Markov Chains

TFs: Justin Zhu

Markov Chains

A Markov Chain is a walk along a discrete **state space** $\{1, 2, \ldots, M\}$. We let X_t denote which element of the state space the walk is on at time t. The Markov Chain is the set of random variables denoting where the walk is at all points in time, $\{X_0, X_1, X_2, \ldots\}$. Each X_i takes on values that are in the state space, so if $X_1 = 3$, then at time 1, we are at state 3.

What makes such a sequence of random variables a Markov Chain is the **Markov Property**, which says that if you want to predict where the chain is at at a future time, you only need to use the present state, and not any past information. In other words, the *given the present, the future and past are conditionally independent*.

Mathematically, this says:

$$P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = P(X_{n+1} = j | X_n = i)$$

In words: Given that my history of states has been $i_0, i_2 \ldots i_n$, the distribution of where my next state will be doesn't depend on any of that history besides i_n, the most recent state.

State Properties

A state is either recurrent or transient.

- If you start at a **Recurrent State**, then you will always return back to that state at some point in the future. *You can leave, but you'll always return at some point.*
- Otherwise you are at a **Transient State**. There is some probability that once you leave you will never return. *There's a chance that you'll leave and never come back*

A state is either periodic or aperiodic.

- If you start at a **Periodic State** of period k, then the GCD of all of the possible number steps it would take to return back is k (which should be > 1).
- Otherwise you are at an **Aperiodic State.** The GCD of all of the possible number of steps it would take to return back is 1.

Transition Matrix

Element q_{ij} in square transition matrix Q is the probability that the chain goes from state i to state j, or more formally:

$$q_{ij} = P(X_{n+1} = j | X_n = i)$$

To find the probability that the chain goes from state i to state j in m steps, take the $(i, j)^{\text{th}}$ element of Q^m.

$$q_{ij}^{(m)} = P(X_{n+m} = j | X_n = i)$$

If X_0 is distributed according to row-vector PMF $\vec{p}$ (e.g. $p_j = P(X_0 = i_j)$), then the marginal PMF of X_n is $\vec{p}Q^n$.

Chain Properties

A chain is **irreducible** if you can get from anywhere to anywhere. An irreducible chain must have all of its states recurrent. A chain is **periodic** if any of its states are periodic, and is **aperiodic** if none of its states are periodic. In an irreducible chain, all states have the same period.

A chain is **reversible** with respect to $\vec{s}$ if $s_i q_{ij} = s_j q_{ji}$ for all i, j. A reversible chain running on $\vec{s}$ is indistinguishable whether it is running forwards in time or backwards in time. Examples of reversible chains include random walks on undirected networks, or any chain with $q_{ij} = q_{ji}$, where the Markov chain would be stationary with respect to $\vec{s} = (\frac{1}{M}, \frac{1}{M}, \dots, \frac{1}{M})$.

Reversibility Condition Implies Stationarity - If you have a PMF $\vec{s}$ on a Markov chain with transition matrix Q, then $s_i q_{ij} = s_j q_{ji}$ for all i, j implies that s is stationary.

Stationary Distribution

Let us say that the vector $\vec{p} = (p_1, p_2, \dots, p_M)$ is a possible and valid PMF of where the Markov Chain is at at a certain time. We will call this vector the stationary distribution, $\vec{s}$, if it satisfies $\vec{s}Q = \vec{s}$. As a consequence, if X_t has the stationary distribution, then all future $X_{t+1}, X_{t+2}, \dots$ also has the stationary distribution.

- If a Markov Chain is irreducible, then it has a unique stationary distribution. In addition, all entries of this stationary distribution are non-zero (which could have been inferred from the fact that all states are recurrent).
  - **Counterexample:** In the Gambler's Ruin problem, which is not irreducible, what ultimately happens to the chain can either be that one's money is always 0 or always N.
- If a Markov Chain is irreducible **and** aperiodic, then it has a unique stationary distribution $\vec{s}$ and

  $$\lim_{n \to \infty} P(X_n = i) = \vec{s}_i$$

  meaning that the chain *converges* to the stationary distribution.
  - **Counterexample:** Imagine a Markov chain which is just a cycle, and hence is periodic. Then, depending on where we start, $P(X_n = i)$ will be either 0 or 1 deterministically, and surely won't converge to the stationary distribution, which is uniform across all nodes in the cycle.

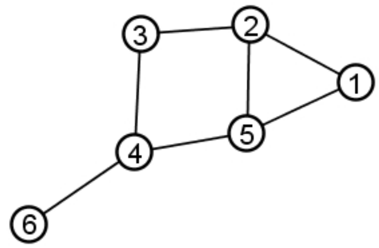
For irreducible, aperiodic chains, the stationary distribution exists, is unique, and s_i is the long-run probability of a chain being at state i. The expected number of steps to return back to i starting from i is $1/s_i$ To solve for the stationary distribution, you can solve for $(Q' - I)(\vec{s})' = 0$. The stationary distribution is uniform if the columns of Q sum to 1.

