CS 70 Discrete Mathematics and Probability Theory
Fall 2025 Sabin, Hug

Discussion 14B

Markov Chains Intro II

Note 26

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.

Irreducibility: A Markov chain is *irreducible* if one can reach any state from any other state in a finite number of steps.

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.

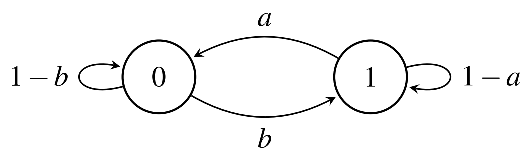
Recall the following definition: The **stationary distribution** (or the **invariant distribution**) of a Markov chain is the row vector π such that $\pi P = \pi$. (That is, transitioning does not change the distribution of states.)

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 Terminology

Note 26

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

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

2 Allen's Umbrella Setup

Note 26

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.

(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).)

(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,

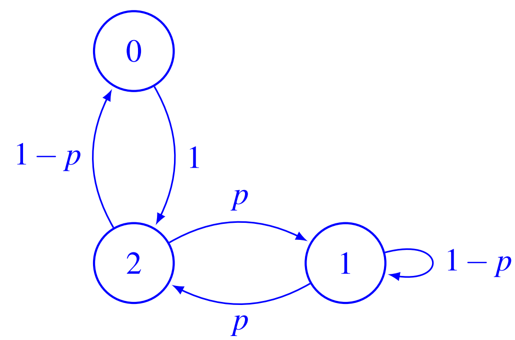
$$\mathbb{P}[X_{n+1} = 2 \mid X_n = 0] = 1.$$

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

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

3 Three Tails

Note 25

You flip a fair coin until you see three tails in a row.

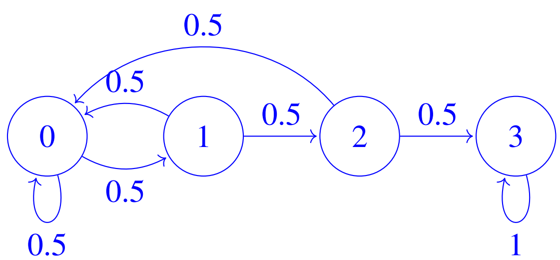
(a) What is the average number of timesteps until you get TTT?

(b) What is the average number of heads that you'll see until you get TTT? *Hint*: Modifying your equations from part (a) slightly to solve the original question.

Solution:

(a) **Solution 1**: We can model this problem as a Markov chain with the following states:

- 0: Currently, we've seen 0 consecutive tails.
- 1: Currently, we've seen 1 consecutive tail.
- 2: Currently, we've seen 2 consecutive tails.
- 3: Currently, we've seen 3 consecutive tails. This concludes the game.



Here are the hitting time equations for the number of **timesteps**, defining $\beta(S)$ as the expected number of timesteps to reach state 3, given that we're currently at state S. Note that we definitively start at state 0.

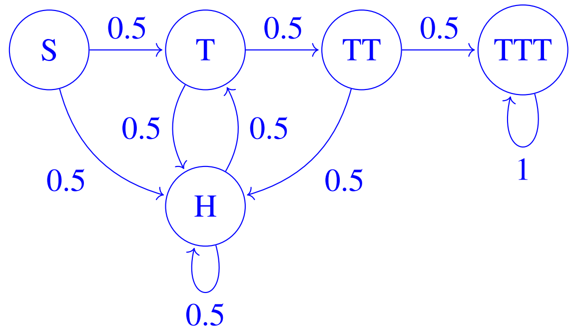
$$\beta(0) = 1 + 0.5\beta(0) + 0.5\beta(1) \tag{1}$$
$$\beta(1) = 1 + 0.5\beta(0) + 0.5\beta(2) \tag{2}$$
$$\beta(2) = 1 + 0.5\beta(0) + 0.5\beta(3) \tag{3}$$
$$\beta(3) = 0 \tag{4}$$

Solving yields $\beta(0) = 14$, $\beta(1) = 12$, $\beta(2) = 8$, and $\beta(3) = 0$. On average, it'll take us 14 timesteps (14 coin flips) before getting TTT.

Solution 2: Alternatively, we can model this problem as a Markov chain with the following states:

- S: Start state, which we are only in before flipping any coins.
- H: We see a head, which means no streak of tails currently exists.
- T: We've seen exactly one tail in a row so far.
- TT: We've seen exactly two tails in a row so far.

- TTT: We've accomplished our goal of seeing three tails in a row and stop flipping.



