Introduction to Stochastic Processes

Assignment 1
Ari Feiglin

1.1 Exercise

Compute the transition matrices of the following two diagrams (shown in the homework's pdf).

The left one has a transition matrix

$$\begin{pmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 1 \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

The right one has a transition matrix

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \\ 0 & 1 & 0 \end{pmatrix}$$

1.2 Exercise

Suppose $\{X_n\}_{n\geq 0}$ is a homogeneus Markov chain whose transition matrix is P.

(1) Show that $\{X_{3n}\}_{n\geq 0}$ is also a Markov chain. What is its transition matrix?

(2) Suppose $\{Y_n\}_{n\geq 0}$ is another homogeneus Markov chain whose transition matrix is Q. Is PQ the transition matrix of some Markov chain?

(1) We will show that if $\{a_n\}_{n\geq 0}$ is strictly increasing then X_{a_n} is a Markov chain. Let $n > 0$ and let $C = \{0, \dots, a_{n-1}\} \setminus \{a_{n-1}, \dots, a_0\}$ then

$$\begin{aligned} &\mathbb{P}(X_{a_n} = s_{a_n} \mid X_{a_{n-1}} = s_{a_{n-1}}, \dots, X_{a_0} = s_{a_0}) \\ &\qquad = \sum_{(s_i)_{i\in C}} \mathbb{P}(X_{a_n} = s_{a_n} \mid X_{a_{n-1}} = s_{a_{n-1}}, \dots, X_0 = s_0) \cdot \mathbb{P}((\forall i \in C) X_i = s_i \mid (\forall i < n) X_{a_i} = a_i) \\ &\qquad = \mathbb{P}(X_{a_n} = s_{a_n} \mid X_{a_{n-1}} = s_{a_{n-1}}) \cdot \sum_{(s_i)_{i\in C}} \mathbb{P}((\forall i \in C) X_i = s_i \mid (\forall i < n) X_{a_i} = a_i) \\ &\qquad = \mathbb{P}(X_{a_n} = s_{a_n} \mid X_{a_{n-1}} = s_{a_{n-1}}) \end{aligned}$$

The second transition is due to the result $\mathbb{P}(X_n = s_n \mid X_m = s_m, \dots, X_0 = s_0) = \mathbb{P}(X_n = s_n \mid X_m = s_m)$ for all $m < n$, which was discussed in lecture. This shows that $\{X_{a_n}\}_{n\geq 0}$ is indeed a Markov chain. Since $3n$ is strictly increasing, this means $\{X_{3n}\}_{n\geq 0}$ is indeed a Markov chain.

(2) Suppose P is a matrix with nonnegative entries such that for every i, $\sum_j P_{ij} = 1$, or equivalently $P\mathbf{1} = \mathbf{1}$ where $\mathbf{1} = (1, \dots, 1)^\top$ (this is as $(P\mathbf{1})_i = \sum_j P_{ij}\mathbf{1}_j = \sum_j P_{ij}$). Then P is the transition matrix of some (homogeneus) Markov chain. This is as we could define functions (random variables) X_n and a probability function such that $\mathbb{P}(X_n = j \mid X_{n-1} = i) = P_{ij}$.

So we must prove that $PQ\mathbf{1} = \mathbf{1}$, which is true since $PQ\mathbf{1} = P(Q\mathbf{1}) = P\mathbf{1} = \mathbf{1}$. And we must also prove that PQ's entries are all nonnegative, but this is trivial as they are all sums of products of entries in P and Q which are nonnegative.

1.3 Exercise

A day in London is either rainy or sunny. The probability that any day is rainy provided it was sunny the previous two days is 0.3, otherwise the likelihood of rain is 0.6. We denote R the state of rain and S of sun, and we define $W_n \in \{R, S\}$ to be the weather on the nth day. Let us define $X_n = (W_{n-1}, W_n)$.

(1) Is $\{W_n\}$ a Markov chain?

(2) Prove that $\{X_n\}$ is a Markov chain and compute its transition matrix.

(3) What is the probability that on the fourth day it is rainy provided it was sunny on the first and second days?

(1) It is not, for if it were then we'd have

$$\mathbb{P}(W_n = R \mid W_{n-1} = S) = \mathbb{P}(W_n = R \mid W_{n-1} = S, W_{n-2} = S) = \mathbb{P}(W_n = R \mid W_{n-1} = S, W_{n-2} = R),$$

but $\mathbb{P}(W_n = R \mid W_{n-1} = S, W_{n-2} = S) = 0.3$ and $\mathbb{P}(W_n = R \mid W_{n-1} = S, W_{n-2} = R) = 0.6$.

