Discussion 14B

Markov Chains Intro II

Note 21

Recall that a Markov chain is defined with the following: the state space $\mathscr{X}$, the transition matrix P, and the initial distribution π_0. This implicitly defines a sequence of random variables X_n with distribution π_n, which denote the state of the Markov chain at timestep n. This sequence of random variables also obey the Markov property: the transition probabilities only depend on the current state, and not any prior states.

A **stationary distribution** (or the **invariant distribution**) of a Markov chain is a row vector π such that $\pi P = \pi$ (That is, transitioning does not change the distribution of states.). The previous equation is called the balance equation, and along with the normalization equation $\sum_i \pi(i) = 1$, we can solve for π.

Irreducibility: A Markov chain is *irreducible* if one can reach any state from any other state in a finite number of steps. An irreducible Markov chain is guaranteed to have a unique invariant distribution.

Periodicity: In an irreducible Markov chain, we define the *period* of a state i as

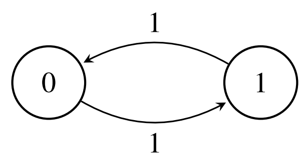
$$d(i) = \gcd\{n > 0 \mid P^n(i,i) = \mathbb{P}[X_n = i \mid X_0 = i] > 0\}.$$

If $d(i) = 1$ for all i, then a Markov chain is *aperiodic*. Otherwise, we say that the Markov chain is *periodic*. One important trait about periods is that they are the same for all states in an irreducible Markov chain. In other words, $d(i) = d(j)$ for all pairs of states i, j.

Fundamental Theorem of Markov Chains: If a Markov chain is irreducible and aperiodic, then for any initial distribution π_0, we have that $\pi_n \to \pi$ as $n \to \infty$, and π is the unique invariant distribution for the Markov chain.

1 Markov Chain Properties

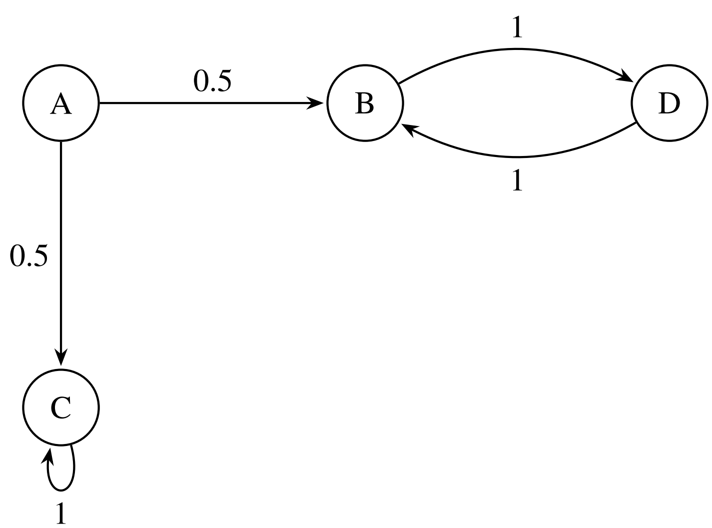
Note 21 In this question, we will build intuition towards the relationship between irreducibility and the invariant distribution(s), and the relationship between periodicity and convergence to the invariant distribution. Consider the following Markov chain.



(a) What is the period of this Markov chain?

(b) Does this Markov chain have a unique invariant distribution? If so, what is it?

(c) Does the distribution of X_n converge to the invariant distribution as $n \to \infty$ for any initial distribution? If not, give a counterexample and describe how the distribution of X_n behaves as $n \to \infty$.



(d) Describe the invariant distribution(s) of this Markov chain. What is the cardinality of the set of invariant distributions?

Solution:

(a) The period is 2, since we go from state 0 to state 1 with probability 1, and then back to state 0, thus creating a cycle of length 2.

(b) The unique invariant distribution is $\pi = (0.5, 0.5)$, since the Markov chain is symmetric and spends equal time in both states.

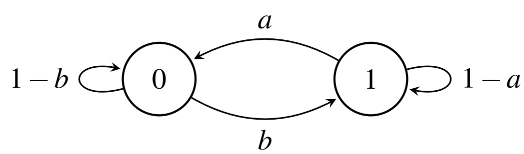
(c) No, the distribution of X_n does not converge to the invariant distribution as it is periodic. For any initial distribution $\pi = (\pi(0), \pi(1))$, we can see that X_n will keep alternatign between $(\pi(0), \pi(1))$ and $(\pi(1), \pi(0))$ as n increases, thus never converging to the invariant distribution.

(d) Notice that any probability mass that enters state C never leaves, and any probability mass that starts at state A will go to state C or state B and never come back. States B and D act just like the Markov Chain in the previous part, so we can see that any distribution of the form $\pi = (0, (1-p)/2, p, (1-p)/2)$ for $p \in [0,1]$ is an invariant distribution. Thus, there are uncountably infinitely many invariant distributions for this Markov chain.

