UC Berkeley
Department of Electrical Engineering and Computer Sciences

EE126: Probability and Random Processes

Solutions 6
Spring 2017

Issued: Thursday, March 2, 2017

Self-graded Scores Due: 5pm, Monday, March 13, 2017
Submit your self-graded scores via the google form:
https://goo.gl/forms/O219EBs50ACR4B1N2. Make sure that you use your **SORTABLE NAME** on bCourses.

Problem 1. Consider the Markov chain with state X_n, $n \geq 0$, shown in Figure 1, where $\alpha, \beta \in (0, 1)$.

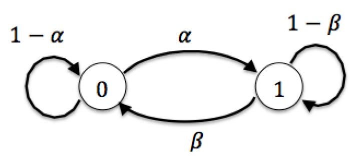

Figure 1: Markov chain for Problem 1

(a) Find the probability transition matrix P and the invariant distribution π of the Markov chain.

(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.

(d) Assume that $X_0 = 0$. Use the result in part (c) to compute the PMF of X_n for all $n \geq 0$. Verify that it converges to the invariant distribution.

Solution:

(a) The probability transition matrix is

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}.$$

By $\pi P = \pi$, we get the invariant distribution $\pi = [\frac{\beta}{\alpha+\beta} \quad \frac{\alpha}{\alpha+\beta}]$.

(b) Since $(P - \lambda_i I)x = 0$ has non-zero solution u_i, we have $\det(P - \lambda_i I) = 0$, i.e, λ_1 and λ_2 are solutions to

$$\begin{vmatrix} 1-\alpha-\lambda & \alpha \\ \beta & 1-\beta-\lambda \end{vmatrix} = \lambda^2 - (2-\alpha-\beta)\lambda + 1 - \alpha - \beta.$$

Then we get $\lambda_1 = 1$, and $\lambda_2 = 1 - \alpha - \beta$. Then we can get u_1 and u_2: $u_1 = [1 \ \ 1]^T$ and $u_2 = [\alpha \ \ -\beta]^T$. Further, we can see that if we let

$$U = [u_1 \ \ u_2] = \begin{bmatrix} 1 & \alpha \\ 1 & -\beta \end{bmatrix},$$

and

$$\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 1-\alpha-\beta \end{bmatrix},$$

we have $PU = U\Lambda$, which is equivalent to $P = U\Lambda U^{-1}$.

(c) We have

$$P^n = U\Lambda U^{-1} \cdots U\Lambda U^{-1} = U\Lambda^n U^{-1}.$$

(d) Let $\pi(n) = [\Pr(X_n = 0) \ \ \Pr(X_n = 1)]$ be the PMF of X_n. Then we have

$$\pi(n) = \pi(0)P^n = \pi(0)U\Lambda^n U^{-1}.$$

Since we have $\pi(0) = [1 \ \ 0]$, by some computation, we get

$$\pi(n) = \begin{bmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n \end{bmatrix}.$$

Then we can see that $\lim_{n\to\infty} \pi(n) = \pi$.

Problem 2. For each of the following case, explain whether $\{Y_n\}_{n\geq 0}$ satisfies the Markov property.

(a) $Y_0, X_1, X_2, \ldots$ are mutually independent discrete random variables and

$$Y_{n+1} = (Y_n + X_{n+1})^{(n+1)} \quad \text{for } n \geq 0.$$

(b) $Y_0 = 1$ and $Y_n = U_1 U_2 U_3 \ldots U_n$, where $U_1, U_2, \ldots$ are independent random variables, each uniformly distributed on the interval $[0, 1]$.

