Chaos within Stochastic Processes: Investigation of Chaotic Behavior for Markov Chains

Ryan Schilt

December 15, 2008

Introduction

This paper will outline an attempt to show that chaos may exist with in the iterative nature of Markov Chains. The impetus for the investigation will come from a brief analysis of periodic markov chains and their resemblance to multidimensional iterative maps. Therefore, the basics of iterative maps will be reviewed as well.

As will be clear in the end, this is a work in progress and very few solid connections have been made, however there appears to be hope within the field of Ergodic Theory, so a short description and an overview of the features it offers to this problem will be discussed.

Discrete Space and Time Markov Chains

In probability, a Markov chain is a process where the probability of something to happen depends only on the previous state. Each Markov chain has an associated probability transition matrix, $p(i,j)$:

$$p(i,j) = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & & \ddots & \\ p_{m,1} & p_{m,2} & \cdots & p_{m,n} \end{bmatrix} \tag{1}$$

Where each $p_{i,j}$ represents the probability of moving to state j, given that you are currently at state i and the each row sums to one. There are many different types of Markov chains that offer good representations of the natural world.

An important property of Markov chains is represented in the Chapman-Kolmogorov equation:

$$p^{m+n}(i,j) = \sum_k p^m(i,k)p^n(k,j) \tag{2}$$

Stated simply, Eq.2 means that if $m = 1$ and $n = 1$ then $p^2 = pp$, or if $m = 2$ and $n = 1$, then $p^3 = p^2p = ppp$. This is valuable because $p^m(i,j)$ is interpreted as the probability of being at state j in m steps, given that you started at state i, which can be determined by multiplying the probability transition matrix with itself m times and looking at the (i,j) element.

Two important properties of a markov chain are it's *period* and *stationarydistribution* (if it exists). When taking n steps, $p^n(i,j)$, one can refer to the period of state, x, as the greatest common division of $I_x = \{n \geqslant 1 : p^n(x,x) > 0\}$. When a chain as a period of one is reffered to as *aperiodic.*

A stationary distribution is a solution to $\pi p = \pi$, where π is a row vector and p is the probability transition matrix. This is important because π becomes an equilibrium state for a markov chain. However, one is not always guareenteed a stationary distribution. As an astute mathematician would note, solving for π is an exercise in linear algebra and only under certain contidions does a solution exist, let alone a unique one.

One final property of a markov chain is called *irreducibility*. A markov chain is irreducible when $\forall x, y \in$ State Space, $x \longrightarrow y$ and $y \longrightarrow x$, or x communicates with y and y communicates with x.

With these definitions, the following Theorem is given without proof[1]:

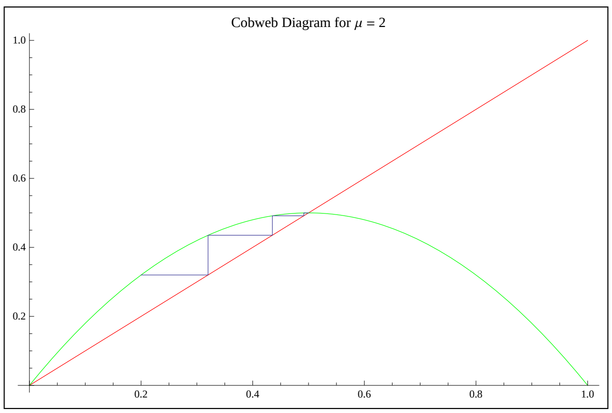
Theorem 1 *Given a Probability transition matrix, p, Suppose it is* irreducible*,* aperiodic*, and has a stationary distribution π. Then as $n \to \infty$, $p^n(x, y) \to \pi(y)$.*