2 Markov Chain Terminology

Note 21

In this question, we will walk you through terms related to Markov chains. Consider the following Markov chain.



(a) For what values of a and b is the above Markov chain irreducible? Reducible?

(b) For $a = 1$, $b = 1$, prove that the above Markov chain is periodic.

(c) For $0 < a < 1$, $0 < b < 1$, prove that the above Markov chain is aperiodic.

(d) Construct a transition probability matrix using the above Markov chain.

(e) Write down the balance equations for this Markov chain and solve them. Assume that the Markov chain is irreducible.

Solution:

(a) The Markov chain is irreducible if both a and b are non-zero. It is reducible if at least one of a and b is 0.

(b) We compute $d(0)$ to find that:

$$d(0) = \gcd\{2, 4, 6, ...\} = 2.$$

This is because if we start at a state X then we can get back to it after taking an even number of steps only (2, 4, 6, 8, etc.), not by taking an odd number of steps (1, 3, 5, 7, etc.). Thus, the chain is periodic with period 2.

(c) We compute $d(0)$ to find that:

$$d(0) = \gcd\{1, 2, 3, ...\} = 1.$$

Thus, the chain is aperiodic. Notice that the self-loops allow us to stay at the same node, thereby letting us get to any other node in an odd *or* even number of steps.

(d) The transition matrix is:

$$\begin{bmatrix} 1-b & b \\ a & 1-a \end{bmatrix}$$

(e) To solve for the stationary distribution, we need to solve for π in $\pi = \pi P$. This gives us the following system of equations:

$$\pi(0) = (1-b)\pi(0) + a\pi(1),$$
$$\pi(1) = b\pi(0) + (1-a)\pi(1).$$

One of the equations is redundant. We throw out the second equation and replace it with $\pi(0) + \pi(1) = 1$. This gives the solution

$$\pi = \frac{1}{a+b}\begin{bmatrix} a & b \end{bmatrix}.$$

3 Allen's Umbrella Setup

Note 21

Every morning, Allen walks from his home to Soda, and every evening, Allen walks from Soda to his home. Suppose that Allen has two umbrellas in his possession, but he sometimes leaves his umbrellas behind. Specifically, before leaving from his home or Soda, he checks the weather. If it is raining outside, he will bring exactly one umbrella (that is, if there is an umbrella where he currently is). If it is not raining outside, he will forget to bring his umbrella. Assume that the probability of rain is $p > 0$.

(a) Model this as a Markov chain. What is $\mathscr{X}$? Write down the transition matrix. (*Hint*: You should have 3 states. Keep in mind that our goal is to construct a Markov chain to solve part (c). You may want to try drawing out a 'naive' Markov chain with more states, and then see if you can combine some states together to get a simpler Markov chain.)

(b) Determine if the distribution of X_n converges to the invariant distribution, and compute the invariant distribution.

(c) In the long term, what is the probability that Allen walks through rain with no umbrella?

Solution:

(a) Let state i represent the situation that Allen has i umbrellas at his current location, for $i = 0$, 1, or 2.

Suppose Allen is in state 0. Then, Allen has no umbrellas to bring, so with probability 1 Allen arrives at a location with 2 umbrellas. That is,

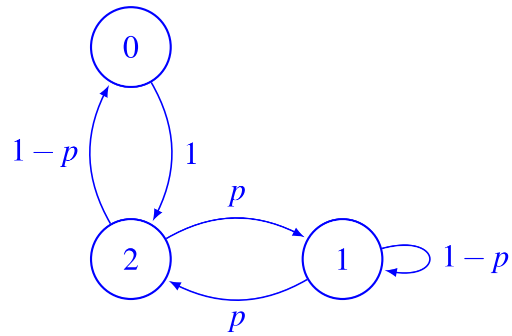
$$\mathbb{P}[X_{n+1} = 2 \mid X_n = 0] = 1.$$

Suppose Allen is in state 1. With probability p, it rains and Allen brings the umbrella, arriving at state 2. With probability $1-p$, Allen forgets the umbrella, so Allen arrives at state 1.

$$\mathbb{P}[X_{n+1} = 2 \mid X_n = 1] = p, \qquad \mathbb{P}[X_{n+1} = 1 \mid X_n = 1] = 1-p$$

Suppose Allen is in state 2. With probability p, it rains and Allen brings the umbrella, arriving at state 1. With probability $1-p$, Allen forgets the umbrella, so Allen arrives at state 0.

$$\mathbb{P}[X_{n+1}=1 \mid X_n=2]=p, \qquad \mathbb{P}[X_{n+1}=0 \mid X_n=2]=1-p$$



We summarize this with the transition matrix

$$P=\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1-p & p \\ 1-p & p & 0 \end{bmatrix}.$$

(b) Observe that the transition matrix has non-zero element in its diagonal, which means the minimum number of steps to transit to state 1 from itself is one. Thus this transition matrix is irreducible and aperiodic, so it converges to its invariant distribution.