(c) Let $\{X_n\}_{n\geq 0}$ be a Markov chain with two states, -1 and 1, and the transition probabilities $P(-1, 1) = P(1, -1) = a$ for $a \in (0, 1)$. Define

$$Y_n = X_0 + X_1 + \cdots + X_n.$$

Solution:

(a) $\{Y_n\}_{n\geq 0}$ satisfies the Markov property. For any $(y_0, y_1, \ldots, y_{n+1}) \in \mathbb{R}^{n+2}$, we have

$$\begin{aligned}
& P(Y_{n+1} = y_{n+1} | Y_n = y_n, Y_{n-1} = y_{n-1}, \ldots, Y_0 = y_0) \\
= \; & P((y_n + X_{n+1})^{n+1} = y_{n+1} | Y_n = y_n, Y_{n-1} = y_{n-1}, \ldots, Y_0 = y_0) \\
= \; & P((y_n + X_{n+1})^{n+1} = y_{n+1}) \\
= \; & P(Y_{n+1} = y_{n+1} | Y_n = y_n)
\end{aligned}$$

where second equality follows from the fact that X_{n+1} is independent from $\{Y_k\}_{k=0}^n$.

(b) $\{Y_n\}_{n\geq 0}$ satisfies the Markov property. For any $(y_1, \ldots, y_{n+1}) \in [0,1]^{n+1}$, we have

$$\begin{aligned}
& P(Y_{n+1} \leq y_{n+1} | Y_n = y_n, Y_{n-1} = y_{n-1}, \ldots, Y_1 = y_1) \\
= \; & P(U_1 U_2 \ldots U_{n+1} \leq y_{n+1} | Y_n = y_n, Y_{n-1} = y_{n-1}, \ldots, Y_1 = y_1) \\
= \; & P(y_n U_{n+1} \leq y_{n+1} | Y_n = y_n, Y_{n-1} = y_{n-1}, \ldots, Y_1 = y_1) \\
= \; & P(y_n U_{n+1} \leq y_{n+1}) \\
= \; & P(Y_{n+1} \leq y_{n+1} | Y_n = y_n)
\end{aligned}$$

where third equality follows from the fact that U_{n+1} is independent from $\{Y_k\}_{k=0}^n$.

(c) If $a \neq \frac{1}{2}$, Y_n does not satisfy the Markov property. Consider the following probability:

$$\Pr(Y_4 = 3 | Y_2 = 1, Y_3 = 2) = \Pr(X_4 = 1 | X_3 = 1) = 1 - a.$$

On the other hand,

$$\Pr(Y_4 = 3 | Y_2 = 3, Y_3 = 2) = \Pr(X_4 = 1 | X_3 = -1) = a.$$

So the distribution of Y_4 given the past is not only dependent on Y_3, which shows that Markov property does not hold.
However, when $a = \frac{1}{2}$, Y_n does satisfy the Markov property. This is because when $a = \frac{1}{2}$, X_i's become i.i.d. with $\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$. Then

$$Y_n = \begin{cases} Y_{n-1} + 1 & \text{with probability } \frac{1}{2}, \\ Y_{n-1} - 1 & \text{with probability } \frac{1}{2}. \end{cases}$$

Then we know that given Y_{n-1}, Y_n is conditional independent of all the previous states $Y_0, \ldots, Y_{n-2}$.

Problem 3. (a) Find the steady-state probabilities $\pi_0, \ldots, \pi_{k-1}$ for the Markov chain in Figure 2. Express your answer in terms of the ratio $\rho = p/q$, where $q = 1 - p$. Pay particular attention to the special case $\rho = 1$.

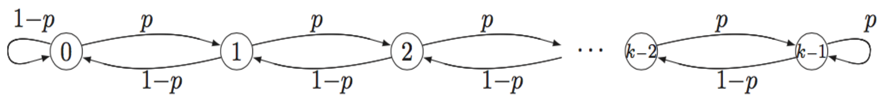

Figure 2: Markov chain for Problem 3

(b) Find the limit of π_0 as k approaches infinity; give separate answers for $\rho < 1$, $\rho = 1$, and $\rho > 1$. Find limiting values of π_{k-1} for the same cases.

