the left-hand side being the total probability of all the ways of placing n balls randomly into three boxes. For the case where $n = 3m$, we have

$$\binom{n}{i\,j\,k} = \frac{n!}{i!j!k!} \leq \binom{n}{m\,m\,m}$$

for all i, j, k, so

$$p_{00}^{(2n)} \leq \binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\binom{n}{m\,m\,m}\left(\frac{1}{3}\right)^{n} \sim \frac{1}{2A^3}\left(\frac{6}{n}\right)^{3/2} \quad \text{as } n \to \infty$$

by Stirling's formula. Hence, $\sum_{m=0}^{\infty} p_{00}^{(6m)} < \infty$ by comparison with $\sum_{n=0}^{\infty} n^{-3/2}$. But $p_{00}^{(6m)} \geq (1/6)^2\, p_{00}^{(6m-2)}$ and $p_{00}^{(6m)} \geq (1/6)^4\, p_{00}^{(6m-4)}$ for all m so we must have

$$\sum_{n=0}^{\infty} p_{00}^{(n)} < \infty$$

and the walk is transient.

Exercises

1.6.1 The rooted binary tree is an infinite graph T with one distinguished vertex R from which comes a single edge; at every other vertex there are three edges and there are no closed loops. The random walk on T jumps from a vertex along each available edge with equal probability. Show that the random walk is transient.

1.6.2 Show that the simple symmetric random walk in $\mathbb{Z}^4$ is transient.

1.7 Invariant distributions

Many of the long-time properties of Markov chains are connected with the notion of an invariant distribution or measure. Remember that a measure λ is any row vector $(\lambda_i : i \in I)$ with non-negative entries. We say λ is *invariant* if

$$\lambda P = \lambda.$$

The terms *equilibrium* and *stationary* are also used to mean the same. The first result explains the term stationary.

Theorem 1.7.1. *Let $(X_n)_{n\geq 0}$ be Markov(λ, P) and suppose that λ is invariant for P. Then $(X_{m+n})_{n\geq 0}$ is also Markov(λ, P).*

Proof. By Theorem 1.1.3, $\mathbb{P}(X_m = i) = (\lambda P^m)_i = \lambda_i$ for all i and clearly, conditional on $X_{m+n} = i$, X_{m+n+1} is independent of $X_m, X_{m+1}, \ldots, X_{m+n}$ and has distribution $(p_{ij} : j \in I)$. $\square$

The next result explains the term equilibrium.

Theorem 1.7.2. *Let I be finite. Suppose for some $i \in I$ that*

$$p_{ij}^{(n)} \to \pi_j \quad \text{as} \quad n \to \infty \quad \text{for all } j \in I.$$

Then $\pi = (\pi_j : j \in I)$ is an invariant distribution.

Proof. We have

$$\sum_{j \in I} \pi_j = \sum_{j \in I} \lim_{n \to \infty} p_{ij}^{(n)} = \lim_{n \to \infty} \sum_{j \in I} p_{ij}^{(n)} = 1$$

and

$$\pi_j = \lim_{n \to \infty} p_{ij}^{(n)} = \lim_{n \to \infty} \sum_{k \in I} p_{ik}^{(n)} p_{kj} = \sum_{k \in I} \lim_{n \to \infty} p_{ik}^{(n)} p_{kj} = \sum_{k \in I} \pi_k p_{kj}$$

where we have used finiteness of I to justify interchange of summation and limit operations. Hence π is an invariant distribution. $\square$

Notice that for any of the random walks discussed in Section 1.6 we have $p_{ij}^{(n)} \to 0$ as $n \to \infty$ for all $i, j \in I$. The limit is certainly invariant, but it is not a distribution!

Theorem 1.7.2 is not a very useful result but it serves to indicate a relationship between invariant distributions and n-step transition probabilities. In Theorem 1.8.3 we shall prove a sort of converse, which is much more useful.

