UC Berkeley
Department of Electrical Engineering and Computer Sciences

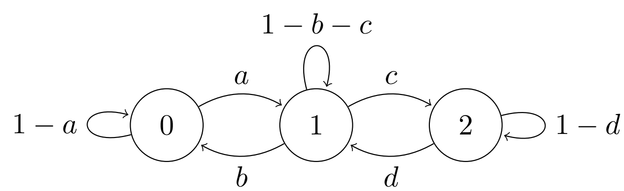
EECS 126: PROBABILITY AND RANDOM PROCESSES

Discussion 8
Fall 2017

---

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. Random Walk on the Cube

Consider the symmetric random walk on the vertices of the 3-dimensional unit cube where two vertices are connected by an edge if and only if the line connecting them is an edge of the cube. In other words, this is the random walk on the graph with 8 nodes each written as a string of 3 bits, so that the vertex set is $\{0,1\}^3$, and where two vertices are connected by an edge if and only if their corresponding bit strings differ in exactly one location.

This random walk is modified so that the nodes 000 and 111 are made absorbing.

(a) What are the communicating classes of the resulting Markov chain? For each class, determine its period, and whether it is transient or recurrent.

(b) For each transient state, what is the probability that the modified random walk started at that state gets absorbed in the state 000?

3. Hidden Markov Models

A hidden Markov model (HMM) is a Markov chain $\{X_n\}_{n=0}^{\infty}$ in which the states are considered "hidden" or "latent". In other words, we do not directly observe $\{X_n\}_{n=0}^{\infty}$. Instead, we observe $\{Y_n\}_{n=0}^{\infty}$, where $Q(x, y)$ is the probability that state x will emit observation y. π_0 is the initial distribution for the Markov chain, and P is the transition matrix.

(a) What is $\mathbb{P}(X_0 = x_0, Y_0 = y_0, \dots, X_n = x_n, Y_n = y_n)$, where n is a positive integer, $x_0, \dots, x_n$ are hidden states, and $y_0, \dots, y_n$ are observations?

(b) What is $\mathbb{P}(X_0 = x_0 \mid Y_0 = y_0)$?

(c) We observe $(y_0, \dots, y_n)$ and we would like to find the most likely sequence of hidden states $(x_0, \dots, x_n)$ which gave rise to the observations. Let

$$U(x_m, m) = \max_{x_{m+1},\dots,x_n \in \mathcal{X}} \mathbb{P}(X_m = x_m, X_{m+1:n} = x_{m+1:n}, Y_{0:n} = y_{0:n})$$

denote the largest probability for a sequence of hidden states beginning at state x_m at time $m \in \mathbb{N}$, along with the observations $(y_0, \dots, y_n)$. Develop a recursion for $U(x_m, m)$ in terms of $U(x_{m+1}, m+1)$, $x_{m+1} \in \mathcal{X}$.