Solution:

(a) Without loss of generality, we can consider $\rho \leq 1$. This is because if $\rho > 1$, we can flip the chain and get a new chain with $\rho \leq 1$. Consider the invariant distribution, using the abbreviation $q = 1 - p$, we have

$$\begin{aligned} \pi_0 &= q\pi_0 + q\pi_1 \\ \pi_j &= p\pi_{j-1} + q\pi_{j+1}; \quad \text{for } 1 \leq j \leq k-2 \\ \pi_{k-1} &= p\pi_{k-2} + p\pi_{k-1}. \end{aligned}$$

Simplifying the first equation, we get $p\pi_0 = q\pi_1$, i.e., $\pi_1 = \rho\pi_0$. Substituting $q\pi_1$ for $p\pi_0$ in the second equation, we get $\pi_1 = q\pi_1 + q\pi_2$. Simplifying the second equation, then, we get $\pi_2 = \rho\pi_1$. We can then use induction. Using the inductive hypothesis $\pi_j = \rho\pi_{j-1}$ (which has been verified for $j = 1, 2$) on $\pi_j = p\pi_{j-1} + q\pi_{j+1}$, we get

$$\pi_{j+1} = \rho\pi_j \quad \text{for } 1 \leq j \leq k-2.$$

Combining these equations, $\pi_j = \rho\pi_{j-1}$ for $1 \leq j \leq k-1$, so $\pi_j = \rho^j\pi_0$ for $1 \leq j \leq k-1$.

$$\pi_0\left(\sum_{j=0}^{k-1}\rho^j\right) = 1 \quad \text{so } \pi_0 = \frac{1-\rho}{1-\rho^k}; \quad \pi_j = \rho^j\frac{1-\rho}{1-\rho^k}.$$

For $\rho = 1$, $\rho^j = 1$ and $\pi_j = 1/k$ for $0 \leq j \leq k-1$.

(b) For state 0,

$$\lim_{k\to\infty}\pi_0 = \begin{cases} \lim_{k\to\infty}\frac{1-\rho}{1-\rho^k} = 1-\rho & \text{for } \rho < 1, \\ \lim_{k\to\infty}\frac{1}{k} = 0 & \text{for } \rho = 1, \end{cases}$$

For state $k-1$, the analogous result is

$$\lim_{k\to\infty}\pi_{k-1} = \begin{cases} \lim_{k\to\infty}\rho^{k-1}\frac{1-\rho}{1-\rho^k} = 0 & \text{for } \rho < 1, \\ \lim_{k\to\infty}\frac{1}{k} = 0 & \text{for } \rho = 1, \end{cases}$$

Problem 4. Consider the Markov chain with 6 states given in Figure 3. Suppose that $P(X_0 = 1) = 1$, and given $X_k = i$, the next state X_{k+1} is one of the two neighbors of i, selected with probability 0.5 each.

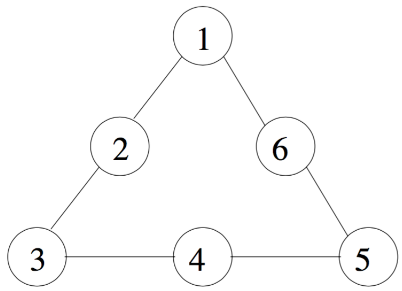

Figure 3: Markov chain for Problem 4

(a) Let τ_C be the first time $k \geq 1$ such that both states 3 and 5 have been visited by time k. Find $E[\tau_C]$.

(b) Let τ_R denote the first time $k \geq \tau_C$ such that $X_k = 1$. That is, τ_R is the first time the process returns to vertex 1 of the triangle after reaching both of the other vertices. Find $E[\tau_R]$.

Solution: Let τ_{ij} denote the first time $k \geq 1$ such that $X_{t+k} = j$ given that $X_t = i$ for any t. Let B denote the combination of nodes $\{3, 4, 5\}$, and let τ_{iB} be defined similar to τ_{ij}, that is, the first time $k \geq 1$ such that $X_{t+k} \in B$ given that $X_t = i$.