Example 1.7.3

Consider the two-state Markov chain with transition matrix

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

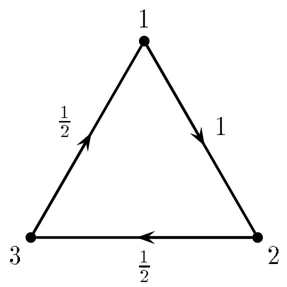
Ignore the trivial cases $\alpha = \beta = 0$ and $\alpha = \beta = 1$. Then, by Example 1.1.4

$$P^n \to \begin{pmatrix} \beta/(\alpha+\beta) & \alpha/(\alpha+\beta) \\ \beta/(\alpha+\beta) & \alpha/(\alpha+\beta) \end{pmatrix} \quad \text{as } n \to \infty,$$

so, by Theorem 1.7.2, the distribution $(\beta/(\alpha+\beta), \alpha/(\alpha+\beta))$ must be invariant. There are of course easier ways to discover this.

Example 1.7.4

Consider the Markov chain $(X_n)_{n \geq 0}$ with diagram



To find an invariant distribution we write down the components of the vector equation $\pi P = \pi$

$$\begin{aligned} \pi_1 &= \tfrac{1}{2}\pi_3 \\ \pi_2 &= \tfrac{1}{2}\pi_1 + \tfrac{1}{2}\pi_3 \\ \pi_3 &= \tfrac{1}{2}\pi_2 + \tfrac{1}{2}\pi_3. \end{aligned}$$

In terms of the chain, the right-hand sides give the probabilities for X_1, when X_0 has distribution π, and the equations require X_1 also to have distribution π. The equations are homogeneous so one of them is redundant, and another equation is required to fix π uniquely. That equation is

$$\pi_1 + \pi_2 + \pi_3 = 1$$

and we find that $\pi = \big(1/5, 2/5, 2/5\big)$.

According to Example 1.1.6

$$p_{11}^{(n)} \to 1/5 \quad \text{as } n \to \infty$$

so this confirms Theorem 1.7.2. Alternatively, knowing that $p_{11}^{(n)}$ had the form

$$p_{11}^{(n)} = a + \left(\frac{1}{2}\right)^n \left(b \cos \frac{n\pi}{2} + c \sin \frac{n\pi}{2}\right)$$

we could have used Theorem 1.7.2 and knowledge of π_1 to identify $a = 1/5$, instead of working out $p_{11}^{(2)}$ in Example 1.1.6.

In the next two results we shall show that every irreducible and recurrent stochastic matrix P has an essentially unique positive invariant measure. The proofs rely heavily on the probabilistic interpretation so it is worth noting at the outset that, for a finite state-space I, the existence of an invariant row vector is a simple piece of linear algebra: the row sums of P are all 1, so the column vector of ones is an eigenvector with eigenvalue 1, so P must have a row eigenvector with eigenvalue 1.

For a fixed state k, consider for each i the *expected time spent in i between visits to k*:

$$\gamma_i^k = \mathbb{E}_k \sum_{n=0}^{T_k - 1} 1_{\{X_n = i\}}.$$

Here the sum of indicator functions serves to count the number of times n at which $X_n = i$ before the first passage time T_k.

Theorem 1.7.5. *Let P be irreducible and recurrent. Then*

(i) $\gamma_k^k = 1$;

(ii) $\gamma^k = (\gamma_i^k : i \in I)$ *satisfies* $\gamma^k P = \gamma^k$;

(iii) $0 < \gamma_i^k < \infty$ *for all* $i \in I$.

Proof. (i) This is obvious. (ii) For $n = 1, 2, \dots$ the event $\{n \le T_k\}$ depends only on $X_0, X_1, \dots, X_{n-1}$, so, by the Markov property at $n-1$

$$\mathbb{P}_k(X_{n-1} = i, X_n = j \text{ and } n \le T_k) = \mathbb{P}_k(X_{n-1} = i \text{ and } n \le T_k) p_{ij}.$$