Fixed Points: a quick review

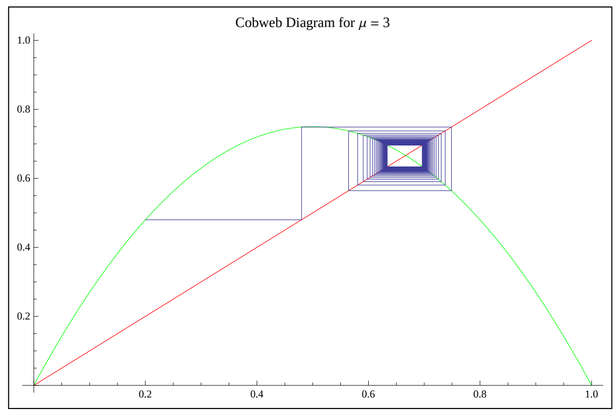
An important topic in Chaos is one-dimensional maps. The most common of thses maps is the logistic map, $x_{n+1} = \mu x_n(1 - x_n)$. By iterating this map, one will see chaotic behavior depending on the value of μ, specifically for $\mu > 3.57$.[2] Figure 1 shows what this iteration looks like in a Cobweb Diagram. The importantance of Figure 1 is that it shows when chaos can be seen. For $\mu = 2$ and $\mu = 3$ (subfigures a) and b)) one can see that the iteration path is very organized and where both are attracted to a fixed point. However, for $\mu = 3.8$, the system is described as chaotic and the iteration path bounces throughout the plot.

This method will become useful as something similar will be used to try and find chaotic behavior within markov chains .

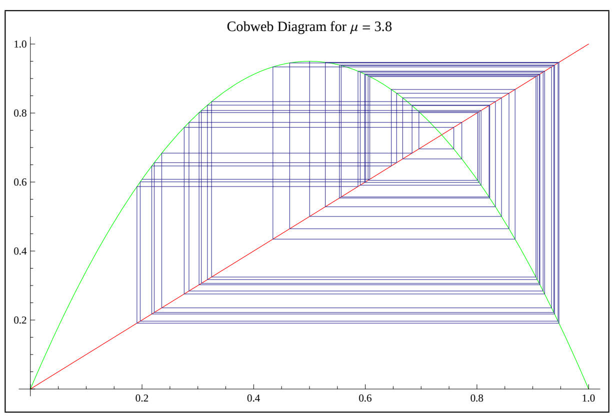
[1] See reference [1], pg. 54



(a)



(b)



(c)

Figure 1: Cobweb diagrams for the logistic map with initial $x_0 = 0.2$, and out to 50 iterations. The green curve is $f(x) = \mu x(1-x)$, the red line is $f(x) = x$, and the blue line show the iterative path that is taken. a) $\mu = 2$, b) $\mu = 3$, and $\mu = 3.8$.

Chaos through Transition Matrix Iteration

Let's begin with an example. Consider the Markov Chain $(X_n)_{n\geqslant 0}$, given in Figure 2, with state space $S = \{1, 2, 3, 4\}$. The transition matrix in Eqn. 3 is both aperiodic and irreducible, so if it assumed

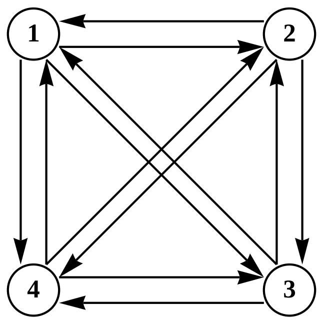


Figure 2: An aperiodic Markov Chain.

$$p(i,j) = \begin{bmatrix} 0.2 & 0.3 & 0.1 & 0.4 \\ 0.6 & 0.2 & 0.1 & 0.1 \\ 0.3 & 0.1 & 0.1 & 0.5 \\ 0.7 & 0.2 & 0.05 & 0.05 \end{bmatrix} \tag{3}$$

that it does have a Stationary Distribution, then Theorem 1 applies. What does the matrix look like at each step?

