Solution 4

Saliba, March 13, 2019

Exercise 1. (**Markov Chain in a library**) In a library with n books, the ith book has probability p_i to be chosen at each request. To make it quicker to find the book the next time, the librarian moves the book to the left end of the shelf. Define the state of a Markov chain at any time to be the list of books we see as we examine the shelf from left to right. Since all the books are distinct, the state space E is the set of all permutations of the set $\{1, 2, \ldots, n\}$. Show that

$$\pi(i_1, \cdots i_n) = p_{i_1} \cdot \frac{p_{i_2}}{1 - p_{i_1}} \cdots \frac{p_{i_n}}{1 - p_{i_1} - \cdots p_{i_{n-1}}}$$

is a stationary distribution.

Solution. The distribution π is stationary for the Markov chain if and only if, supposing that there exists $n > 0$ such that $X_n \sim \pi$, X_{n+1} has also the same distribution π.
Suppose that X_n is distributed according to π. Notice that if we are in state $(i_1, \cdots, i_n)$ at time $n + 1$, then the only possibilities for the chain at time n are

$$S = \{(i_1, \cdots, i_n), \ \ (i_2, i_1, i_3 \cdots, i_n), \ \ (i_2, i_3, i_1, i_4, \cdots, i_n), \ \cdots \ (i_2, i_3, i_4, \cdots, i_n, i_1)\}. \tag{1}$$

Hence, to reach $(i_1, \cdots, i_n)$ at time $n+1$, we choose the book i_1 at time n (starting from one of the states in S given by (1)). We get

$$\begin{aligned} \mathbb{P}(X_{n+1} = (i_1, \cdots, i_n)) &= \sum_{v \in S} \mathbb{P}(X_n = v)\mathbb{P}(X_{n+1} = (i_1, \cdots, i_n) \mid X_n = v) \qquad (2) \\ &= p_{i_1} \left(\pi(i_1, \cdots, i_n) + \pi(i_2, i_1, i_3, \cdots, i_n) + \cdots + \pi(i_2, i_3, i_4, \cdots, i_n, i_1)\right), \end{aligned}$$

where we supposed that X_n is distributed according to π. It remains to show that the right term of (2) is equal to $\pi(i_1, \cdots, i_n)$. To simplify the computations, we suppose that $n = 3$. We then obtain, noticing that $1 - p_{i_1} - p_{i_2} = p_{i_3}$,

$$\begin{aligned} p_{i_1} \frac{p_{i_2}}{1 - p_{i_1}} &\stackrel{?}{=} p_{i_1} \cdot (p_{i_1} \frac{p_{i_2}}{1 - p_{i_1}} + p_{i_2} \frac{p_{i_1}}{1 - p_{i_2}} + p_{i_2} \frac{p_{i_3}}{1 - p_{i_2}}) && \iff \\ p_{i_1} \frac{p_{i_2}}{1 - p_{i_1}} \cdot (1 - p_{i_1}) &\stackrel{?}{=} p_{i_1} \cdot (p_{i_2} \frac{p_{i_1}}{1 - p_{i_2}} + p_{i_2} \frac{p_{i_3}}{1 - p_{i_2}}) && \iff \\ p_{i_1} p_{i_2} (1 - \frac{p_{i_1}}{1 - p_{i_2}}) &\stackrel{?}{=} \frac{p_{i_1} p_{i_2} p_{i_3}}{1 - p_{i_2}} && \iff \\ \frac{p_{i_1} p_{i_2} p_{i_3}}{1 - p_{i_2}} &= \frac{p_{i_1} p_{i_2} p_{i_3}}{1 - p_{i_2}}. \end{aligned}$$

Since the last equality is always verified, we deduce that all the others are also verified, and thus $\mathbb{P}(X_{n+1} = (i_1, i_2, i_3)) = \pi(i_1, i_2, i_3)$. Since the state (i_1, i_2, i_3) is random, we deduce that X_{n+1} is distributed according to π.
The argument can be easily generalized to any n, and hence π is the stationary distribution of the system.

Exercise 2. Let P be a transition matrix on a finite state space E.

(a) Prove the following linear algebra result: Given a matrix Q, Q and Q^t have the same eigenvalues.
Use this to prove that P has a stationary distribution π(i.e. a probability measure $\pi P = \pi$).

(b) Find an example, when E is an infinite state space, for which P doesn't have any stationary distribution.

Solution. (a) A real number λ is an eigenvalue for a matrix Q if and only if $\det(Q-\lambda I)=0$ where I is the identity matrix of the same size as Q. We can write:

$$\det(Q^t-\lambda I)=\det(Q^t-\lambda I^t)=\det\left((Q-\lambda I)^t\right)=\det(Q-\lambda I).$$

This shows that Q and Q^t have the same eigenvalues.
Since P is a transition matrix, we have

$$P\mathbf{1}=\mathbf{1},$$

where $\mathbf{1}$ is a column vector of 1's of size $|E|$. Therefore, $\lambda=1$ is an eigenvalue for P. By the above result, we know that there exists a vector $\mathbf{v}$ such that

$$P^t\mathbf{v}=\mathbf{v}\Longleftrightarrow \mathbf{v^t}P=\mathbf{v^t}.$$

Remark:
Using the Perron-Frobenius theorem (not seen in this course), we can choose the eigenvector $\mathbf{v}=(v_1,v_2,\cdots)$ such that $v_i\geq 0$ for all i.
The vector $\pi:=\mathbf{v^t}/(\|\mathbf{v^t}\|)$ is a stationary distribution for P.

