CS 70 Discrete Mathematics and Probability Theory
Summer 2021 **Discussion 07C**

1 Markov Chain Basics

A Markov chain is a sequence of random variables X_n, $n = 0,1,2,\ldots$. Here is one interpretation of a Markov chain: X_n is the state of a particle at time n. At each time step, the particle can jump to another state. Formally, a Markov chain satisfies the Markov property:

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i), \tag{1}$$

for all n, and for all sequences of states $i_0, \ldots, i_{n-1}, i, j$. In other words, the Markov chain does not have any memory; the transition probability only depends on the current state, and not the history of states that have been visited in the past.

(a) In lecture, we learned that we can specify Markov chains by providing three ingredients: $\mathscr{X}$, P, and π_0. What do these represent, and what properties must they satisfy?

(b) If we specify $\mathscr{X}$, P, and π_0, we are implicitly defining a sequence of random variables X_n, $n = 0,1,2,\ldots$, that satisfies (1). Explain why this is true.

(c) Calculate $\mathbb{P}(X_1 = j)$ in terms of π_0 and P. Then, express your answer in matrix notation. What is the formula for $\mathbb{P}(X_n = j)$ in matrix form?

Solution:

(a) $\mathscr{X}$ is the set of states, which is the range of possible values for X_n. In this course, we only consider finite $\mathscr{X}$.

P contains the transition probabilities. $P(i,j)$ is the probability of transitioning from state i to state j. It must satisfy $\sum_{j\in\mathscr{X}} P(i,j) = 1 \; \forall i \in \mathscr{X}$, which says that the probability that *some* transition occurs must be 1. Also, the entries must be non-negative: $P(i,j) \geq 0 \; \forall i,j \in \mathscr{X}$. A matrix satisfying these two properties is called a stochastic matrix.

Note that we allow states to transition to themselves, i.e. it is possible for $P(i,i) > 0$.

π_0 is the initial distribution, that is, $\pi_0(i) = \mathbb{P}(X_0 = i)$. Similarly, we let π_n be the distribution of X_n. Since π_0 is a probability distribution, its entries must be non-negative and $\sum_{i\in\mathscr{X}} \pi_0(i) = 1$.

(b) The sequence of random variables X_n, $n = 0,1,2,\ldots$, is defined in the following way:

- X_0 has distribution π_0, i.e. $\mathbb{P}(X_0 = i) = \pi_0(i)$.
- X_{n+1} has distribution given by

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i) = P(i,j),$$

for all $n = 0,1,2,\ldots$.

It is important to realize the connection between the Markov property (1) and the transition matrix P. P contains information about the transition probabilities in one step. If the Markov property did not hold, then P would not be enough to specify the distribution of X_{n+1}. Conversely, if we only specify P, then we are implicitly assuming that the transition probabilities do not depend on anything other than the current state. Note that this convention is different from what EE16A uses, if you have taken that class/are taking it right now.

(c) By the Law of Total Probability,

$$\mathbb{P}(X_1 = j) = \sum_{i \in \mathscr{X}} \mathbb{P}(X_1 = j, X_0 = i) = \sum_{i \in \mathscr{X}} \mathbb{P}(X_0 = i)\mathbb{P}(X_1 = j \mid X_0 = i) = \sum_{i \in \mathscr{X}} \pi_0(i)P(i, j).$$

If we write $\pi_1(j) = \mathbb{P}(X_1 = j)$ and π_0 as row vectors, then in matrix notation we have

$$\pi_1 = \pi_0 P.$$

The effect of a transition is right-multiplication by P. After n time steps, we have

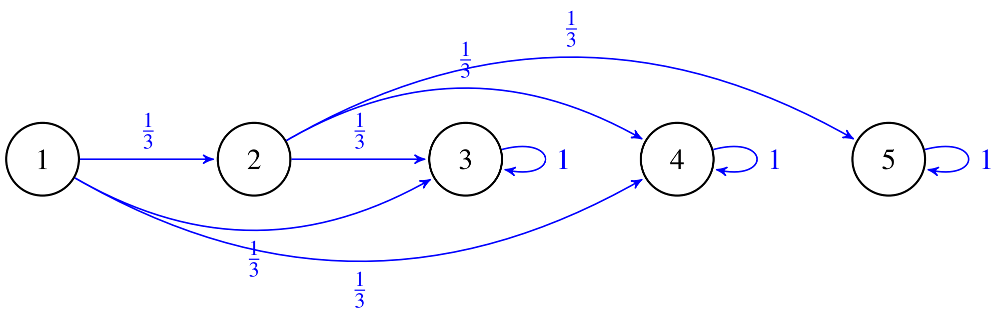
$$\pi_n = \pi_0 P^n.$$

At this point, it should be mentioned that many calculations involving Markov chains are very naturally expressed with the language of matrices. Consequently, Markov chains are very well-suited for computers, which is one of the reasons why Markov chain models are so popular in practice.