We can write the first step equations and solve for $\beta(S)$. The equations are as follows:

$$\beta(S) = 1 + 0.5\beta(T) + 0.5\beta(H) \tag{5}$$
$$\beta(H) = 1 + 0.5\beta(H) + 0.5\beta(T) \tag{6}$$
$$\beta(T) = 1 + 0.5\beta(TT) + 0.5\beta(H) \tag{7}$$
$$\beta(TT) = 1 + 0.5\beta(H) + 0.5\beta(TTT) \tag{8}$$
$$\beta(TTT) = 0 \tag{9}$$

From (6), we see that

$$0.5\beta(H) = 1 + 0.5\beta(T)$$

and can substitute that into (7) to get

$$0.5\beta(T) = 0.5\beta(TT) + 2.$$

Substituting this into (8), we can deduce that $\beta(TT) = 8$. This allows us to conclude that $\beta(T) = 10$, $\beta(H) = 12$, and $\beta(S) = 14$.

(b) **Solution 1**: Now, we can modify the first step equations from (a) and solve for $\beta_H(S)$ defined as the expected number of heads we see before we reach state 3, given that we're currently at state S. We will provide some motivation for how to set up the equations. In the previous equations, we had $\beta(0) = 1 + 0.5\beta(0) + 0.5\beta(1)$ which represents that we take a timestep, and move into either state 0 or state 1 with equal probability by flipping a coin. We can think of the same equation in an alternate viewpoint: consider instead flipping the coin to decide whether to move into state 0 or state 1, and in either case, we took a timestep, so we add 1 to our counter. Thus, the same equation can be rewritten as $\beta(0) = 0.5(1 + \beta(0)) + 0.5(1 + \beta(1))$.

Now consider how we may edit these equations to account for the number of **heads**, instead of the number of **timesteps**. We can modify the same equation to be $\beta_H(0) = 0.5(1 + \beta_H(0)) + 0.5(\beta_H(1))$ we only add 1 to our counter if we move into state 0, and we do not add anything to our counter if we move into state 1. We can set up the whole system of equations as follows:

$$\beta_H(0) = 0.5(1+\beta_H(0)) + 0.5(\beta_H(1)) \quad (10)$$
$$\beta_H(1) = 0.5(1+\beta_H(0)) + 0.5(\beta_H(2)) \quad (11)$$
$$\beta_H(2) = 0.5(1+\beta_H(0)) + 0.5(\beta_H(3)) \quad (12)$$
$$\beta_H(3) = 0 \quad (13)$$

Solving these equations yields $\beta_H(0) = 7$, $\beta_H(1) = 6$, $\beta_H(2) = 4$. This is the same answer we got before, resulting in an expected number of 7 heads before we see three tails in a row. On average, we expect to see 7 heads before flipping three tails in a row.

Solution 2:

We can write the first step equations and solve for $\beta(S)$, only counting heads that we see since we are not looking for the total number of flips. The equations are as follows:

$$\beta(S) = 0.5\beta(T) + 0.5\beta(H) \quad (14)$$
$$\beta(H) = 1 + 0.5\beta(H) + 0.5\beta(T) \quad (15)$$
$$\beta(T) = 0.5\beta(TT) + 0.5\beta(H) \quad (16)$$
$$\beta(TT) = 0.5\beta(H) + 0.5\beta(TTT) \quad (17)$$
$$\beta(TTT) = 0 \quad (18)$$

From (15), we see that

$$0.5\beta(H) = 1 + 0.5\beta(T)$$

and can substitute that into (16) to get

$$0.5\beta(T) = 0.5\beta(TT) + 1.$$

Substituting this into (17), we can deduce that $\beta(TT) = 4$. This allows us to conclude that $\beta(T) = 6$, $\beta(H) = 8$, and $\beta(S) = 7$.

Note on Symmetry: You may have noticed that the expected number of heads (7) we see before we see three tails in a row is half the expected number of timesteps (14). We can't directly say that this follows as a consequence of heads and tails being symmetric, as not every coin flip in our experiment ends up having the same distribution when conditioned on the fact that we end on TTT. For example, if we end on TTT, we know that the last three flips must all be tails, and have no chance of being heads. However, symmetry still ends up being applicable. What if we were to revisit the previous set of equations, but instead calculate the expected number of **tails** we see before we see three tails in a row? We can set up the equations as follows:

$$\beta_T(0) = 0.5(\beta_T(0)) + 0.5(1+\beta_T(1)) \quad (19)$$
$$\beta_T(1) = 0.5(\beta_T(0)) + 0.5(1+\beta_T(2)) \quad (20)$$
$$\beta_T(2) = 0.5(\beta_T(0)) + 0.5(1+\beta_T(3)) \quad (21)$$
$$\beta_T(3) = 0 \quad (22)$$

The $(+1)$ term ends up being applied to the latter term, as we are counting tails instead of heads. Now, we notice that these two sets of equations are actually the exact same, when we distribute out the $0.5*1$ term from every equation. Thus, the expected number of tails we see before we see three tails in a row is the same as the expected number of heads we see before we see three tails in a row, which is 7.