(a) Note that

$$\tau_{1B} = \begin{cases} (1 + \tau_{2B}) & \text{if next state is 2} \\ (1 + \tau_{6B}) & \text{if next state is 6} \end{cases}$$

$$\tau_{2B} = \begin{cases} 1 & \text{if next state is 3} \\ (1 + \tau_{1B}) & \text{if next state is 1} \end{cases}$$

$$\tau_{6B} = \begin{cases} 1 & \text{if next state is 5} \\ (1 + \tau_{1B}) & \text{if next state is 1} \end{cases}$$

Then, we can obtain 3 equations for $E[\tau_{1B}], E[\tau_{2B}]$ and $E[\tau_{6B}]$:

$$E[\tau_{1B}] = \frac{1}{2}(1 + E[\tau_{2B}]) + \frac{1}{2}(1 + E[\tau_{6B}])$$

$$E[\tau_{2B}] = E[\tau_{6B}] = \frac{1}{2} + \frac{1}{2}(1 + E[\tau_{1B}])$$

Solving these equations, we find $E[\tau_{1B}] = 4$.

To compute $E[\tau_C]$, we need $E[\tau_{35}] = E[\tau_{53}]$. Note that, due to symmetry, we have

$$E[\tau_{35}] = E[\tau_{15}],$$

$$E[\tau_{45}] = E[\tau_{65}].$$

Therefore, we can write

$$E[\tau_{25}] = \frac{1}{2}(1 + E[\tau_{35}]) + \frac{1}{2}(1 + E[\tau_{15}]) = 1 + E[\tau_{35}]$$

$$E[\tau_{35}] = \frac{1}{2}(1 + E[\tau_{25}]) + \frac{1}{2}(1 + E[\tau_{45}])$$

$$E[\tau_{45}] = \frac{1}{2} + \frac{1}{2}(1 + E[\tau_{35}])$$

and obtain

$$E[\tau_{35}] = 8.$$

As a result, $E[\tau_C] = E[\tau_{1B}] + E[\tau_{35}] = 12$.

(b) Note that $E[\tau_R] = E[\tau_C + \tau_{31}] = E[\tau_C + \tau_{51}]$ and due to symmetry $E[\tau_{31}] = E[\tau_{51}] = E[\tau_{35}] = 8$. Therefore, $E[\tau_R] = 20$.

Problem 5. You have a database of an infinite number of movies. Each movie has a rating that is uniformly distributed in $[0, 5]$ and you want to find two movies such that the sum of their ratings is greater than 7.5. Assume that you choose movies from the database one by one and keep the movie with the highest rating. You stop when you find that the sum of the ratings of the last movie you have chosen and the movie with the highest rating among all the previous movies is greater than 7.5. What is the expected number of movies you will have to choose?

Solution: We always remember the maximum rating seen up to now. Define it as x to be the state of the system. Then we define $\beta(x)$ to be the average remaining number of movies we have to check to get two ratings that sum to greater than 7.5. We are interested in computing $\beta(0)$ using first step equations. First note that if $x < 2.5$, then it is impossible to have two ratings that sum to larger than 7.5 in the next step. This shows that $\beta(x) = \beta(0)$ for $x < 2.5$. If $3.75 \leq x \leq 5$, then either the next movie has rating larger than $7.5 - x$ (with probability $\frac{x-2.5}{5}$) which ends the experiment, or it is less than $7.5 - x$ (with probability $\frac{7.5-x}{5}$) and since $x > 3.75$ it is also less than x. In this case, we remember x as the largest rating seen. Thus, the first step equation for $3.75 \leq x \leq 7.5$ is

$$\beta(x) = \frac{x - 2.5}{5} \times 1 + \frac{7.5 - x}{5}(1 + \beta(x))$$

which leads to $\beta(x) = \frac{5}{x-2.5}$ if $3.75 \leq x \leq 5$. If $2.5 \leq x < 3.75$, then either the next movie with rating y is such that $y > 7.5 - x$ (with probability $\frac{x-2.5}{5}$) which ends the experiment, or $y < x$ (with probability $\frac{x}{5}$), or it is larger than x but less than $7.5 - x$. For these values, $x < y < 7.5 - x$, one should remember the new largest seen rating which is y. Note that y is in an small interval of length dy with probability $\frac{dy}{5}$. Thus, the first step equation is

$$\beta(x) = \frac{x - 2.5}{5} \times 1 + \frac{x}{5}(1 + \beta(x)) + \int_x^{7.5-x} \frac{1}{5}(1 + \beta(y))dy.$$

