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 $\lambda$ is any row vector $(\lambda_i : i \in I)$ with non-negative entries. We say $\lambda$ 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($\lambda, P$) and suppose that $\lambda$ is invariant for $P$. Then $(X_{m+n})_{n \geq 0}$ is also Markov($\lambda, P$).

**Proof.** By Theorem 1.1.3, $P(X_m = i) = (\lambda P^m)_i = \lambda_i$ 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

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

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

**Proof.** We have

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

and

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

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

Notice that for any of the random walks discussed in Section 1.6 we have $p^{(n)}_{ij} \to 0$ 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

$$P^n \rightarrow \left( \frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right) \text{ as } n \rightarrow \infty,$$

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 $\pi P = \pi$

$$\pi_1 = \frac{1}{2}\pi_3$$

$$\pi_2 = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_3$$

$$\pi_3 = \frac{1}{2}\pi_2 + \frac{1}{2}\pi_3.$$

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

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

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

According to Example 1.1.6

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

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 $\pi_1$ 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
1.7 Invariant distributions

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 \( k \):

\[
\gamma_i^k = E_k \sum_{n=0}^{T_{i-1}} I_{\{X_n = i\}}.
\]

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) \( \gamma_i^k = 1 \);
(ii) \( \gamma_i^k = (\gamma_i^k : i \in I) \) satisfies \( \gamma^k P = \gamma^k \);
(iii) \( 0 < \gamma_i^k < \infty \) for all \( i \in I \).

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

\[
P_k(X_{n-1} = i, X_n = j \text{ and } n \leq T_k) = P_k(X_{n-1} = i \text{ and } n \leq T_k)p_{ij}.
\]

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

\[
\gamma_j^k = E_k \sum_{n=1}^{T_k} I_{\{X_n = j\}} = E_k \sum_{n=1}^{\infty} I_{\{X_n = j \text{ and } n \leq T_k\}}
\]

\[
= \sum_{n=1}^{\infty} P_k(X_n = j \text{ and } n \leq T_k)
\]

\[
= \sum_{i \in I} \sum_{n=1}^{\infty} P_k(X_{n-1} = i, X_n = j \text{ and } n \leq T_k)
\]

\[
= \sum_{i \in I} p_{ij} \sum_{n=1}^{\infty} P_k(X_{n-1} = i \text{ and } n \leq T_k)
\]

\[
= \sum_{i \in I} p_{ij} E_k \sum_{m=0}^{T_k-1} I_{\{X_m = i \text{ and } m \leq T_k-1\}}
\]

\[
= \sum_{i \in I} p_{ij} E_k \sum_{m=0}^{T_k-1} I_{\{X_m = i\}} = \sum_{i \in I} \gamma_i^k p_{ij}.
\]

(iii) Since \( P \) is irreducible, for each state \( i \) there exist \( n, m \geq 0 \) with \( p_{ik}^{(n)} p_{ki}^{(m)} > 0 \).
Then \( \gamma_i^k \geq \gamma_k p_{ki}^{(n)} > 0 \) and \( \gamma_i^k p_{ki}^{(n)} \leq \gamma_k^k = 1 \) by (i) and (ii). \( \square \)
Theorem 1.7.6. Let $P$ be irreducible and let $\lambda$ be an invariant measure for $P$ with $\lambda_k = 1$. Then $\lambda \geq \gamma^k$. If in addition $P$ is recurrent, then $\lambda = \gamma^k$.