Since P is recurrent, under $\mathbb{P}_k$ we have $T_k < \infty$ and $X_0 = X_{T_k} = k$ with probability one. Therefore

$$\begin{aligned}
\gamma_j^k &= \mathbb{E}_k \sum_{n=1}^{T_k} 1_{\{X_n = j\}} = \mathbb{E}_k \sum_{n=1}^{\infty} 1_{\{X_n = j \text{ and } n \le T_k\}} \\
&= \sum_{n=1}^{\infty} \mathbb{P}_k(X_n = j \text{ and } n \le T_k) \\
&= \sum_{i \in I} \sum_{n=1}^{\infty} \mathbb{P}_k(X_{n-1} = i, X_n = j \text{ and } n \le T_k) \\
&= \sum_{i \in I} p_{ij} \sum_{n=1}^{\infty} \mathbb{P}_k(X_{n-1} = i \text{ and } n \le T_k) \\
&= \sum_{i \in I} p_{ij} \mathbb{E}_k \sum_{m=0}^{\infty} 1_{\{X_m = i \text{ and } m \le T_k - 1\}} \\
&= \sum_{i \in I} p_{ij} \mathbb{E}_k \sum_{m=0}^{T_k - 1} 1_{\{X_m = i\}} = \sum_{i \in I} \gamma_i^k p_{ij}.
\end{aligned}$$

(iii) Since P is irreducible, for each state i there exist $n, m \ge 0$ with $p_{ik}^{(n)}, p_{ki}^{(m)} > 0$. Then $\gamma_i^k \ge \gamma_k^k p_{ki}^{(m)} > 0$ and $\gamma_i^k p_{ik}^{(n)} \le \gamma_k^k = 1$ by (i) and (ii). □

Theorem 1.7.6. *Let P be irreducible and let λ be an invariant measure for P with $\lambda_k = 1$. Then $\lambda \ge \gamma^k$. If in addition P is recurrent, then $\lambda = \gamma^k$.*

Proof. For each $j \in I$ we have

$$\begin{aligned}
\lambda_j &= \sum_{i_0 \in I} \lambda_{i_0} p_{i_0 j} = \sum_{i_0 \ne k} \lambda_{i_0} p_{i_0 j} + p_{kj} \\
&= \sum_{i_0, i_1 \ne k} \lambda_{i_1} p_{i_1 i_0} p_{i_0 j} + \Big(p_{kj} + \sum_{i_0 \ne k} p_{k i_0} p_{i_0 j} \Big) \\
&\qquad \vdots \\
&= \sum_{i_0, \dots, i_n \ne k} \lambda_{i_n} p_{i_n i_{n-1}} \dots p_{i_0 j} \\
&\quad + \Big(p_{kj} + \sum_{i_0 \ne k} p_{k i_0} p_{i_0 j} + \dots + \sum_{i_0, \dots, i_{n-1} \ne k} p_{k i_{n-1}} \dots p_{i_1 i_0} p_{i_0 j} \Big) \\
&\ge \mathbb{P}_k(X_1 = j \text{ and } T_k \ge 1) + \mathbb{P}_k(X_2 = j \text{ and } T_k \ge 2) \\
&\quad + \dots + \mathbb{P}_k(X_n = j \text{ and } T_k \ge n) \\
&\to \gamma_j^k \quad \text{as } n \to \infty.
\end{aligned}$$

So $\lambda \geq \gamma^k$. If P is recurrent, then γ^k is invariant by Theorem 1.7.5, so $\mu = \lambda - \gamma^k$ is also invariant and $\mu \geq 0$. Since P is irreducible, given $i \in I$, we have $p_{ik}^{(n)} > 0$ for some n, and $0 = \mu_k = \sum_{j \in I} \mu_j p_{jk}^{(n)} \geq \mu_i p_{ik}^{(n)}$, so $\mu_i = 0$. □

Recall that a state i is recurrent if

$$\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 1$$

and we showed in Theorem 1.5.3 that this is equivalent to

$$\mathbb{P}_i(T_i < \infty) = 1.$$

If in addition the *expected return time*

$$m_i = \mathbb{E}_i(T_i)$$

is finite, then we say i is *positive recurrent*. A recurrent state which fails to have this stronger property is called *null recurrent*.

