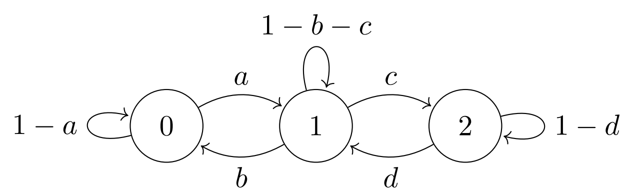
UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 08
Spring 2024

1. **Markov Chain Big Theorem**

For this problem we will consider the following three-state chain and illustrate the ideas behind the Markov chain convergence theorem. Here, $a, b, c, d \in (0, 1)$.



a. Let $T_0 = \min\{n \in \mathbb{Z}_+ : X_n = 0\}$ be the first passage time to state 0. Let $\mu_y := \mathbb{E}_0[\sum_{n=0}^{T_0-1} \mathbb{1}\{X_n = y\}]$ for $y = 0, 1, 2$ be the mean number of visits to state y, starting at 0 and ending right before we return to 0. Explain why $\mu = \mu P$.

b. Therefore, if we define π to be μ after we normalize it so that the entries sum to 1, π is a stationary distribution. Why is π unique?

c. Now deduce that $\pi_0 = 1/\mathbb{E}_0[T_0]$. In words, $\mathbb{E}_0[T_0]$ is the mean return time from state 0 to itself.

d. Explain why the fraction of times $\sum_{m=1}^{n} \mathbb{1}\{X_m = 0\}$, where n is a positive integer, converges a.s. to π_0 as $n \to \infty$. (Hint: Define $T_0^{(1)} := T_0$ and for integers $k \geq 2$, define

$$T_0^{(k)} = \min\{n > T_0^{(k-1)} : X_n = 0\} - T_0^{(k-1)}$$

to be the additional time it takes to return to 0 for the kth time. Then $T_0^{(1)}, T_0^{(2)}, T_0^{(3)}, \ldots$ are i.i.d. and one can apply the SLLN.)

e. Consider two copies of the above chain $(X_n, Y_n)_{n\in\mathbb{N}}$, where the chains move independently of each other, Y_0 is picked from the stationary distribution, and X_0 is started from any fixed state x. Explain why the two chains will meet after a finite time, and think about why this implies that the chain started from state x converges in distribution to the stationary distribution π.

2. **Two-State Chain with Linear Algebra**

Consider the Markov chain $(X_n,\ n \in \mathbb{N})$, shown in Figure 1, where $\alpha, \beta \in (0, 1)$.

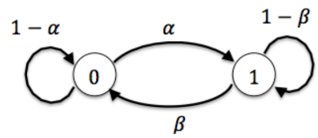


Figure 1: Markov chain for this Problem

a. Find the probability transition matrix P.

b. Find two real numbers λ_1 and λ_2 such that there exists two non-zero vectors u_1 and u_2 such that $Pu_i = \lambda_i u_i$ for $i = 1, 2$. Further, show that P can be written as $P = U \Lambda U^{-1}$, where U and Λ are 2×2 matrices and Λ is a diagonal matrix.

   Hint: This is called the eigendecomposition of a matrix.

c. Find P^n in terms of U and Λ for each $n \in \mathbb{N}$.

d. Assume that $X_0 = 0$. Use the result in part (c) to compute the PMF of X_n for all $n \in \mathbb{N}$.

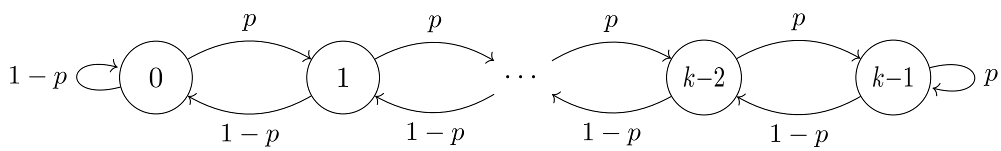
e. What does the fraction of time spent in state 0, $n^{-1} \sum_{i=1}^{n} \mathbb{1}\{X_i = 0\}$, converge to (almost surely) as $n \to \infty$?

