MARKOV CHAINS

Remarks on Example Sheet 2

2.1 Let i ($\neq s$) be a state of the chain, and define $n_i = \min\{n : p_{is}(n) > 0\}$. If $X_0 = i$ and $X_{n_i} = s$ then, with probability one, X makes no visit to i during the intervening period $[1, n_i - 1]$; this follows from the minimality of n_i. Now s is absorbing, and hence

$$\mathbb{P}(\text{no return to } i \mid X_0 = i) \geq \mathbb{P}(X_{n_i} = s \mid X_0 = i) > 0.$$

2.2 If $r = 1$, then state i is absorbing for $i \geq 1$; also, 0 is transient unless $a_0 = 1$.

Assume $r < 1$ and let $J = \sup\{j : a_j > 0\}$. The states $0, 1, \ldots, J$ form an irreducible persistent class; they are aperiodic if $r > 0$. All other states are transient. For $0 \leq i \leq J$, the recurrence time T_i of i satisfies $\mathbb{P}(T_i = 1) = r$. If $T_i > 1$ then T_i may be expressed as the sum of

$$\begin{aligned} T_i^{(1)} &:= \text{ time to reach 0, given that the first step is leftwards,} \\ T_i^{(2)} &:= \text{ time spent in excursions from 0 not reaching } i, \\ T_i^{(3)} &:= \text{ time taken to reach } i \text{ in final excursion.} \end{aligned}$$

It is easy to see that $\mathbb{E}\big(T_i^{(1)}\big) = 1 + (i-1)/(1-r)$ if $i \geq 1$, since the waiting time at each intermediate point has mean $(1-r)^{-1}$. The number N of such 'small' excursions has mass function $\mathbb{P}(N = n) = \alpha_i(1-\alpha_i)^n$, $n \geq 0$, where $\alpha_i = \sum_{j=i}^{\infty} a_j$; hence $\mathbb{E}(N) = (1-\alpha_i)/\alpha_i$. Each such small excursion has mean duration

$$\sum_{j=0}^{i-1} \left(\frac{j}{1-r} + 1 \right) \frac{a_j}{1-\alpha_i} = 1 + \sum_{j=0}^{i-1} \frac{j a_j}{(1-\alpha_i)(1-r)}$$

and therefore

$$\mathbb{E}\big(T_i^{(2)}\big) = \frac{1}{\alpha_i} \left\{ (1-\alpha_i) + \sum_{j=0}^{i-1} \frac{j a_j}{1-r} \right\}.$$

By a similar argument,

$$\mathbb{E}\big(T_i^{(3)}\big) = \frac{1}{\alpha_i} \sum_{j=i}^{\infty} \left(1 + \frac{j-i}{1-r} \right) a_j.$$

Combining this information, we obtain that

$$\mathbb{E}(T_i) = r + (1-r)\mathbb{E}\big(T_i^{(1)} + T_i^{(2)} + T_i^{(3)}\big) = \frac{1}{\alpha_i} \left(1 - r + \sum_{j=0}^{\infty} j a_j \right), \quad i \geq 1,$$

and a similar argument yields $\mathbb{E}(T_0) = 1 + \sum_j j a_j/(1-r)$. The apparent simplicity of these formulae suggests the possibility of an easier derivation; see Exercise (6.4.25). Clearly $\mathbb{E}(T_i) < \infty$ for $i \leq J$ whenever $\sum_j j a_j < \infty$, a condition which certainly holds if $J < \infty$.

Alternatively, compute the invariant distribution π, and set $\mathbb{E}(T_i) = 1/\pi_i$.

2.3 All states are absorbing if $p = 0$. Assume henceforth that $p \neq 0$. Diagonalize $\mathbf{P}$ to obtain $\mathbf{P} = \mathbf{B}\boldsymbol{\Lambda}\mathbf{B}^{-1}$ where

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \quad \mathbf{B}^{-1} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \end{pmatrix},$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-2p & 0 \\ 0 & 0 & 1-4p \end{pmatrix}.$$

Therefore

$$\mathbf{P}^n = \mathbf{B}\boldsymbol{\Lambda}^n\mathbf{B}^{-1} = \mathbf{B} \begin{pmatrix} 1 & 0 & 0 \\ 0 & (1-2p)^n & 0 \\ 0 & 0 & (1-4p)^n \end{pmatrix} \mathbf{B}^{-1}$$

whence $p_{ij}(n)$ is easily found.

In particular, $p_{11}(n) = \frac{1}{4} + \frac{1}{2}(1-2p)^n + \frac{1}{4}(1-4p)^n$.

2.4 (a) By symmetry, all states have the same mean-recurrence time. Using the renewal-process argument of the last solution, the common value equals 8, being the number of vertices of the cube. Hence $\mu_v = 8$.

Alternatively, let s be a neighbour of v, and let t be a neighbour of s other than v. In the obvious notation, by symmetry,

$$\begin{aligned} \mu_v &= 1 + \tfrac{3}{4}\mu_{sv}, & \mu_{sv} &= 1 + \tfrac{1}{4}\mu_{sv} + \tfrac{1}{2}\mu_{tv}, \\ \mu_{tv} &= 1 + \tfrac{1}{2}\mu_{sv} + \tfrac{1}{4}\mu_{tv} + \tfrac{1}{4}\mu_{wv}, & \mu_{wv} &= 1 + \tfrac{1}{4}\mu_{wv} + \tfrac{3}{4}\mu_{tv}, \end{aligned}$$

a system of equations which may be solved to obtain $\mu_v = 8$.
(b) Using the above equations, $\mu_{wv} = 40/3$, whence $\mu_{vw} = 40/3$ by symmetry.
(c) The required number X satisfies $\mathbb{P}(X = n) = \theta^{n-1}(1-\theta)^2$ for $n \geq 1$, where θ is the probability that the first return of the walk to its starting point precedes its first visit to the diametrically opposed vertex. Therefore

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} n\theta^{n-1}(1-\theta)^2 = 1.$$

2.5 This is just to check they have registered that recurrence and transience are class properties, and a finite class is recurrent if and only if it is closed.

