatrix} \tag{26}$$

$$T\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}b\\e\end{bmatrix} \tag{27}$$

These resulting vectors are the verticies of the probability triangle given after one iteration. Since $x_1 + x_2 = 1$, (24) becomes

$$x_1\begin{bmatrix}a\\d\end{bmatrix} + x_2\begin{bmatrix}b\\e\end{bmatrix} + \begin{bmatrix}1-x_1-x_2\\1-x_1-x_2\end{bmatrix}\begin{bmatrix}c\\f\end{bmatrix} = x_1\begin{bmatrix}a\\d\end{bmatrix} + x_2\begin{bmatrix}b\\e\end{bmatrix} \tag{28}$$

This shows that, from the initial state to the next timestep, the limit point maintains the same barrycentric coordinates (i.e., x_1, x_2, and 0). Additionally, suppose $\vec{x} = r_1\vec{\alpha} + r_2\vec{\beta} + r_3\vec{\gamma}$, where r_1, r_2, and r_3 are the barrycentric weights and $\vec{\alpha}$, $\vec{\beta}$, and $\vec{\gamma}$ are the three vertices of any current set of all probabilities. Then

$$T\vec{x} = B(r_1\vec{\alpha} + r_2\vec{\beta} + r_3\vec{\gamma}) + \zeta \tag{29}$$

Since $r_1 + r_2 + r_3 = 1$, (29) becomes

$$Br_1\vec{\alpha} + Br_2\vec{\beta} + Br_3\vec{\gamma} + (r_1 + r_2 + r_3)\zeta \tag{30}$$

$$= r_1(B\vec{\alpha} + \zeta) + r_2(B\vec{\beta} + \zeta) + r_3(B\vec{\gamma} + \zeta) \tag{31}$$

$$= r_1T\vec{\alpha} + r_2T\vec{\beta} + r_3T\vec{\gamma} \tag{32}$$

This shows that the next iteration of an arbitrary set has the same barrycentric coordinates (r_1, r_2, and r_3). Therefore, by induction, the limit point has the same barrycentric coordinates each iteration. □

4.5 Decomposition of A

In an attempt to understand how the transition matrix A evolved over time and what that looked like geometrically, I researched ways to decompose A. First, I began with an singular value decomposition and then an eigenvalue decomposition. Both of these placed certain restrictions on the type of stochastic matrix A which could be productive to look into (e.g. is A always diagonalizable). The main purpose of this decomposition was to determine, given a certain matrix A, how it could be made up of translations, rotations, shears, and reflections. This served the greater purpose of understanding what A (or really B) was doing geometrically. Things we considered were: does T rotate the triangle the same amount each time, does the triangle shrink the same each time, are angle measurements preserved under iterations, and are there families of similar stochastic matrices? One basic result I found was that, if you permuted A, then the resulting triangle would cover the exact same area but have its vertices reassigned. That is, each triangle corresponded to a matrix family of $3! = 6$ permutations. Also, we considered the question of what kind of matrix M could multiply A such that AM was still stochastic. By assuming both A and M are stochastic and multiplying them out, it was clear it would always yield a stochastic matrix. In the other direction, if A and AM were stochastic, then the columns of M had to add up to 1, but did not necessarily have to have entries lying between 0 and 1. That is, it implied a slightly weaker condition than being stochastic.

4.6 Rock-Paper-Scissors, an Example

Let's get back to the Koopman Operator. To calculate an example, consider a Rock-Paper-Scissors tournament. By analyzing how a certain player's strategy evolves over time, we can determine a Markov chain (Figure 5) with probabilities reflecting what that player is likely to throw giving what they just threw.

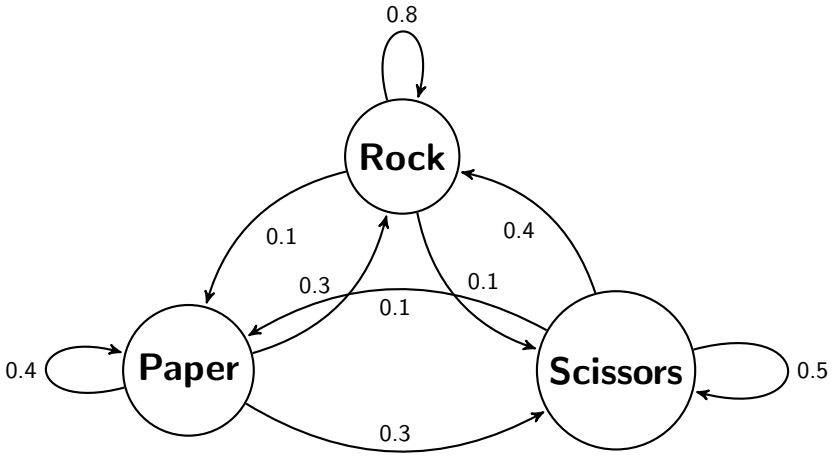


Figure 5: A Rock-Paper-Scissors Strategy Probability

4.6.1 Markov Chain

Let r_t, p_t, and s_t be the probabilities at time t of throwing rock, paper, and scissors respectively. Given these probabilities, we can write the transition matrix A such that,

$$x_{t+1} = Ax_t \text{ becomes } \begin{bmatrix} r_{t+1} \\ p_{t+1} \\ s_{t+1} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.3 & 0.4 \\ 0.1 & 0.4 & 0.1 \\ 0.1 & 0.3 & 0.5 \end{bmatrix} \begin{bmatrix} r_t \\ p_t \\ s_t \end{bmatrix} \tag{33}$$