This is an integral equation. If we differentiate both sides with respect to x we get the following differential equation.

$$\beta'(x) = \frac{1}{5} + \frac{1}{5}(1+\beta(x)) + \frac{x\beta'(x)}{5} + (-1)\frac{1}{5}(1+\beta(7.5-x)) - \frac{1}{5}(1+\beta(x)).$$

Note that for differentiating the integral we use the Leibniz rule which is as follows.

$$\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)\,dt\right) = f(x,b(x))\,b'(x) - f(x,a(x))\,a'(x) + \int_{a(x)}^{b(x)} f_x(x,t)\,dt.$$

The equation will be simplified to

$$\beta'(x) = \frac{-\beta(7.5-x)}{5-x}.$$

Since $3.75 \le 7.5 - x \le 5$ for $2.5 \le x \le 3.75$, we have that $\beta(7.5-x) = \frac{5}{5-x}$. Thus,

$$\beta'(x) = -\frac{5}{(5-x)^2}.$$

Integrating both sides we have,

$$\beta(x) = -\frac{5}{5-x} + C,$$

where C is a constant to be determined by the boundary condition of differential equation. The boundary condition here is $\beta(3.75) = \frac{5}{3.75-2.5} = 4$. Thus, $C = 8$. Finally,

$$\beta(0) = \beta(2.5) = \frac{-5}{2.5} + 8 = 6.$$

Thus the expected number of movies that we have to check is 6.

Other solutions: Any movie whose rating is less than 2.5 is clearly useless. Thus, we assume that movie ratings are uniformly distributed in $[2.5, 5.0]$, find the expected number of movies to pick, and multiply it by 2 to obtain the solution. Among the n chosen movies, denote by $X_{(1)}^n$ the maximum rating and by $X_{(2)}^n$ the second maximum rating. The (stopping) time T is defined such that $X_{(1)}^T + X_{(2)}^T > 7.5$, but $X_{(1)}^{T-1} + X_{(2)}^{T-1} \le 7.5$. Clearly, $T \ge 2$.

To simplify calculation, we normalize X_i so that it is distributed in $[0,1]$. With such normalized random variables, the (stopping) time T can be also normalized such that $X_{(1)}^T + X_{(2)}^T > 1$, but $X_{(1)}^{T-1} + X_{(2)}^{T-1} \le 1$. In other words, T is the number of (unit) uniform random variables to draw until the sum of the maximum and the second maximum is no less than 1. As T is a non-negative integer random variable,

$$\mathbb{E}[T] = \sum_{t=1}^{\infty} P(T \ge t) = 2 + \sum_{t=3}^{\infty} P(T \ge t).$$

Note that $P(T \ge t) = P(X_{(1)}^{t-1} + X_{(2)}^{t-1} \le 1)$. For any $t \ge 2$, the joint distribution of $(X_{(1)}^t, X_{(2)}^t)$ is

$$f_{X_{(1)}^t, X_{(2)}^t}(x_1, x_2) = t(t-1)x_2^{t-2}\mathbf{1}\{1 \ge x_1 \ge x_2 \ge 0\}.$$

Thus,

$$\begin{aligned}
P(X_{(1)}^t + X_{(2)}^t \leq 1) &= \int_0^1 \int_0^1 f_{X_{(1)}^t, X_{(2)}^t}(x_1, x_2)\mathbf{1}\{x_1 + x_2 \leq 1\} dx_2 dx_1 \\
&= t(t-1) \int_{x_2=0}^{\frac{1}{2}} (1 - 2x_2) x_2^{t-2} dx_2 \\
&= \frac{1}{2^{t-1}}.
\end{aligned}$$

Thus, $\mathbb{E}[T] = 2 + \sum_{t=3}^{\infty} \frac{1}{2^{t-2}} = 3$, giving us 6 as the final answer for the original problem.