1.9 Time reversal

For Markov chains, the past and future are independent given the present. This property is symmetrical in time and suggests looking at Markov chains with time running backwards. On the other hand, convergence to equilibrium shows behaviour which is asymmetrical in time: a highly organised state such as a point mass decays to a disorganised one, the invariant distribution. This is an example of entropy increasing. It suggests that if we want complete time-symmetry we must begin in equilibrium. The next result shows that a Markov chain in equilibrium, run backwards, is again a Markov chain. The transition matrix may however be different.

Theorem 1.9.1. *Let P be irreducible and have an invariant distribution π. Suppose that $(X_n)_{0\le n\le N}$ is Markov(π, P) and set $Y_n = X_{N-n}$. Then $(Y_n)_{0\le n\le N}$ is Markov$(\pi, \widehat{P})$, where $\widehat{P} = (\widehat{p}_{ij})$ is given by*

$$\pi_j \widehat{p}_{ji} = \pi_i p_{ij} \quad \textit{for all } i, j$$

and $\widehat{P}$ is also irreducible with invariant distribution π.

Proof. First we check that $\widehat{P}$ is a stochastic matrix:

$$\sum_{i\in I} \widehat{p}_{ji} = \frac{1}{\pi_j} \sum_{i\in I} \pi_i p_{ij} = 1$$

since π is invariant for P. Next we check that π is invariant for $\widehat{P}$:

$$\sum_{j\in I} \pi_j \widehat{p}_{ji} = \sum_{j\in I} \pi_i p_{ij} = \pi_i$$

since P is a stochastic matrix.

We have

$$\begin{aligned} &P(Y_0 = i_0, Y_1 = i_1, \dots, Y_N = i_N) \\ &\quad = P(X_0 = i_N, X_1 = i_{N-1}, \dots, X_N = i_0) \\ &\quad = \pi_{i_N} p_{i_N i_{N-1}} \dots p_{i_1 i_0} = \pi_{i_0} \widehat{p}_{i_0 i_1} \dots \widehat{p}_{i_{N-1} i_N} \end{aligned}$$

so, by Theorem 1.1.1, $(Y_n)_{0\le n\le N}$ is Markov$(\pi, \widehat{P})$. Finally, since P is irreducible, for each pair of states i, j there is a chain of states $i_0 = i, i_1, \dots, i_{n-1}, i_n = j$ with $p_{i_0 i_1} \dots p_{i_{n-1} i_n} > 0$. Then

$$\widehat{p}_{i_n i_{n-1}} \dots \widehat{p}_{i_1 i_0} = \pi_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i_n} / \pi_{i_n} > 0$$

so $\widehat{P}$ is also irreducible. $\square$

The chain $(Y_n)_{0\le n\le N}$ is called the *time-reversal* of $(X_n)_{0\le n\le N}$.

A stochastic matrix P and a measure λ are said to be in *detailed balance* if

$$\lambda_i p_{ij} = \lambda_j p_{ji} \quad \text{for all } i, j.$$

Though obvious, the following result is worth remembering because, when a solution λ to the detailed balance equations exists, it is often easier to find by the detailed balance equations than by the equation $\lambda = \lambda P$.

Lemma 1.9.2. *If P and λ are in detailed balance, then λ is invariant for P.*

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

Let $(X_n)_{n\geq 0}$ be Markov(λ, P), with P irreducible. We say that $(X_n)_{n\geq 0}$ is *reversible* if, for all $N \geq 1$, $(X_{N-n})_{0\leq n\leq N}$ is also Markov(λ, P).

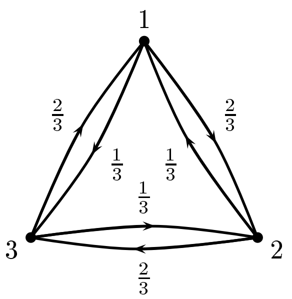
Theorem 1.9.3. *Let P be an irreducible stochastic matrix and let λ be a distribution. Suppose that $(X_n)_{n\geq 0}$ is Markov(λ, P). Then the following are equivalent:*
(a) *$(X_n)_{n\geq 0}$ is reversible;*
(b) *P and λ are in detailed balance.*

Proof. Both (a) and (b) imply that λ is invariant for P. Then both (a) and (b) are equivalent to the statement that $\widehat{P} = P$ in Theorem 1.9.1. $\square$

We begin a collection of examples with a chain which is not reversible.

Example 1.9.4

Consider the Markov chain with diagram:



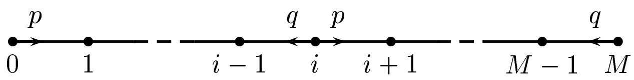
The transition matrix is

$$P = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \end{pmatrix}$$

and $\pi = (1/3, 1/3, 1/3)$ is invariant. Hence $\widehat{P} = P^T$, the transpose of P. But P is not symmetric, so $P \neq \widehat{P}$ and this chain is not reversible. A patient observer would see the chain move clockwise in the long run: under time-reversal the clock would run backwards!

Example 1.9.5

Consider the Markov chain with diagram:



where $0 < p = 1 - q < 1$. The non-zero detailed balance equations read

$$\lambda_i p_{i,i+1} = \lambda_{i+1} p_{i+1,i} \quad \text{for } i = 0, 1, \dots, M-1.$$

So a solution is given by

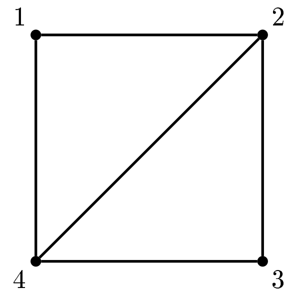
$$\lambda = \big((p/q)^i : i = 0, 1, \dots, M\big)$$

and this may be normalised to give a distribution in detailed balance with P. Hence this chain is reversible.