Theorem 1.7.7. *Let P be irreducible. Then the following are equivalent:*

(i) *every state is positive recurrent;*

(ii) *some state i is positive recurrent;*

(iii) *P has an invariant distribution, π say. Moreover, when* (iii) *holds we have $m_i = 1/\pi_i$ for all i.*

Proof. (i) $\Rightarrow$ (ii) This is obvious.
(ii) $\Rightarrow$ (iii) If i is positive recurrent, it is certainly recurrent, so P is recurrent. By Theorem 1.7.5, γ^i is then invariant. But

$$\sum_{j \in I} \gamma_j^i = m_i < \infty$$

so $\pi_j = \gamma_j^i / m_i$ defines an invariant distribution.
(iii) $\Rightarrow$ (i) Take any state k. Since P is irreducible and $\sum_{i \in I} \pi_i = 1$ we have $\pi_k = \sum_{i \in I} \pi_i p_{ik}^{(n)} > 0$ for some n. Set $\lambda_i = \pi_i / \pi_k$. Then λ is an invariant measure with $\lambda_k = 1$. So by Theorem 1.7.6, $\lambda \geq \gamma^k$. Hence

$$m_k = \sum_{i \in I} \gamma_i^k \leq \sum_{i \in I} \frac{\pi_i}{\pi_k} = \frac{1}{\pi_k} < \infty \tag{1.7}$$

and k is positive recurrent.

To complete the proof we return to the argument for (iii) $\Rightarrow$ (i) armed with the knowledge that P is recurrent, so $\lambda = \gamma^k$ and the inequality (1.7) is in fact an equality. □

Example 1.7.8 (Simple symmetric random walk on $\mathbb{Z}$)

The simple symmetric random walk on $\mathbb{Z}$ is clearly irreducible and, by Example 1.6.1, it is also recurrent. Consider the measure

$$\pi_i = 1 \quad \text{for all } i.$$

Then

$$\pi_i = \tfrac{1}{2}\pi_{i-1} + \tfrac{1}{2}\pi_{i+1}$$

so π is invariant. Now Theorem 1.7.6 forces any invariant measure to be a scalar multiple of π. Since $\sum_{i\in\mathbb{Z}} \pi_i = \infty$, there can be no invariant distribution and the walk is therefore null recurrent, by Theorem 1.7.7.

Example 1.7.9

The existence of an invariant measure does not guarantee recurrence: consider, for example, the simple symmetric random walk on $\mathbb{Z}^3$, which is transient by Example 1.6.3, but has invariant measure π given by $\pi_i = 1$ for all i.

Example 1.7.10

Consider the asymmetric random walk on $\mathbb{Z}$ with transition probabilities $p_{i,i-1} = q < p = p_{i,i+1}$. In components the invariant measure equation $\pi P = \pi$ reads

$$\pi_i = \pi_{i-1}p + \pi_{i+1}q.$$

This is a recurrence relation for π with general solution

$$\pi_i = A + B(p/q)^i.$$

So, in this case, there is a two-parameter family of invariant measures – uniqueness up to scalar multiples does not hold.

Example 1.7.11

Consider a *success-run chain* on $\mathbb{Z}^+$, whose transition probabilities are given by

$$p_{i,i+1} = p_i, \quad p_{i0} = q_i = 1 - p_i.$$

Then the components of the invariant measure equation $\pi P = \pi$ read

$$\begin{aligned}\pi_0 &= \sum_{i=0}^{\infty} q_i \pi_i,\\ \pi_i &= p_{i-1}\pi_{i-1}, \quad \text{for } i \geq 1.\end{aligned}$$

Suppose we choose p_i converging sufficiently rapidly to 1 so that

$$p = \prod_{i=0}^{\infty} p_i > 0$$

which is equivalent to

$$\sum_{i=0}^{\infty} q_i = \infty.$$

Then for any solution of $\pi P = \pi$ we have

$$\pi_i = \left(\prod_{j=0}^{i-1} p_j \right) \pi_0 \geq p \pi_0$$

and so

$$\pi_0 \geq p\pi_0 \sum_{i=0}^{\infty} q_i.$$

This last equation forces either $\pi_0 = 0$ or $\pi_0 = \infty$, so there is no invariant measure.