$$p^3(i,j) = \begin{bmatrix} 0.3820 & 0.2450 & 0.0910 & 0.2820 \\ 0.4670 & 0.2245 & 0.0843 & 0.2243 \\ 0.3980 & 0.2425 & 0.0898 & 0.2698 \\ 0.4795 & 0.2213 & 0.0836 & 0.2156 \end{bmatrix} \tag{4}$$

$$p^6(i,j) = \begin{bmatrix} 0.4318 & 0.2331 & 0.0872 & 0.2480 \\ 0.4243 & 0.2349 & 0.0877 & 0.2531 \\ 0.4303 & 0.2334 & 0.0873 & 0.2490 \\ 0.4232 & 0.2351 & 0.0878 & 0.2539 \end{bmatrix} \tag{5}$$

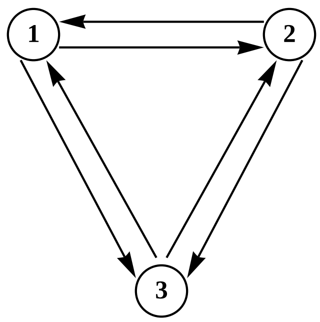
$$p^{15}(i,j) = \begin{bmatrix} 0.4277 & 0.2340 & 0.0875 & 0.2508 \\ 0.4277 & 0.2340 & 0.0875 & 0.2508 \\ 0.4277 & 0.2340 & 0.0875 & 0.2508 \\ 0.4277 & 0.2340 & 0.0875 & 0.2508 \end{bmatrix} \tag{6}$$

$$p^{20}(i,j) = \begin{bmatrix} 0.4277 & 0.2340 & 0.0875 & 0.2508 \\ 0.4277 & 0.2340 & 0.0875 & 0.2508 \\ 0.4277 & 0.2340 & 0.0875 & 0.2508 \\ 0.4277 & 0.2340 & 0.0875 & 0.2508 \end{bmatrix} \tag{7}$$

As Eqn's 4-7 show, each row of the matrix tends toward the standard distribution. Another way to think about this is that each row defines a point in $\mathbb{R}^n$, and in this case it would be a four dimensional point. Furthermore, at each iteration of the matrix one could track the point each row defines to see how the iterations are behaving. Let's see what this looks like in $\mathbb{R}^3$ with a periodic chain.

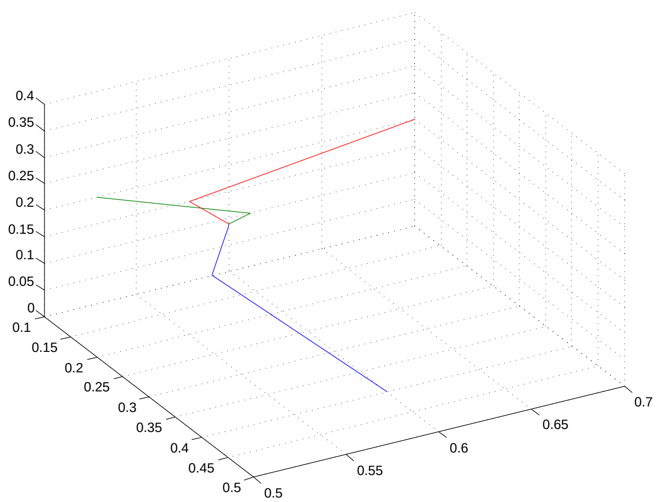
Periodic Chain

Consider a probability transition matrix of the following form: This transition matrix is 2-periodic and



$$p(i,j) = \begin{bmatrix} 0 & p_{1,2} & p_{1,3} \\ p_{2,1} & 0 & p_{2,3} \\ p_{3,1} & p_{3,2} & 0 \end{bmatrix} \quad (8)$$

is irreducible, therefore Thereom 1 cannot be used, and there is no way of knowing what will happen when $p^n(i,j)$ is calculated for $n \geqslant 2$. But the visualization described above can be of some help. Below are several figures each with a different initial transition matrix.



$$p(i,j) = \begin{bmatrix} 0.4000 & 0.6000 & 0.0000 \\ 0.2000 & 0.5000 & 0.3000 \\ 0.1000 & 0.7000 & 0.2000 \end{bmatrix} \quad (9)$$

Figure 3: 2-periodic Markov Chain.

