UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 07
Spring 2023

1. **Mutual Information for Markov Chain**

   In the proof of Homework 06 Q4, we stated without proof the fact that $H(X \mid Y) \leq H(X \mid \hat{X})$, where $\hat{X} = g(Y)$. Here, we will explore why this inequality is true. We define the *conditional mutual information* between random variables X and Y given Z to be

$$I(X;Y \mid Z) \coloneqq \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y \mid z)}{p(x \mid z)\, p(y \mid z)}.$$

   a. Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain. Show that $I(X_{n-1}; X_{n+1} \mid X_n) = 0$ for any $n \geq 1$.
   b. Give an interpretation of part a.
   c. Show that $I(X;Y \mid Z) = H(X \mid Z) - H(X \mid Y, Z)$. Returning to the setting of Homework 06 Q4, conclude that $H(X \mid Y) \leq H(X \mid \hat{X})$.
   Hint: Show that $I(X; \hat{X} \mid Y) = 0$ using part a.

Solution:

a. By the Markov property, $X = X_{n-1}$ and $Y = X_{n+1}$ are conditionally independent given $Z = X_n$. That is, $p(X \mid Z) \cdot p(Y \mid Z) = p(X, Y \mid Z)$. Then

$$I(X;Y \mid Z) = \mathbb{E}\left(\log \frac{p(X,Y \mid Z)}{p(X \mid Z)\, p(Y \mid Z)}\right) = \mathbb{E}(\log 1) = 0.$$

b. Given the current state of a Markov chain, no information can be gained about the past by observing the future, and vice versa.

c. As we have seen, by the linearity of expectation,

$$\begin{aligned} I(X;Y \mid Z) &= \mathbb{E}\left(\log \frac{p(X,Y \mid Z)}{p(X \mid Z)\, p(Y \mid Z)}\right) \\ &= \mathbb{E}(-\log p(X \mid Z)) + \mathbb{E}(\log p(X \mid Y, Z)) \\ &= H(X \mid Z) - H(X \mid Y, Z). \end{aligned}$$

Now, by part a, because X and $\hat{X} = g(Y)$ are conditionally independent given Y, we have $I(X; \hat{X} \mid Y) = 0$, or $H(X \mid Y) = H(X \mid \hat{X}, Y)$, which also equals $H(X \mid \hat{X}) - I(X; Y \mid \hat{X})$. Conditional mutual information is nonnegative by Jensen's inequality, and therefore $H(X \mid Y) \leq H(X \mid \hat{X})$.

2. **Umbrellas**

A professor has n umbrellas, $n \geq 1$. Every morning, she commutes from her home to her office, and every night she commutes from her office back home. On every commute, if it is raining outside and there is at least one umbrella at her starting location, she takes an umbrella with her; otherwise, she does not take any umbrellas.

Assume that on each commute, it rains with probability $p \in (0,1)$, independently of all other times. Give the state space and transition probabilities for the Markov chain corresponding to the number of umbrellas the professor has at her current location.

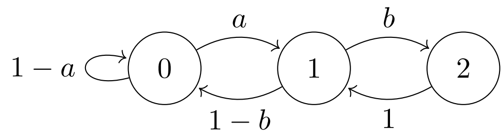
Solution: The state space is $\mathcal{X} \coloneqq \{0, \ldots, n\}$. Suppose that the professor is in state $i \in \mathcal{X}$. If it does not rain, she will land on state $n-i$; otherwise, she will land on state $\max\{n-i+1, n\}$. Thus, for $i \in \mathcal{X} \setminus \{0\}$,

$$\begin{aligned} P(i, n-i) &= 1-p, \\ P(i, n-i+1) &= p, \end{aligned}$$

and $P(0,n) = 1$.

3. **Three-State Chain**

Consider the following Markov chain, where $0 < a, b < 1$.



a. Calculate $\mathbb{P}(X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 1 \mid X_0 = 0)$.

b. Show that the Markov chain is irreducible and aperiodic.

c. Find the invariant or stationary distribution.