Exercises

1.7.1 Find all invariant distributions of the transition matrix in Exercise 1.2.1.

1.7.2 Gas molecules move about randomly in a box which is divided into two halves symmetrically by a partition. A hole is made in the partition. Suppose there are N molecules in the box. Show that the number of molecules on one side of the partition just after a molecule has passed through the hole evolves as a Markov chain. What are the transition probabilities? What is the invariant distribution of this chain?

1.7.3 A particle moves on the eight vertices of a cube in the following way: at each step the particle is equally likely to move to each of the three adjacent vertices, independently of its past motion. Let i be the initial vertex occupied by the particle, o the vertex opposite i. Calculate each of the following quantities:

(i) the expected number of steps until the particle returns to i;
(ii) the expected number of visits to o until the first return to i;
(iii) the expected number of steps until the first visit to o.

1.7.4 Let $(X_n)_{n\geq 0}$ be a simple random walk on $\mathbb{Z}$ with $p_{i,i-1} = q < p = p_{i,i+1}$. Find

$$\gamma_i^0 = \mathbb{E}_0 \left(\sum_{n=0}^{T_0 - 1} 1_{\{X_n = i\}} \right)$$

and verify that

$$\gamma_i^0 = \inf_{\lambda} \lambda_i \quad \text{for all } i$$

where the infimum is taken over all invariant measures λ with $\lambda_0 = 1$. (Compare with Theorem 1.7.6 and Example 1.7.10.)

1.7.5 Let P be a stochastic matrix on a finite set I. Show that a distribution π is invariant for P if and only if $\pi(I - P + A) = a$, where $A = (a_{ij} : i, j \in I)$ with $a_{ij} = 1$ for all i and j, and $a = (a_i : i \in I)$ with $a_i = 1$ for all i. Deduce that if P

is irreducible then $I - P + A$ is invertible. *Note that this enables one to compute the invariant distribution by any standard method of inverting a matrix.*

1.8 Convergence to equilibrium

We shall investigate the limiting behaviour of the n-step transition probabilities $p_{ij}^{(n)}$ as $n \to \infty$. As we saw in Theorem 1.7.2, if the state-space is finite and if for some i the limit exists for all j, then it must be an invariant distribution. But, as the following example shows, the limit does not always exist.

Example 1.8.1

Consider the two-state chain with transition matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then $P^2 = I$, so $P^{2n} = I$ and $P^{2n+1} = P$ for all n. Thus $p_{ij}^{(n)}$ fails to converge for all i, j.

Let us call a state i *aperiodic* if $p_{ii}^{(n)} > 0$ for all sufficiently large n. We leave it as an exercise to show that i is aperiodic if and only if the set $\{n \geq 0 : p_{ii}^{(n)} > 0\}$ has no common divisor other than 1. This is also a consequence of Theorem 1.8.4. The behaviour of the chain in Example 1.8.1 is connected with its periodicity.

Lemma 1.8.2. *Suppose P is irreducible and has an aperiodic state i. Then, for all states j and k, $p_{jk}^{(n)} > 0$ for all sufficiently large n. In particular, all states are aperiodic.*

Proof. There exist $r, s \geq 0$ with $p_{ji}^{(r)}, p_{ik}^{(s)} > 0$. Then

$$p_{jk}^{(r+n+s)} \geq p_{ji}^{(r)} p_{ii}^{(n)} p_{ik}^{(s)} > 0$$

for all sufficiently large n. $\square$

Here is the main result of this section. The method of proof, by coupling two Markov chains, is ingenious.

Theorem 1.8.3 (Convergence to equilibrium). *Let P be irreducible and aperiodic, and suppose that P has an invariant distribution π. Let λ be any distribution. Suppose that $(X_n)_{n\geq 0}$ is Markov(λ, P). Then*

$$\mathbb{P}(X_n = j) \to \pi_j \quad \text{as } n \to \infty \text{ for all } j.$$

In particular

$$p_{ij}^{(n)} \to \pi_j \quad \text{as } n \to \infty \text{ for all } i, j.$$

Proof. We use a coupling argument. Let $(Y_n)_{n\geq 0}$ be Markov(π, P) and independent of $(X_n)_{n\geq 0}$. Fix a reference state b and set $T = \inf\{n \geq 1 : X_n = Y_n = b\}$.