(b) If E is of infinite dimension, we can find examples for which $\|\mathbf{v^t}\|=\infty$ which implies $\pi\equiv\mathbf{0}$ and so is not a distribution anymore. We can, for example, consider the Markov chain in exercise 2a) in Serie 3. In this case we have

$$\pi_i=(\pi P)_i=\sum_j \pi_j P_{ji}=q\pi_i+p\pi_{i+1}.$$

This implies that $\pi_i=\pi_{i+1}$ for all $i\geq 0$ and so $\sum_i \pi_i=\infty$. In other words, π is not a stationary distribution.

Exercise 3. (**Gambler's ruin**) Assume that a gambler is making bets for 1 dollar on fair coin flips, and that she will abandon the game when her fortune falls to 0 or reaches n dollar. Let X_t be the Markov chain on $\{0,\dots,n\}$ describing the gambler's fortune at time t, that is, $\mathbb{P}(X_{t+1}=k+1\mid X_t=k)=\mathbb{P}(X_{t+1}=k-1\mid X_t=k)=1/2$, $k=1,\dots,n-1$, and $\mathbb{P}(X_{t+1}=0\mid X_t=0)=\mathbb{P}(X_{t+1}=n\mid X_t=n)=1$. Let T be the time required to be absorbed at one of 0 or n. Assume that $X_0=k$, where $0\leq k\leq n$.

(i). Find the probability $\mathbb{P}_k(X_T=n)$ for the gambler to reach n dollars with initial capital k.

(ii). Compute $\mathbb{E}_k[T]$, the expected time to reach n or 0 starting from k.

Solution. (i). Let p_k be the probability that the gambler reaches a fortune of n before ruin, given that she starts with k dollars. We solve simultaneously for $p_0,p_1,\cdots,p_n$. Clearly $p_0=0$ and $p_n=1$. By the total probability formula, we have

$$p_k=\frac{1}{2}p_{k-1}+\frac{1}{2}p_{k+1},\quad 1\leq k\leq n-1. \tag{3}$$

It's easy to solve this system to obtain $p_k = \frac{k}{n}$ for $0 \le k \le n$ (by defining $\Delta_k := p_k - p_{k-1}$ similarly to the next part). Another method to solve this recursive equation is by using its characteristic equation:

$$x = \frac{1}{2} + \frac{1}{2}x^2 \iff x^2 - 2x + 1 = 0. \tag{4}$$

If we have, in general, two distinct roots $x_1,\ x_2$ for this equation, then the solution of the recursive equation would be of the form

$$p_k = Ax_1^k + Bx_2^k,$$

where A and B are constants.
But in our case, the equation (4) has one root $x_0 = 1$ with multiplicity 2. In this case, the solution for (3) is given by

$$p_k = x_0^k(Ak + B) = Ak + B,$$

where A and B are constants. Since p_0=0 and $p_n = 1$, we get:

$$p_k = \frac{k}{n},\ 0 \le k \le n.$$

(ii). We write f_k for the expected time $\mathbb{E}_k[T]$ to be absorbed, starting at position k. Clearly, $f_0 = f_n = 0$. For $1 \le k \le n-1$, we have, conditioning on the first step

$$f_k = \frac{1}{2}(1 + f_{k+1}) + \frac{1}{2}(1 + f_{k-1}).$$

To solve this system, we let $\Delta_k = f_k - f_{k-1}$. It is easy to verify that $\Delta_k = \Delta_{k+1} + 2$. Using that $\sum_{k=1}^n \Delta_k = 0$, we obtain

$$\Delta_k = n - 1 - 2(k-1), \implies f_k = k(n-k),\ 0 \le k \le n.$$

Exercise 4. A company issues n different types of coupons. A collector desires a complete set. We suppose each coupon he acquires is equally likely one of the n types. Let X_l denote the number of different types represented among the collector's first l coupons. Clearly X_0 = 0. Let T_k be the total number of coupons accumulated when the collection first contains k distinct coupons.

(i). Find the distribution of $T_k - T_{k-1}$ for $k \le n$, that is, the time it takes to obtain the kth new coupon.

(ii). Find the expected value of the (random) time T representing the number of coupons needed to collect in order to have all coupon types.