To solve for the invariant distribution, we set $\pi P = \pi$, or $\pi(P-I)=0$. This yields the balance equations

$$\begin{bmatrix} \pi(0) & \pi(1) & \pi(2) \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 0 & -p & p \\ 1-p & p & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

As usual, one of the equations is redundant. We replace the last column by the normalization condition $\pi(0)+\pi(1)+\pi(2)=1$.

$$\begin{bmatrix} \pi(0) & \pi(1) & \pi(2) \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 0 & -p & 1 \\ 1-p & p & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

Now solve for the distribution:

$$\begin{bmatrix} \pi(0) & \pi(1) & \pi(2) \end{bmatrix} = \frac{1}{3-p} \begin{bmatrix} 1-p & 1 & 1 \end{bmatrix}$$

(c) Allen walks through rain with no umbrella if and only if it is raining when we take the transition from state 0 to 2 (i.e. Allen had no umbrellas, and moved to a location with 2 umbrellas). Note that given that we are in state 0, we must always take this transition with probability 1, so it suffices to compute the probability that it rains *and* we are in state 0.

Since the invariant distribution has $\pi(0) = \frac{1-p}{3-p}$, and it rains with probability p, the probability of walking through rain with no umbrella in the long term is

$$\mathbb{P}[\text{rain} \cap \text{no umbrella}] = p \cdot \frac{1-p}{3-p} = \frac{p(1-p)}{3-p}.$$

4 Can it be a Markov Chain?

Note 21

(a) A fly flies in a straight line in unit-length increments. Each second it moves to the left with probability 0.3, right with probability 0.3, and stays put with probability 0.4. There are two spiders at positions 1 and m and if the fly lands in either of those positions it is captured.

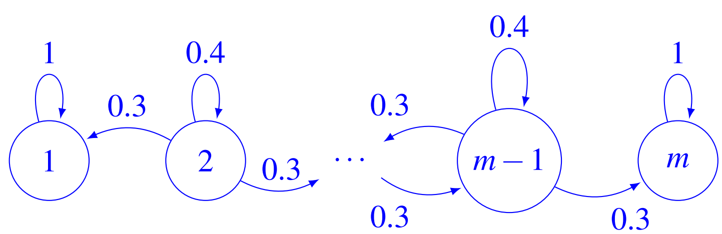
Given that the fly starts at state i, where $1 < i < m$, model this process as a Markov Chain. (Don't forget to specify the initial distribution!)

(b) Take the same scenario as in the previous part with $m = 4$. Let $Y_n = 0$ if at time n the fly is in position 1 or 2 and let $Y_n = 1$ if at time n the fly is in position 3 or 4.

Provide the state space for Y_n. Is the process Y_n a Markov chain?

Solution:

(a) We can draw the Markov chain as such:



The initial distribution is $\pi_0(i) = 1$, and $\pi_0(j) = 0$ for $j \neq i$.

(b) The state space is $\{0, 1\}$, the set of possible values that Y_n can take on.

Y_n cannot be a Markov chain because the memoryless property is violated.

For example, say $\mathbb{P}[X_0 = 2] = \mathbb{P}[X_0 = 3] = 1/2$ and $\mathbb{P}[X_0 = 1] = \mathbb{P}[X_0 = 4] = 0$. Then

$$\begin{aligned}
\mathbb{P}[Y_2 = 0 \mid Y_1 = 1, Y_0 = 0] &= \mathbb{P}[X_2 \in \{1,2\} \mid X_1 = 3, X_0 = 2] \\
&= \mathbb{P}[X_2 = 2 \mid X_1 = 3] = 0.3 \\
\mathbb{P}[Y_2 = 0 \mid Y_1 = 1, Y_0 = 1] &= \mathbb{P}[Y_2 = 0, Y_1 = 1, Y_0 = 1] / \mathbb{P}[Y_1 = 1, Y_0 = 1] \\
&= \mathbb{P}[X_2 = 2, X_1 = 3, X_0 = 3] / (\mathbb{P}[X_1 = 3, X_0 = 3] + \mathbb{P}[X_1 = 4, X_0 = 3]) \\
&= \frac{0.5 \cdot 0.4 \cdot 0.3}{0.5 \cdot 0.4 + 0.5 \cdot 0.3} = \frac{6}{35}
\end{aligned}$$

If Y was Markov, then $\mathbb{P}[Y_2 = 0 \mid Y_1 = 1, Y_0 = 0] = \mathbb{P}[Y_2 = 0 \mid Y_1 = 1] = \mathbb{P}[Y_2 = 0 \mid Y_1 = 1, Y_0 = 1]$. However, $0.3 > 6/35$, and so Y cannot be Markov.