Solution:

a. By the Markov property, this probability is

$$P(0,1) \cdot P(1,0) \cdot P(0,0) \cdot P(0,1) = a \cdot (1-b) \cdot (1-a) \cdot a = a^2(1-a)(1-b).$$

b. The chain is irreducible because its transition diagram is strongly connected — there is a path from any state to any other state — and it is aperiodic because there is a self-loop.

c. To find the stationary distribution, let us solve the balance equations:

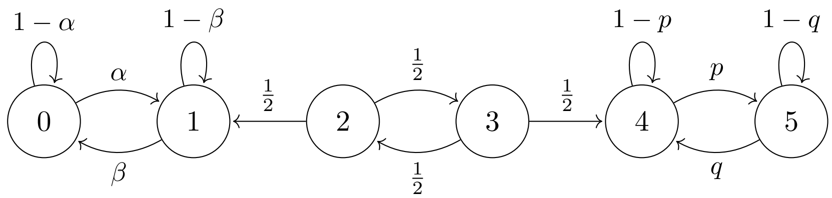
$$\pi(2) = b\pi(1),\ \pi(1) = a\pi(0) + \pi(2),\ \pi(0) + \pi(1) + \pi(2) = 1,$$

from which we find the solution

$$\begin{bmatrix}\pi(0) & \pi(1) & \pi(2)\end{bmatrix} = \frac{1}{1-b+a+ab}\begin{bmatrix}1-b & a & ab\end{bmatrix}.$$

4. **Reducible Markov Chain**

Consider the following Markov chain, where $\alpha, \beta, p, q \in (0, 1)$.



a. Find all the recurrent and transient classes.

b. Given that we start in state 2, what is the probability we reach state 0 before state 5?

c. What are all of the possible stationary distributions of this chain?
   Hint: Consider the recurrent classes.

d. Suppose we start with initial distribution $\pi_0 := \begin{bmatrix} 0 & 0 & \gamma & 1-\gamma & 0 & 0 \end{bmatrix}$ for some $\gamma \in [0, 1]$. Does the distribution of the chain converge, and if so, to what?

Solution:

a. The classes are $\{0, 1\}$ (recurrent), $\{2, 3\}$ (transient), and $\{4, 5\}$ (recurrent).

b. Let T_0 and T_5 denote the time it takes to reach states 0 and 5 respectively. Note that exactly one of T_0 and T_5 will be finite. We can set up first-step equations to compute the hitting probability $\mathbb{P}_2(T_0 < T_5) = \mathbb{P}(T_0 < T_5 \mid X_0 = 2)$:

$$\mathbb{P}_2(T_0 < T_5) = \frac{1}{2} + \frac{1}{2}\mathbb{P}_3(T_0 < T_5)$$
$$\mathbb{P}_3(T_0 < T_5) = \frac{1}{2}\mathbb{P}_2(T_0 < T_5).$$

   From this, we find that $\mathbb{P}_2(T_0 < T_5) = \frac{2}{3}$.

c. No transient state can support nonzero probability mass at stationarity, so any stationary distribution must be supported on the states $\{0, 1, 4, 5\}$. Now, if we restrict our attention to only the states $\{0, 1\}$, we have an irreducible chain with stationary distribution

$$\pi_1 := \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \alpha \end{bmatrix}.$$

   Similarly, the states $\{4, 5\}$ form an irreducible chain with stationary distribution

$$\pi_2 := \frac{1}{p+q}\begin{bmatrix} q & p \end{bmatrix}.$$

   Any stationary distribution for the entire chain will be a convex combination of π_1 and π_2, depending on the total amount of stationary mass in each recurrent class. Explicitly, the stationary distributions are of the form

$$\pi = \begin{bmatrix} c\dfrac{\beta}{\alpha+\beta} & c\dfrac{\alpha}{\alpha+\beta} & 0 & 0 & (1-c)\dfrac{q}{p+q} & (1-c)\dfrac{p}{p+q} \end{bmatrix}$$

   for $c \in [0, 1]$.

d. The distribution will indeed converge, even without irreducibility. Intuitively, probability mass gradually leaves the transient states $\{2,3\}$ until eventually, all of the probability mass is supported on the recurrent states. The two recurrent classes then each settle into their own comfortable equilibrium.