(2) Let us say that a sequence of random variables $\{X_n\}_{n\geq 0}$ is a *k-dependent Markov chain* if for every $n \geq 0$,

$$\mathbb{P}(X_{n+k} = s_{n+k} \mid X_{n+k-1} = s_{n+k-1}, \dots, X_0 = s_0) = \mathbb{P}(X_{n+k} = s_{n+k} \mid X_{n+k-1} = s_{n+k-1}, \dots, X_n = s_n)$$

(ie. X_n is dependent only on the previous k measurements.) Then let us define $Y_n = (X_{n-k+1}, \dots, X_n)$ for $n \geq k-1$. We claim that $\{Y_n\}_{n\geq k-1}$ is a Markov chain.

Notice that if $Y_n = (s_{n-k+1}, \dots, s_n)$ then $Y_{n+1} = (s_{n-k+2}, \dots, s_n, s_{n+1})$, so we need only deal with events where the first $k-1$ indexes of Y_{n+1} agree with the last $k-1$ of Y_n. And so let us denote $\vec{s_i} = (s_{i-k+1}, \dots, s_i)$ and

$$\begin{aligned}\mathbb{P}(Y_n = \vec{s}_n \mid Y_{n-1} = \vec{s}_{n-1}, \dots, Y_{k-1} = \vec{s}_{k-1}) &= \mathbb{P}(X_{n-k+1} = s_{n-k+1}, \dots, X_n = s_n \mid X_{n-1} = s_{n-1}, \dots, X_0 = s_0)\\ &= \mathbb{P}(X_n = s_n \mid X_{n-1} = s_{n-1}, \dots, X_0 = s_0)\\ &= \mathbb{P}(X_n = s_n \mid X_{n-1} = s_{n-1}, \dots, X_{n-k} = s_{n-k})\end{aligned}$$

And

$$\begin{aligned}\mathbb{P}(Y_n = \vec{s}_n \mid Y_{n-1} = \vec{s}_{n-1}) &= \mathbb{P}(X_{n-k+1} = s_{n-k+1}, \dots, X_n = s_n \mid X_{n-k} = s_{n-k}, \dots, X_{n-1} = s_{n-1})\\ &= \mathbb{P}(X_n = s_n \mid X_{n-1} = s_{n-1}, \dots, X_{n-k} = s_{n-k})\end{aligned}$$

Thus we have that $\mathbb{P}(Y_n = \vec{s}_n \mid Y_{n-1} = \vec{s}_{n-1}, \dots, Y_{k-1} = \vec{s}_{k-1}) = \mathbb{P}(Y_n = \vec{s}_n \mid Y_{n-1} = \vec{s}_{n-1})$, meaning $\{Y_n\}_{n\geq k-1}$ is indeed a Markov chain. (We take the states of Y_n to be S^k, as $\mathbb{P}\big(Y_n \in S^k\big) = \mathbb{P}(X_n \in S, \dots, X_{n-k+1} \in S) = 1$.)

Now, W_n is a 2-dependent Markov chain as W_n is dependent only on W_{n-1} and W_{n-2}, and thus $X_n = (W_{n-1}, W_n)$ is a Markov chain as we just showed. And as already computed,

$$\mathbb{P}(X_n = (s_2, s_1) \mid X_{n-1} = (s_3, s_2)) = \mathbb{P}(W_n = s_1 \mid W_{n-1} = s_2, W_{n-2} = s_3)$$

So if we order the states of X_n by $\{(R,R), (S,R), (R,S), (S,S)\}$ we get

$$P = \begin{pmatrix} 0.6 & 0 & 0.4 & 0 \\ 0.6 & 0 & 0.4 & 0 \\ 0 & 0.6 & 0 & 0.4 \\ 0 & 0.3 & 0 & 0.7 \end{pmatrix}$$

(3) Here we'd like to compute $\mathbb{P}(W_4 = R \mid X_2 = (S,S))$, and this is just equal to $\mathbb{P}(X_4 = (R,R) \mid X_2 = (S,S)) + \mathbb{P}(X_4 = (S,R) \mid X_2 = (S,S))$. In order to compute X_4's distribution given that $X_2 = (S,S)$ what we do is compute $(0,0,0,1)P^2$ (as $X_2 \sim (0,0,0,1)$ and P^2 gives the two-step transition matrix of $\{X_n\}$). Computing P^2 gives

$$\begin{pmatrix} 0.36 & 0.24 & 0.24 & 0.16 \\ 0.36 & 0.24 & 0.24 & 0.16 \\ 0.36 & 0.12 & 0.24 & 0.28 \\ 0.18 & 0.21 & 0.12 & 0.49 \end{pmatrix}$$

Then $(0,0,0,1)P^2$ is simply the fourth line of P^2 which is $(0.18, 0.21, 0.12, 0.49)$ and so

$$\mathbb{P}(X_4 = (R,R) \mid X_2 = (S,S)) + \mathbb{P}(X_4 = (S,R) \mid X_2 = (S,S)) = 0.18 + 0.21 = 0.39$$