2.6 The chain is transient if and only if $\sum_{j=1}^{\infty} j^{-\alpha} < \infty$ by the calculation of 1.13. Transience means only finitely many visits to each state, recurrence means infinitely many visits to every state, and one or the other holds. Hence

$$\mathbb{P}(X_n \to \infty \text{ as } n \to \infty) = \begin{cases} 1 & \text{if } \alpha > 1 \\ 0 & \text{if } \alpha \leq 1. \end{cases}$$

Students may not notice that the $\alpha \leq 1$ case gives a recurrent walk even though there is a greater tendency to move right than left.

2.7 Draw the tree! Set X_n to be the distance from R at time n. Then $(X_n)_{n\geq 0}$ is a biased random walk on $\{0, 1, \dots\}$ with $p_{i,i+1} = \frac{2}{3}$, $p_{i,i-1} = \frac{1}{3}$, which is known to be transient, so the random walk on T is certainly transient.

2.8 Project onto $\mathbb{Z}^3$ and remove times at which the walk stays still (using the strong Markov property). So obtain the simple symmetric random walk in $\mathbb{Z}^3$ which is known to be transient.

2.9 Invariant distributions are $\{(\frac{1}{2}p, 0, 1-p, 0, \frac{1}{2}p)\colon\ 0 \leq p \leq 1\}$. The point is to show how non-uniqueness occurs when the matrix is not irreducible.

2.10 The state-space is $\{0, 1, \dots, N\}$ and the transition probabilities are given by

$$p_{i,i+1} + p_{i,i-1} = 1, \quad p_{i,i-1} = \frac{i}{N}, \quad i = 1, \dots, N.$$

The underlying physical model makes it clear that the equilibrium distribution is $\mathrm{Bin}(N, \frac{1}{2})$, which may also be discovered more laboriously from the transition probabilities.

2.11 Students must be drilled to check for irreducibility and aperiodicity before using $p_{ij}(n) \to \pi_j$ for an invariant distribution π.

2.12 The invariant distribution π satisfies $\pi_1 = p\pi_3$ and $\pi_3 = \frac{1}{3}\pi_2$, so $\pi = (p, 3, 1)/(p+4)$.

2.13 Draw the flow diagram, symmetrically. The invariant measure π satisfies

$$\pi_{123} = \alpha_1(\pi_{123} + \pi_{231} + \pi_{213}), \quad \pi_{132} = \alpha_1(\pi_{132} + \pi_{312} + \pi_{321})$$

so $\pi_{123} + \pi_{132} = \alpha_1$, similarly $\pi_{231} + \pi_{213} = \alpha_2$ and $\pi_{123} = \alpha_1(\pi_{123} + \alpha_2)$.

2.14 This is partly covered in lectures. Try to solve the detailed balance equations $\pi_i p_{ij} = \pi_j p_{ji}$. In case (b) no solution exists if $p \neq \frac{1}{2}$. In case (c) there is no solution with $\sum_{i=0}^{\infty} \pi_i = 1$ if $p \geq \frac{1}{2}$. Otherwise P is reversible. It may be worth stressing that all birth and death processes are reversible and that this is a good way to find invariant distributions for such processes.

2.15 Look for a solution to the equations $\pi_i p_{ij} = \pi_j p_{ji}$. The only non-trivial cases of interest are those with $j = i+1$, and therefore $\lambda_i \pi_i = \mu_{i+1}\pi_{i+1}$ for $0 \leq i < b$, with solution

$$\pi_i = \pi_0 \frac{\lambda_0 \lambda_1 \cdots \lambda_{i-1}}{\mu_1 \mu_2 \cdots \mu_i}, \qquad 0 \leq i \leq b,$$

an empty product being interpreted as 1. The constant π_0 is chosen in order that the π's sum to 1, and the chain is therefore time-reversible.

2.16 Let π be the stationary distribution of X, and suppose X is time-reversible. We have that $\pi_i p_{ij} = p_{ji}\pi_j$ for all i, j, and furthermore $\pi_i > 0$ for all i. Hence

$$\pi_i p_{ij} p_{jk} p_{ki} = p_{ji} \pi_j p_{jk} p_{ki} = p_{ji} p_{kj} \pi_k p_{ki} = p_{ji} p_{kj} p_{ik} \pi_i$$

as required when $n = 3$. A similar calculation is valid when $n > 3$.

Suppose conversely that the given display holds for all finite sequences of states. Sum over all values of the subsequence $j_2, \ldots, j_{n-1}$ to deduce that $p_{ij}(n-1)p_{ji} = p_{ij}p_{ji}(n-1)$, where $i = j_1$, $j = j_n$. Take the limit as $n \to \infty$ to obtain $\pi_j p_{ji} = p_{ij}\pi_i$ as required for time-reversibility.