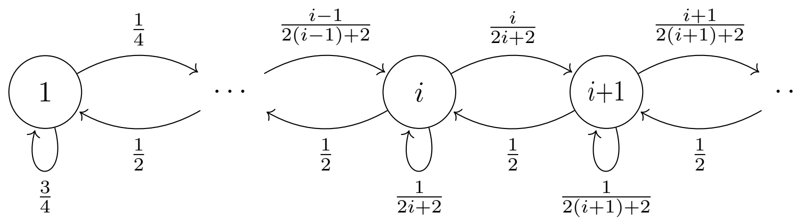
   Let us use the previous parts to find the limiting distribution. We start in state 2 with probability γ, and we end up in the recurrent class $\{0,1\}$ with further probability $\frac{2}{3}$. By symmetry, the probability that we end in $\{0,1\}$ starting from state 3 is $\frac{1}{3}$. Thus, the total probability mass which settles into $\{0,1\}$ is

$$\frac{2\gamma}{3} + \frac{1-\gamma}{3} = \frac{1}{3} + \frac{\gamma}{3},$$

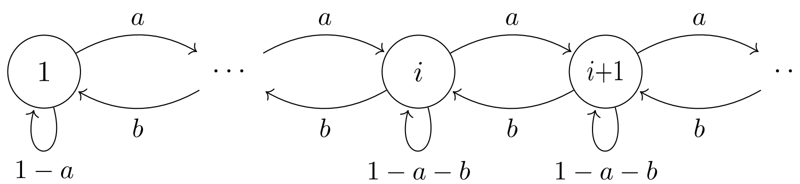
   and the probability mass which settles in $\{4,5\}$ is $\frac{2}{3} - \frac{\gamma}{3}$. Therefore the chain converges to the stationary distribution found in part c with parameter $c = \frac{1}{3} + \frac{\gamma}{3}$.

5. **Markov Chains with Countably Infinite State Space**

a. Show that the Markov chain with state space $\mathbb{Z}^+$ and the following transition diagram is not positive recurrent. Also find the expected time it takes to return to state i starting from i for any $i \in \mathbb{Z}^+$.



b. Let $0 < a < b < a + b \le 1$. Consider now the Markov chain with state space $\mathbb{Z}^+$ and the following transition diagram:



Show that a stationary distribution of this Markov chain is given by

$$\pi(i) = \left(\frac{a}{b}\right)^{i-1}\left(1 - \frac{a}{b}\right).$$

Also compute the expected time it takes to return to state i starting from i.

Solution:

a. By the big theorem, an irreducible Markov chain is not positive recurrent if it does not have a stationary distribution. The given chain is a birth-death chain, so any stationary distribution π will satisfy the detailed balance equations, namely

$$\pi(i) \cdot P(i, i+1) = \pi(i+1) \cdot P(i+1, i) \quad \text{for all } i \in \mathbb{Z}^+.$$

With the given transition probabilities, this means

$$\pi(i+1) = \frac{i}{i+1}\pi(i) = \frac{i}{i+1}\frac{i-1}{i}\pi(i-1) = \cdots = \frac{1}{i+1}\pi(1).$$

A stationary distribution must also satisfy $\sum_{i=1}^{\infty} \pi(i) = 1$ in order to be a valid probability distribution. However,

$$\sum_{i=1}^{\infty} \pi(i) = \pi(1) \sum_{i=1}^{\infty} \frac{1}{i} = \pi(1) \cdot \infty$$

means that it is impossible to assign a value to $\pi(1)$ such that $\sum_{i=1}^{\infty} \pi(i) = 1$. Therefore this chain does not admit a stationary distribution. The expected return time for any state is ∞ as the chain is not positive recurrent.

b. We first observe that

$$\left(\frac{a}{b}\right)^{i-1}\left(1-\frac{a}{b}\right) \geq 0,$$
$$\sum_{i=1}^{\infty}\left(\frac{a}{b}\right)^{i-1}\left(1-\frac{a}{b}\right) = \frac{1}{1-\frac{a}{b}}\left(1-\frac{a}{b}\right) = 1,$$

so π is a valid probability distribution. This Markov chain is also a birth-death chain, so we are left to verify that π satisfies the detailed balance equations:

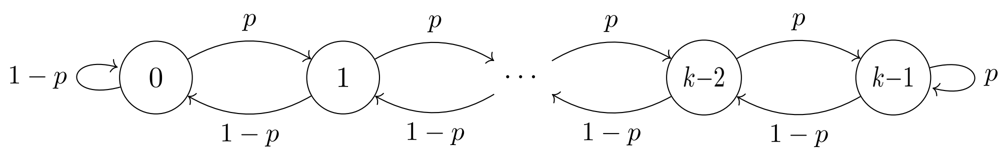
$$\pi(i)\cdot P(i,i+1) = \left(\frac{a}{b}\right)^{i-1}\left(1-\frac{a}{b}\right)a = \left(\frac{a}{b}\right)^{i}\left(1-\frac{a}{b}\right)b = \pi(i+1)\cdot P(i+1,i).$$

In this case, the expected return time is $\frac{1}{\pi(i)} = (\frac{b}{a})^{i-1}\frac{b}{b-a}$.

Remark. Part b shows that the stationary distribution of the infinite birth-death chain with bias p is Geometric$(1-\frac{p}{q})$. The last line is a rather magical result — once we have the stationary distribution, we immediately know the expected amount of time it takes between successive visits to any given state!

6. **Finite Random Walk**

Let $0 < p < 1$, and consider the following finite *random walk* with bias p on $\mathcal{X} = \{0, \dots, k-1\}$, also known as the finite *birth-death chain*.



a. Find the stationary distribution π.
Hint: Write $q = 1 - p$ and define $r := \frac{p}{q}$. Be careful when $r = 1$.

b. Find the limit of $\pi(0)$ and $\pi(k-1)$, as functions of k, as $k \to \infty$.

Solution:

a. Let us solve the detailed balance equations:

$$p \cdot \pi(i-1) = q \cdot \pi(i) \quad \text{for all } i = 1, \dots, k-1,$$

or $\pi(i) = r\pi(i-1)$. Iterating this recurrence relation, we have $\pi(i) = r^i \pi(0)$, so

$$\sum_{i=0}^{k-1} \pi(i) = \pi(0) \sum_{i=0}^{k-1} r^i = \pi(0) \frac{1-r^k}{1-r} = 1.$$

We can then solve for $\pi(0) = \frac{1-r}{1-r^k}$ and $\pi(i) = r^i \frac{1-r}{1-r^k}$ in general. However, this formula is undefined when $r = 1$, or $p = \frac{1}{2}$. Instead, we find that $\pi(i) \equiv \frac{1}{k}$ for all $i = 0, \dots, k-1$. In short, the stationary distribution is given by

$$\pi(i) = \begin{cases} r^i \frac{1-r}{1-r^k} & \text{if } p \neq \frac{1}{2} \\ \frac{1}{k} & \text{if } p = \frac{1}{2}. \end{cases}$$

b. First, the limit of $\pi(0)$ is $\lim_{k\to\infty} \frac{1}{k} = 0$ if $r = 1$ and $(1-r)\lim_{k\to\infty} \frac{1}{1-r^k}$ if $r \neq 1$. Now, $1/(1-r^k) \to 1$ if $r < 1$ and $\to 0$ if $r > 1$. Therefore

$$\lim_{k\to\infty} \pi(0) = \begin{cases} 1-r & \text{if } r < 1 \\ 0 & \text{if } r \geq 1. \end{cases}$$

For $\pi(k-1)$, we know that the limit is also 0 if $r = 1$, so let $r \neq 1$. Then

$$\lim_{k\to\infty} \pi(k-1) = \lim_{k\to\infty} \frac{r^{k-1} - r^k}{1 - r^k} = \lim_{k\to\infty} \frac{1-r}{\frac{1}{r^{k-1}} - r} = \begin{cases} 0 & \text{if } r < 1 \\ \frac{r-1}{r} & \text{if } r > 1. \end{cases}$$

Therefore the limit is

$$\lim_{k\to\infty} \pi(k-1) = \begin{cases} 0 & \text{if } r \leq 1 \\ 1 - \frac{1}{r} & \text{if } r > 1. \end{cases}$$