These are only a few of the transition matrices that were investigated, however most of them had patterns similar to Figures 3-6 and none were found that didn't have a stationary distribution.

These results were not quite as expected. Finding patters similar to the Logistic Map of Figure 1 would have revealed direct chaotic behavior, however these iterations seem fairly regular.

By simply looking at examples and not "seeing" chaotic behavior, does not mean that it doesn't exist. There are many different types of transition matrices and only an really small subset were investigated. To show definitively that chaos can or cannot happen, we will now move into an area of mathematics known as Ergodic Theory.

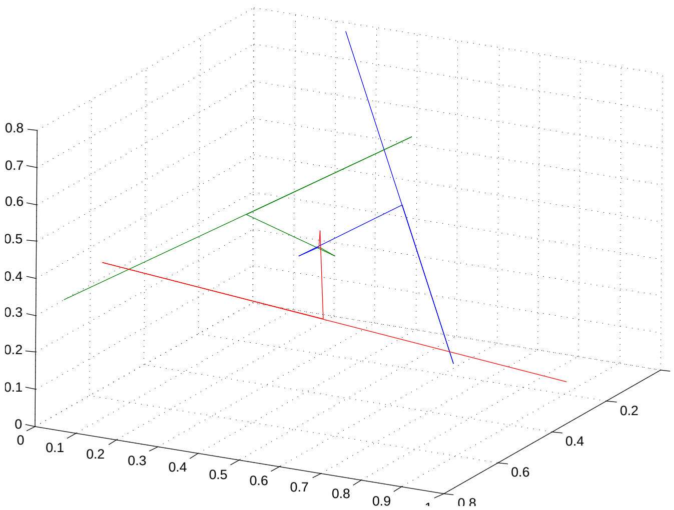


Figure 4: 2-periodic Markov Chain.

$$p(i,j) = \begin{bmatrix} 0.0000 & 0.2238 & 0.7762 \\ 0.6991 & 0.0000 & 0.3009 \\ 0.1386 & 0.8614 & 0.0000 \end{bmatrix} \tag{10}$$

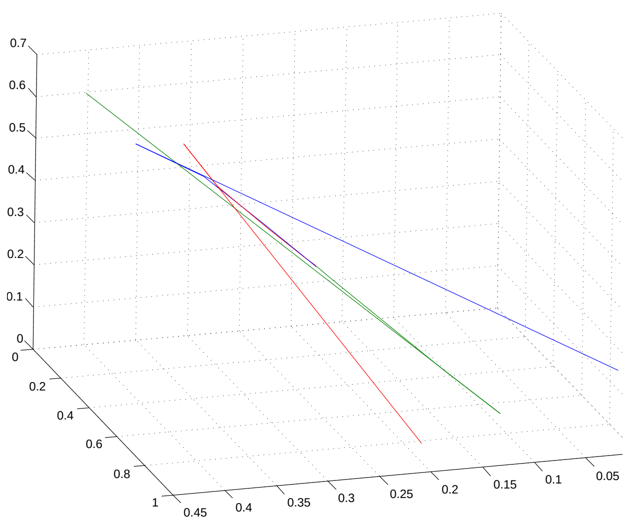


Figure 5: 2-periodic Markov Chain.

$$p(i,j) = \begin{bmatrix} 0.0000 & 0.8530 & 0.1470 \\ 0.4018 & 0.0000 & 0.5982 \\ 0.1839 & 0.8161 & 0.0000 \end{bmatrix} \tag{11}$$