3. **Random Walk on an Undirected Graph**

   Consider a random walk on an undirected connected finite graph (that is, define a Markov chain where the state space is the set of vertices of the graph, and at each time step, transition to a vertex chosen uniformly at random out of the neighborhood of the current vertex). What is the stationary distribution π? Your answer may depend on $\deg(v)$ (i.e., the degree of a vertex v) for some v. *Hint*: assume first that the chain is *reversible*.

4. **Finite Random Walk**

Let $0 < p < 1$, and consider the following finite *random walk* with bias p on $\mathcal{X} = \{0, \ldots, k-1\}$, 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 $k \to \infty$.

5. **Metropolis–Hastings**

We will prove some properties of the *Metropolis–Hastings* algorithm, an example of Markov Chain Monte Carlo (MCMC) sampling that you will see more of in lab. The goal of MH is to draw samples from a distribution $p(x)$; the algorithm assumes that

- We can compute $p(x)$ up to a normalizing constant C via $f(x)$, and
- We have a proposal distribution $g(x, \cdot)$.

The steps in making a transition are:

i. Propose the next state y according to the distribution $g(x, \cdot)$.

ii. Accept the proposal with probability

$$A(x, y) = \min\left\{1, \frac{f(y)}{f(x)} \frac{g(y, x)}{g(x, y)}\right\}.$$

iii. If the proposal is accepted, move the chain to y; otherwise, stay at x.

Remark. The normalizing factor $C = 1/\sum_{x\in\mathcal{X}} f(x)$ is sometimes called the *partition function*, and it can be difficult to compute for large sets $\mathcal{X}$, even if $f(x)$ is efficient to compute.

In the following, we will verify that the Metropolis–Hastings chain has stationary distribution p, and in fact approaches stationarity after running for some time, at which point we can draw samples from p by sampling from the chain.

a. The key to why Metropolis–Hastings works is the **detailed balance equations**. Suppose we have a finite irreducible Markov chain on a state space $\mathcal{X}$ with transition probability matrix P. Show that if there exists a distribution π on $\mathcal{X}$ satisfying detailed balance,

$$\pi(x)P(x, y) = \pi(y)P(y, x) \quad \text{for all } x, y \in \mathcal{X},$$

then $\pi P = \pi$ is a stationary distribution of the chain.

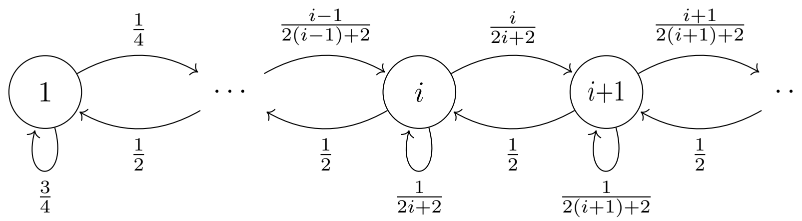
b. Returning to the Metropolis–Hastings chain, find $P(x, y)$. For simplicity, assume $x \neq y$.

c. Show that the target distribution $p(x)$ satisfies the detailed balance equations for $P(x, y)$, and conclude that $p(x)$ is the stationary distribution of the chain.

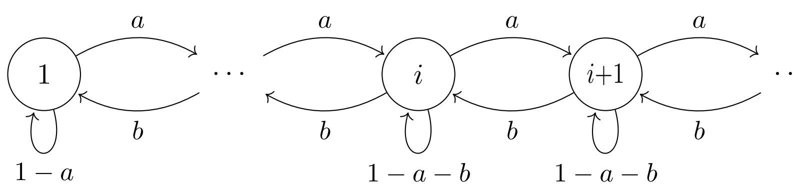
d. If the chain is aperiodic, then it will converge to the stationary distribution. If not, we can force the chain to be aperiodic by considering the **lazy chain**: on each transition, the chain decides not to move with probability $\frac{1}{2}$, independently of the propose-accept step. Explain why the lazy chain is aperiodic, and explain why the stationary distribution is the same as before.

6. **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 $0 < a < b < a + b \leq 1$. 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.