2 Skipping Stones

We consider a simple Markov chain model for skipping stones on a river, but with a twist: instead of trying to make the stone travel as far as possible, you want the stone to hit a target. Let the set of states be $\mathscr{X} = \{1,2,3,4,5\}$. State 3 represents the target, while states 4 and 5 indicate that you have overshot your target. Assume that from states 1 and 2, the stone is equally likely to skip forward one, two, or three steps forward. If the stone starts from state 1, compute the probability of reaching our target before overshooting, i.e. the probability of $\{3\}$ before $\{4,5\}$.

Solution: Here is the Markov Chain we are working with:



Let $\alpha(i)$ denote the probability of reaching the target before overshooting, starting at state i. Then:

$$\begin{aligned}
\alpha(5) &= 0 \\
\alpha(4) &= 0 \\
\alpha(3) &= 1 \\
\alpha(2) &= \frac{1}{3}\alpha(3) + \frac{1}{3}\alpha(4) + \frac{1}{3}\alpha(5) = \frac{1}{3} \\
\alpha(1) &= \frac{1}{3}\alpha(2) + \frac{1}{3}\alpha(3) + \frac{1}{3}\alpha(4) = \frac{1}{9} + \frac{1}{3}
\end{aligned}$$

Therefore, $\alpha(1) = 1/9 + 1/3 = 4/9$.

3 Consecutive Flips

Suppose you are flipping a fair coin (one Head and one Tail) until you get the same side 3 times (Heads, Heads, Heads) or (Tails, Tails, Tails) in a row.

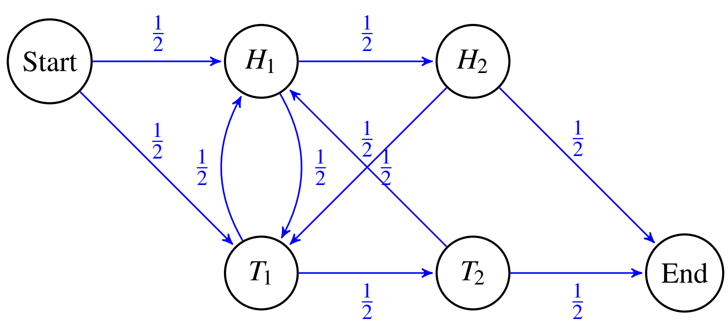
(a) Construct an Markov chain that describes the situation with a start state and end state.

(b) Given that you have flipped a (Tails, Heads) so far, what is the expected number of flips to see the same side three times?

(c) What is the expected number of flips to see the same side three times, beginning at the start state?

Solution:

(a) The appropriate Markov chain has 6 states: Start, H_1, H_2, T_1, T_2, and End.
For starting node, there is an outgoing edge to H_1 and T_1, each with equal probability of $1/2$.
For H_1, there is an outgoing edge to H_2 and T_1, each with equal probability of $1/2$.
For H_2, there is an outgoing edge to End and T_1, each with equal probability of $1/2$.
For T_1, there is an outgoing edge to H_1 and T_2, each with equal probability of $1/2$.
For T_2, there is an outgoing edge to H_1 and End, each with equal probability of $1/2$.



Note that you can merge H_1 and T_1 into one single state S_1; and merge H_2 and T_2 into a single state S_2 to obtain a 4-state Markov chain that also describes the situation. In this case, S_1 represents the state where you saw the side you just flipped exactly "one time in a row"—meaning that the last flip (if it exists) was the other side; S_2 represents the state where you saw the side you flipped exactly two times in a row.

(b) If you got a Tails and then a Heads, you are currently in the H_1 state. Thus, we must calculate the expected number of flips to end from H_1. Thus we will do this with a system of equations. Since we are not trying to solve for the starting state, we have 5 unknowns that depend on 5 linearly independent equations. Let $\beta(i)$ be the expected number of flips to reach the end state starting from state i. Then we have:

$$\begin{aligned}
\beta(H_1) &= 1 + 0.5\beta(T_1) + 0.5\beta(H_2) \\
\beta(H_2) &= 1 + 0.5\beta(\text{End}) + 0.5\beta(T_1) \\
\beta(T_1) &= 1 + 0.5\beta(T_2) + 0.5\beta(H_1) \\
\beta(T_2) &= 1 + 0.5\beta(\text{End}) + 0.5\beta(H_1) \\
\beta(\text{End}) &= 0
\end{aligned}$$

If we solve this system of equations, we get $\beta(H_1) = 6, \beta(H_2) = 4, \beta(T_1) = 6, \beta(T_2) = 4$.

(c) $\beta(S) = 1 + 0.5\beta(H_1) + 0.5\beta(T_1) = 1 + 0.5 \cdot 6 + 0.5 \cdot 6 = 7$.