CS 70 Discrete Mathematics and Probability Theory
Spring 2025 Rao

Discussion 14A

1 Markov Chains Intro I

Note 22

A **Markov chain** models an experiment with states, transitioning between states with some probability. A Markov chain is uniquely defined with the following variables:

- $\mathscr{X}$ is the set of possible states in the Markov chain. For this course, we'll only be working with Markov chains with a finite state space.
- X_n is a random variable denoting the state of the Markov chain at timestep n.
- P is the transition matrix. The element row i and column j in the matrix is defined as

  $$P(i,j) = \mathbb{P}[X_{n+1} = j \mid X_n = i].$$

  In particular, this is the probability that we transition from state i to state j.
- π_0 is the initial distribution; it is a row vector, where $\pi_0(i) = \mathbb{P}[X_0 = i]$. (Similarly, π_n is the distribution of states at timestep n; we have $\pi_n(i) = \mathbb{P}[X_n = i]$.)

Markov chains also have the **Markov property**:

$$\mathbb{P}[X_{n+1} = j \mid X_n = i, X_{n-1} = a_{n-1}, \ldots, X_0 = a_0] = \mathbb{P}[X_{n+1} = j \mid X_n = i].$$

That is, the next state depends only on the current state, and not on any prior states (this is also known as the memoryless property of Markov chains).

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.)

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*.

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.

(a) Consider the transition matrix P of a Markov chain.

(i) Is it always true that every *row* of P sums to the same value? If so, state this value and briefly explain why this makes sense. If not, briefly explain why.

(ii) Is it always true that every *column* of P sums to the same value? If so, state this value and briefly explain why this makes sense. If not, briefly explain why.

(b) Compute $\mathbb{P}[X_1 = j]$ in terms of π_0 and P. Then, express your answer in matrix notation—that is, give an expression for the row vector π_1, where $\pi_1(j) = \mathbb{P}[X_1 = j]$. Generalize your answer to express π_n in matrix form in terms of n, π_0, and P.

(c) Note that we only need to provide $\mathscr{X}$, P, and π_0 in order to uniquely define a Markov chain; the random variables X_n are implicitly defined.

(i) Explain how you can compute the distributions of the random variables X_n for $n \geq 0$ using only these parameters. (*Hint*: Part (b) can be helpful.)

(ii) The Markov property is also implicit in this definition of a Markov chain. If the Markov property *does not hold*, are $\mathscr{X}$, P, and π_0 sufficient to compute the distributions of X_n for $n \geq 0$? Justify your answer.

Solution:

(a) (i) Yes, every row must sum to 1. Note that the element at row i and column j gives the probability of transitioning from state i to state j; the sum of the elements of a row gives us the sum of all transition probabilities *out* of state i. The fact that this must sum to 1 means that we will always transition from any given state to *some* next state—we must do something at every timestep.

(ii) No, there are no restrictions on the sum of each column. The sum here would represent the sum of all transition probabilities *into* a state j, which has no inherent restrictions; these probabilities depend on the starting state, not the ending state.

(b) 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.

(c) (i) The important insight here is that π_n is exactly the distribution of X_n, where X_n takes on values in $\mathscr{X}$. We can compute the distribution of X_n as $\mathbb{P}[X_n = i] = \pi_n(i)$, where $\pi_n = \pi_0 P^n$.

(ii) If the Markov property does not hold, then P would not be sufficient to determine π_n from π_{n-1}; we'd need to know additional information about how the transition probabilities depend on the entire history of states.

2 Can it be a Markov Chain?

Note 22

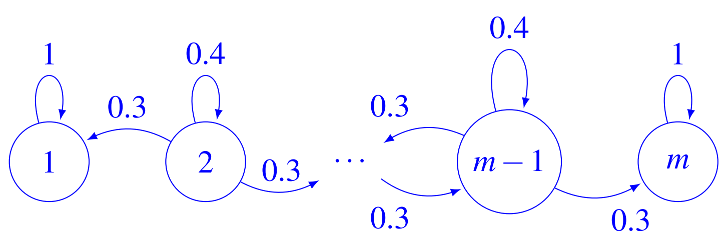
(a) A fly flies in a straight line in unit-length increments. Each second it moves to the left with probability 0.3, right with probability 0.3, and stays put with probability 0.4. There are two spiders at positions 1 and m and if the fly lands in either of those positions it is captured.

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. 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) No, 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.

3 Allen's Umbrella Setup

Note 22

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,

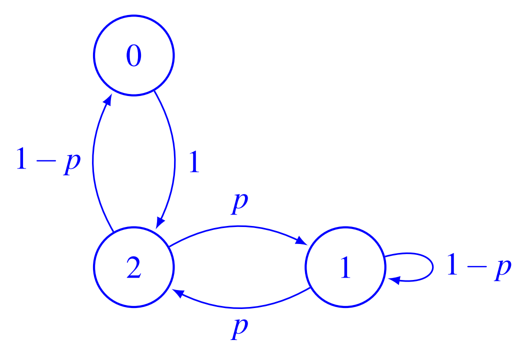
$$\mathbb{P}[X_{n+1} = 2 \mid X_n = 0] = 1.$$

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

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