Projecting down, we derive that

$$Tx_t = \begin{bmatrix} 0.4 & -0.1 \\ 0 & 0.3 \end{bmatrix} x_t + \begin{bmatrix} 0.4 \\ 0.1 \end{bmatrix} \tag{34}$$

Since $\alpha = \text{tr}(B) = 0.7$ and $\beta = \det(B) = 0.12$, the eigenvalues are

$$\lambda_2 \approx 0.842443 \text{ and } \lambda_3 \approx -0.142443 \tag{35}$$

We can plot the state of all possible probabilities as it evolves over time (Figure 6). As can be seen, the probabilities converge to a limit point, which is the steady state eigenvector (fixed point) corresponding the the eigenvalue of A, $\lambda_1 = 1$. This eigenvector is approximately

$$v_1 \approx \begin{bmatrix} 0.642857 \\ 0.142857 \\ 0.214286 \end{bmatrix} \tag{36}$$

showing that this player will settle into a strategy that involves using rock about 64% of the time.

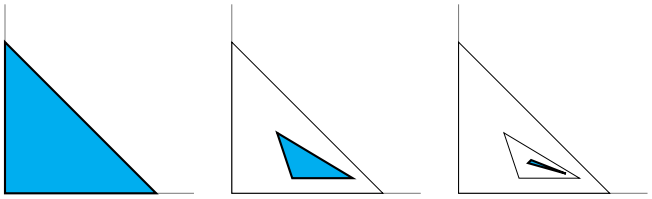


Figure 6: Evolution of Rock-Paper-Scissors Probabilities

4.6.2 Koopman Operator

In order to set up a Koopman Operator, we need an observable. For this reason, imagine that, every round of the game, you need to purchase the option that you choose. Say, Scissors are \$2.00, Paper is \$0.50, and Rocks are \$0.01. Then we can define the **expected marginal cost** as $g(x_t) = 0.01r_t + 0.5p_t + 2s_t$. We can choose a function such as this to be our observable. Following this, the Koopman Operator (with respect to g) is

$$Ug(x_t) = g \circ T(x_t) = g(Ax_t) = 0.01r_{t+1} + 0.5p_{t+1} + 2s_{t+1} \tag{37}$$

$$= 0.258r_t + 0.803p_t + 1.054s_t, \text{ or} \tag{38}$$

$$= \begin{bmatrix} 0.258 & 0.803 & 1.054 \end{bmatrix} \cdot x_t \tag{39}$$

Note this is the same as $A^T \begin{bmatrix} 0.01 \\ 0.5 \\ 2 \end{bmatrix}$. As it turns out, in Markov Chains, expected value lends itself very readily as an observable for the Koopman Operator. If we restrict observables to linear combinations within the state,

$$[g(x_t) = E[C] = c_1x_1 + c_2x_2 + c_3x_3 = \tag{40}$$

$$\begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix} \cdot x_t = \vec{C} \cdot x_t \tag{41}$$

Then, because the transpose A^T is the adjoint for the real-valued finite dimensional space (Markov Chain state space),

$$Ug(x_t) = g(Ax_t) = (Ax_t) \cdot \vec{C} = x_t \cdot (A^T\vec{C}) \tag{42}$$

In this application, we see that the Koopman Operator is fundamentally tied to the adjoint. This enables us to find the desired observable for the next timestep without knowing the state after the next timestep.

4.7 Miscellaneous Remarks

4.7.1 Hidden Markov Model

It seems there is a connection between a hidden Markov model and using the Koopman operator on a Markov model. That is, if the chain is hidden and the observable is all that is known, is the Koopman Operator just a way to operate with a hidden Markov model?

4.7.2 Literature Crawl

In trying to find sources that had described what we found as basic results, I stumbled across several papers with relevant terms and information. **Dynamic Mode Decomposition** (DMD) can be thought of as an algorithm for finding Koopman modes (and might use dimensional reduction). **Linear Inverse Modelling** is a kind of DMD that uses empirical orthogonal functions (EOF) from a stochastic linear Markov system. I also read briefly about the **Ulam Matrix/Ulam Method**, which is a technique to approximate the Perron-Frobenius operator by a stochastic matrix. The **Stochastic Koopman Operator** is a composition of a skew-flow map. I read about it in connection with Markovian random dynamical systems, which I thought might be relevant. It was in reading about a **Markovian Dynamical System** that we found the information pertaining expected value and the adjoint. Finally, I found a tool called a Koopman matrix, which is an extended DMD that evaluates an approximation for the Koopman Operator.

5 Future Work

Since much of the project was learning various topics, there are many specific questions to dive into and investigate. Some of these ideas include:

- Analyze differences and/or convergence properties of the discrete Koopman Operator with the semigroup of continuous Koopman Operators.
- Analyze the implications of certain functional space requirements (Hilbert Space, Banach Space, L^2 Space, etc.).
- Perform Spectrum analysis on an eigenbasis expansion of the Koopman Operator.
- Calculate the Koopman Operator for more complicated dynamical systems.
- Compare properties of the Koopman Operator with geometric properties of the 3-state Markov Chain (map to the complex unit disc, convex geometry, etc.)

In the end, I am very thankful for the opportunity to do this research.

6 References

- Arbabi, Hasan. *Introduction to Koopman operator theory of dynamical systems.*
- Kubrusly, Carlos. *Spectral Theory of Operators on Hilbert Spaces.*

- Mexić, Igor. *Koopman Operator, Geometry, and Learning of Dynamical Systems.*
- Mardt, Andreas, Luca Pasquali, Frank Noe, Hao Wu. *Deep learning Markov and Koopman models with physical constraints.*