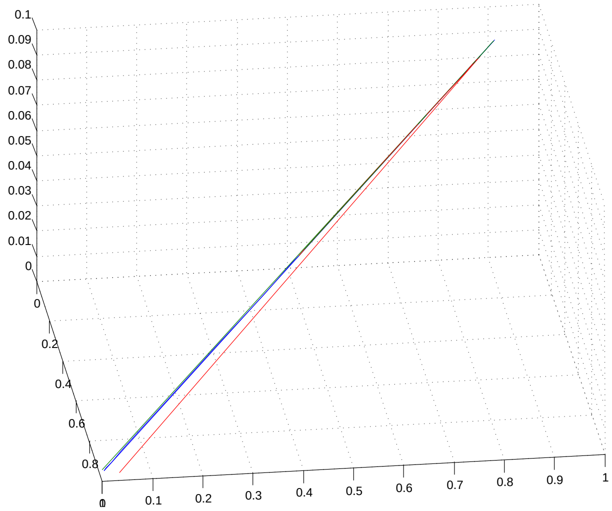


Figure 6: 2-periodic Markov Chain.

$$p(i,j) = \begin{bmatrix} 0.0000 & 0.9133 & 0.0867 \\ 0.9961 & 0.0000 & 0.0039 \\ 0.9619 & 0.0381 & 0.0000 \end{bmatrix} \tag{12}$$

Ergodic Theory Insights

Some important insights may come from the Birkhoff's Ergodic Theorem[3], which is stated below:

Theorem 2 *Let T be an ergodic endomorphism of the probability space X and let $f : X \to \mathbb{R}$ be a real-valued measurable function. Then for almost every $x \in X$ we have,*

$$\frac{1}{n}\sum_{j=1}^{n} f \circ T^j(x) \to \int f dm \tag{13}$$

as $n \to \infty$.

The left-hand side of Eqn.13 just says how often the orbit of x (that is the points, x, Tx, T^2x, $\ldots$) lies in A, and the right-hand side is just the measure of A. Thus, for an ergodic endomorphism, "space-averages = time-averages almost everywhere." Moreover, if T is continuous and uniquely ergodic with Borel measure m and f is continuous, then we can replace the almost everywhere convergence in Eqn.13 with "everywhere."

In terms of the Markov problem, X will be a countable or finite set, and μ is a function such that, $\mu : X \to \mathbb{R}$, with $\mu(x) \geqslant 0, \forall x \in X$ and $\sum_{x \in S} \mu(x) = 1$. μ is then a probability measure. This makes μ the function f in Eqn.13. Any morkov chain can be considered a collection of probability measures indexed by X such that: $p : X \times X \to \mathbb{R}$. Which would mean that for each $x \in X$, $p(x,y)$ (a function of from $y \in X$ to $\mathbb{R}$) is a probability measure over X.

Now, define the space of all probability measures over X as V. Then, the transformation $T : V \to V$ can be defined as:

$$T_\mu(y) = \sum_{x \in X} \mu(x)p(x, y) \tag{14}$$

Now if p is aperiodic and irreducible the there exists a probability measure, $\pi \in V$, such that $T_\mu^n \to \pi$ as $n \to \infty$, and irregardless of μ. As an analogy for iterative maps, μ is kind of like the initial value for some linear map, p, which has fixed point, π. However, before this analogy can be an equality it needs to be shown that V is a compact metric space and that under the defined metric, T is a contraction.

Conclusion

The Chapman-Kolmogorov equation first lead to the idea that perhaps when a probability transition matrix is periodic, the matrix could be multiplied revealing chaotic behavior. However, buy using Cobweb diagram methods from one-dimensional maps, the stationary distributions were tracked showing little chaotic behavior. Only a small subset of all the possible 3×3, periodic, and irreducible transition matrices were looked at. It is possible that the chaotic behavior does still exist. In order to settle this remaining question, the problem might be able to be put into a form that satisfies Birkoff's Ergodic Theorem. Once this is done, the iteration of the transition matrix will be in the more elegant notation of the area-preserving map T.

Bibliography

[1] Rick Durrett. *Essentials of Stochastic Processes.* Springer, 2004.

[2] Steven Strogatz. *Nonlinear Dyamics and Chaos.* Perseus Books Group, 1994.

[3] Eric W. Weisstein. *Birkhoff's Ergodic Theorem.* From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/BirkhoffsErgodicTheorem.html`.