Random Walk on Undirected Network

If you have a certain number of nodes with undirected edges between them, and a chain can pick any edge uniformly at random and move to another node, then this is a random walk on an undirected network. The stationary distribution can be easily calculated. Let d_i be the degree of the ith node, meaning the number of edges connected to this node. Then, we have:

$$\vec{s}_i = \frac{d_i}{\sum_i d_i}$$

For example, in the below graph:



The stationary distribution would be proportional to: $(w_1, w_2, w_3, w_4, w_5, w_6) = (2, 3, 2, 3, 3, 1)$ and therefore, it would be $\left(\frac{2}{14}, \frac{3}{14}, \frac{2}{14}, \frac{3}{14}, \frac{3}{14}, \frac{1}{14}\right)$

Practice Problems

Example 1. Two-State Markov Chain.
Suppose X_n is a two-state Markov chain with transition matrix

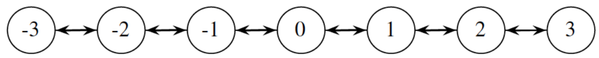
$$Q = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$$

(a) Find the stationary distribution $\vec{s} = (s_0, s_1)$ of X_n by solving $\vec{s}Q = \vec{s}$.

(b) Show that this Markov Chain is reversible under the stationary distribution found in part (a)

(c) Let $Z_n = (X_{n-1}, X_n)$. Is Z_n a Markov chain? If so, what are the states and transition matrix?

Example 2. Symmetrical Chain.
A Markov chain $X_0, X_1, X_2 \dots$ with state space $\{3, 2, 1, 0, 1, 2, 3\}$ proceeds as follows. The chain starts at $X_0 = 0$. If X_n is not an endpoint (-3 or 3), then X_{n+1} is $X_n + 1$ or $X_n - 1$, each with probability $1/2$. Otherwise, the chain gets reflected off the endpoint, i.e., from 3 it always goes to 2 and from -3 it always goes to -2. A diagram of the chain is shown below.



(a) Is $|X_0|, |X_1|, |X_2|, \dots$ a Markov Chain?

(b) Define the sign function $S(x)$ as follows:

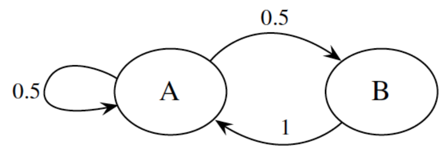
$$S(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

Is $S(X_0), S(X_1), S(X_2), \dots$ a Markov Chain?

(c) Find the stationary distribution of the original chain: $X_1, X_2, X_3 \dots$

(d) Find a simple way to modify some of the transition probabilities q_{ij} for $i, j \in \{-3, 3\}$ to make the stationary distribution of the modified chain uniform over the states.

Example 3. Not a Markov Chain.
A Markov chain has two states, A and B, with transitions as follows:



Suppose we do not get to observe this Markov chain, which we'll call $X_1, X_2, \ldots$. Instead, whenever the chain transitions from A back to A, we observe a 0, and whenever it changes states, we observe a 1. Let the sequence of 0's and 1's be called $Y_0, Y_1, Y_2, \ldots$ For example, if the X chain starts out as

$$A, A, B, A, B, A, A, \ldots$$

then the Y chain starts out as

$$0, 1, 1, 1, 1, 0, \ldots$$

(a) Show that $Y_0, Y_1, Y_2, \ldots$ is not a Markov Chain.

(b) In the past, when we have encountered processes that are not directly Markov chains, our remedy was to create a new state Z_n which would represent not only the current state X_n that the chain is on, but also remember the past m states. Thus, Z_n is not just one state, but a tuple of states that represent the chain's history:

$$Z_n = \left(X_n, X_{n-1}, X_{n-2} \ldots X_{n-(m-1)}\right)$$

Show that such a trick will not work for $Y_0, Y_1, Y_2 \ldots$. That is, no matter how large m is, $\{Z_n\}$ will never be a Markov Chain.

Example 4. Balls and Urns.
There are two urns with a total of $2N$ distinguishable balls. Initially, the first urn has N white balls and the second urn has N black balls. At each stage, we pick a ball at random from each urn and interchange them. Let X_n be the number of black balls in the first urn at time n. This is a Markov chain on the state space $\{0, 1, 2, \ldots N\}$.

(a) Give the transition probabilities of the chain.

(b) Set up, but do not evaluate, one or more equations that you would use to prove that

$$\vec{s}_i = \frac{\binom{N}{i}\binom{N}{N-i}}{\binom{2N}{N}}$$

is the stationary distribution for the above Markov Chain.

Extension: Show that the equation(s) you set up for part (b) are in fact true.