Solution. (i). By the strong Markov property, we can write

$$\begin{aligned}\mathbb{P}(T_k - T_{k-1} = l) &= \mathbb{P}(X_{T_{k-1}+l} = k, X_{T_{k-1}} = k-1, \cdots, X_{T_{k-1}+l-1} = k-1) \\ &= \mathbb{P}(X_{T_{k-1}+l} = k \mid X_{T_{k-1}+l-1} = k-1) \\ &\times \prod_{i=1}^{l-1} \mathbb{P}(X_{T_{k-1}+i} = k-1 \mid X_{T_{k-1}+i-1} = k-1) = \left(\frac{k-1}{n}\right)^{l-1} \frac{n-k+1}{n}.\end{aligned}$$

This shows that $T_k - T_{k-1}$ has a geometric distribution with probability of success equal to $\frac{n-k+1}{n}$.

(ii). We can write T as a function of T_k's as follows

$$T = T_n = T_1 + (T_2 - T_1) + \cdots + (T_n - T_{n-1}).$$

We use this to obtain

$$\mathbb{E}[T] = \sum_{k=1}^{n} \mathbb{E}[T_k - T_{k-1}] = \sum_{k=1}^{n} \frac{n}{n-k+1} = n \sum_{k=1}^{n} \frac{1}{k}.$$

Exercise 5. (a) A transition matrix P defined on a state space E and a distribution λ have the *detailed balance property* if

$$\lambda_j P_{ji} = \lambda_i P_{ij}, \ \forall i, j \in E.$$

Show that in this case, λ is a stationary distribution for P.

(b) Consider two urns each of which contains m balls; b of these $2m$ balls are black, and the remaining $2m - b$ are white. We say that the system is at state i if the first urn contains i black balls and $m - i$ white balls while the second contains $b - i$ black balls and $m - b + i$ white balls. Each trial consists of choosing a ball at random from each urn and exchanging the two. Let X_n be the state of the system after n exchanges have been made. X_n is a Markov chain.

(1) Compute its transition probability.

(2) Verify (using (a)) that the stationary distribution is given by

$$\pi(i) = \frac{\binom{b}{i}\binom{2m-b}{m-i}}{\binom{2m}{m}}.$$

(3) Can you give a simple intuitive explanation why the formula in (2) gives the right answer?

Solution. (a) We need to show that $\lambda P = \lambda$. Using the detailed balance equations, we get easily for all $i \in I$:

$$(\lambda P)_i = \sum_{j \in I} \lambda_j p_{ji} = \sum_{j \in I} \lambda_i p_{ij} = \lambda_i.$$

(b) (1) Let $p(i, i+1)$ be the probability to go from i black balls to $i+1$ in the first urn after one step (for $i \leq b \wedge m$). This events happens if and only if we choose a white ball in the first urn and a black ball in the second one, so we have

$$p(i, i+1) = \frac{m-i}{m} \cdot \frac{b-i}{m}.$$

Similarly, we have $p(i, i-1) = \frac{i}{m} \cdot \frac{m-b+i}{m}$. After one step, the number of black balls can remain unchange i, or go to either $i+1$ or $i-1$. Therefore

$$p(i,i) = 1 - p(i, i+1) - p(i, i-1) = \frac{i}{m} \cdot \frac{b-i}{m} + \frac{m-i}{m} \cdot \frac{m-b+i}{m}.$$

(2) A sufficient condition for π to be a stationary distribution is to verify the detailed balance property:

$$\pi(i)p(i,i+1) = \pi(i+1)p(i+1,i), i \in [0, b \wedge m]. \tag{5}$$

If $|i-j| > 1$, the equalities $\pi(i)p(i,j) = \pi(j)p(j,i)$ are clearly satisfied since $p(i,j) = p(j,i) = 0$. By developping the left term in (5), we get

$$\begin{aligned}
\binom{2m}{m}\pi(i)m^2 p(i,i+1) &= \binom{b}{i}\binom{2m-b}{m-i}(m-i)(b-i) \\
&= \frac{b!}{i!(b-i-1)!} \cdot \frac{(2m-b)!}{(m-i-1)!(m-b+i)!} \\
&= \binom{b}{i+1}(i+1)\binom{2m-b}{m-i+1}(m-b+i+1) \\
&= m^2 \cdot \binom{b}{i+1}\binom{2m-b}{m-i+1}\frac{i+1}{m}\frac{m-b+i+1}{m} \\
&= \binom{2m}{m}\pi(i+1)m^2 p(i+1,i).
\end{aligned}$$

This shows that π verifies the detailed balance equations and so is the unique stationary distribution of the system.

(3) We know that if X_0 is distributed according to the stationary distribution π, then X_n is also distributed according to π for all $n \geq 1$. Suppose that we number the balls from 1 to $2m$ and we arrange them randomly (by a permutation $\sigma \in S_{2m}$) and we put the first m balls in the first urn (this corresponds to the definition of π in (b)). By exchanging two balls randomly chosen from the first and second urn, the new setting of the balls is "as random as" before. In other words, if the $2m$ balls are arranged in a uniform way at time $t = 0$, they will intuitively still be uniformly ordered after one step of the process.