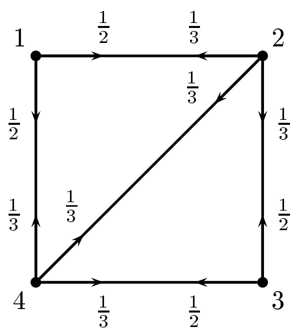
If p were much larger than q, one might argue that the chain would tend to move to the right and its time-reversal to the left. However, this ignores the fact that we reverse the chain *in equilibrium*, which in this case would be heavily concentrated near M. An observer would see the chain spending most of its time near M and making occasional brief forays to the left, which behaviour is symmetrical in time.

Example 1.9.6 (Random walk on a graph)

A *graph* G is a countable collection of states, usually called *vertices*, some of which are joined by *edges*, for example:



Thus a graph is a partially drawn Markov chain diagram. There is a natural way to complete the diagram which gives rise to the random walk on G. The *valency* v_i of vertex i is the number of edges at i. We have to assume that every vertex has finite valency. The random walk on G picks edges with equal probability:



Thus the transition probabilities are given by

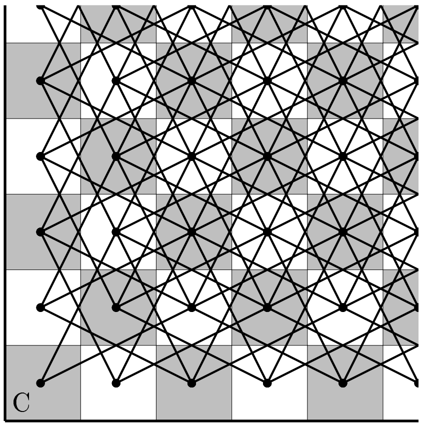
$$p_{ij} = \begin{cases} 1/v_i & \text{if } (i,j) \text{ is an edge} \\ 0 & \text{otherwise.} \end{cases}$$

We assume G is connected, so that P is irreducible. It is easy to see that P is in detailed balance with $v = (v_i : i \in G)$. So, if the total valency $\sigma = \sum_{i \in G} v_i$ is finite, then $\pi = v/\sigma$ is invariant and P is reversible.

Example 1.9.7 (Random chessboard knight)

A random knight makes each permissible move with equal probability. If it starts in a corner, how long on average will it take to return?

This is an example of a random walk on a graph: the vertices are the squares of the chessboard and the edges are the moves that the knight can take:



The diagram shows a part of the graph. We know by Theorem 1.7.7 and the preceding example that

$$E_c(T_c) = 1/\pi_c = \sum_i (v_i/v_c)$$

so all we have to do is identify valencies. The four corner squares have valency 2, and the eight squares adjacent to the corners have valency 3. There are 20 squares of valency 4, 16 of valency 6, and the 16 central squares have valency 8. Hence

$$E_c(T_c) = \frac{8 + 24 + 80 + 96 + 128}{2} = 168.$$

Alternatively, if you enjoy solving sets of 64 simultaneous linear equations, you might try finding π from $\pi P = \pi$, or calculating $E_c(T_c)$ using Theorem 1.3.5!

Exercises

1.9.1 In each of the following cases determine whether the stochastic matrix P, which you may assume is irreducible, is reversible:

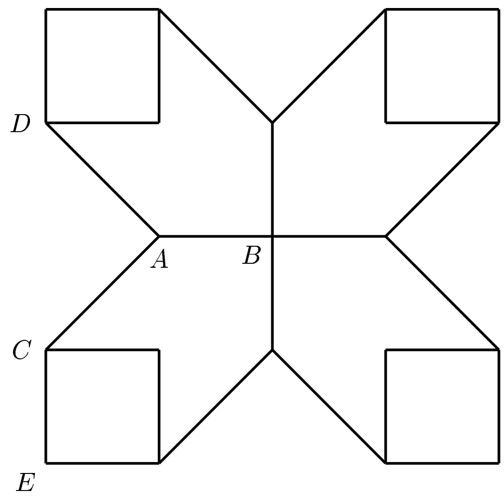
(a) $\begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$; (b) $\begin{pmatrix} 0 & p & 1-p \\ 1-p & 0 & p \\ p & 1-p & 0 \end{pmatrix}$;

(c) $I = \{0, 1, \dots, N\}$ and $p_{ij} = 0$ if $|j - i| \geq 2$;

(d) $I = \{0, 1, 2, \dots\}$ and $p_{01} = 1$, $p_{i,i+1} = p$, $p_{i,i-1} = 1 - p$ for $i \geq 1$;

(e) $p_{ij} = p_{ji}$ for all $i, j \in S$.

1.9.2 Two particles X and Y perform independent random walks on the graph shown in the diagram. So, for example, a particle at A jumps to B, C or D with equal probability $1/3$.



Find the probability that X and Y ever meet at a vertex in the following cases:

(a) X starts at A and Y starts at B;

(b) X starts at A and Y starts at E. For $I = B, D$ let M_I denote the expected time, when both X and Y start at I, until they are once again both at I. Show that $9M_D = 16M_B$.