Proof. For each $j \in I$ we have
\[
\lambda_j = \sum_{i_0 \in I} \lambda_{i_0} p_{i_0,j} = \sum_{i_0 \neq k} \lambda_{i_0} p_{i_0,j} + p_{k,j} \\
= \sum_{i_0, i_1 \neq k} \lambda_{i_0} p_{i_0,i_1} p_{i_1,j} + \left( p_{k,j} + \sum_{i_0, i_1 \neq k} p_{k,i_0} p_{i_0,j} \right) \\
\ldots \\
= \sum_{i_0, \ldots, i_n \neq k} \lambda_{i_0} p_{i_0,i_1} \ldots p_{i_n,j} \\
+ \left( p_{k,j} + \sum_{i_0, \ldots, i_n \neq k} p_{k,i_0} p_{i_0,j} + \ldots + \sum_{i_0, \ldots, i_{n-1} \neq k} p_{k,i_{n-1}} \ldots p_{i_n,j} \right) \\
\geq P_k(X_1 = j \text{ and } T_k \geq 1) + P_k(X_2 = j \text{ and } T_k \geq 2) \\
\ldots + P_k(X_n = j \text{ and } T_k \geq n) \\
\to \gamma_j^k \text{ as } n \to \infty.
\]
So $\lambda \geq \gamma^k$. If $P$ is recurrent, then $\gamma^k$ 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 $0 = \mu_k = \sum_{j \in I} \mu_{ij} p_{ij}^{(n)} \geq \mu_i p_{ik}^{(n)}$, so $\mu_i = 0$. \(\square\)

Recall that a state $i$ is recurrent if
\[P_i(X_n = i \text{ for infinitely many } n) = 1\]
and we showed in Theorem 1.5.3 that this is equivalent to
\[P_i(T_i < \infty) = 1.\]
If in addition the expected return time
\[m_i = E_i(T_i)\]
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, $\pi$ 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, $\gamma^k$ is then invariant. But
\[\sum_{j \in I} \gamma_j^k = m_i < \infty\]

1.7 Invariant distributions

so \( \pi_j = \gamma_j^k/m_k \) defines an invariant distribution.

\[ (\text{iii}) \Rightarrow (\text{i}) \] Take any state \( k \). Since \( P \) is irreducible and \( \sum_{i \in I} \pi_i = 1 \) we have

\[ \pi_k = \sum_{i \in I} \pi_i \phi_i^{(n)} > 0 \text{ for some } n. \]

Set \( \lambda_i = \pi_i / \pi_k \). Then \( \lambda \) is an invariant measure with \( \lambda_k = 1 \). So by Theorem 1.7.6, \( \lambda \geq \gamma^k \). Hence

\[ m_k = \sum_{i \in I} \gamma_i^k \leq \sum_{i \in I} \pi_i = \frac{1}{\pi_k} < \infty \quad (1.7) \]

and \( k \) is positive recurrent.

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

Example 1.7.8 (Simple symmetric random walk on \( Z \))

The simple symmetric random walk on \( Z \) 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

\[ \pi_i = \frac{1}{2} \pi_{i-1} + \frac{1}{2} \pi_{i+1}, \]

so \( \pi \) is invariant. Now Theorem 1.7.6 forces any invariant measure to be a scalar multiple of \( \pi \). Since \( \sum_{i \in Z} \pi_i = \infty \), 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 \( Z^3 \), which is transient by Example 1.6.3, but has invariant measure \( \pi \) given by \( \pi_i = 1 \) for all \( i \).

Example 1.7.10

Consider the asymmetric random walk on \( Z \) 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 \( \pi \) 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 \( Z^+ \), whose transition probabilities are given by

\[ p_{i,i+1} = p_i, \quad p_0 = q_i = 1 - p_i. \]
Then the components of the invariant measure equation \( \pi P = \pi \) read
\[
\pi_0 = \sum_{i=0}^{\infty} q_i \pi_i,
\]
\[
\pi_i = p_{i-1} \pi_{i-1}, \quad \text{for } i \geq 1.
\]

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
\[
\pi_i = \left( \prod_{j=0}^{i-1} p_j \right) \pi_0 \geq p^n \pi_0
\]
and so
\[
\pi_0 \geq p^n \sum_{i=0}^{\infty} q_i.
\]

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
\[
\gamma_i^0 = E_0 \left( \sum_{n=0}^{T_i-1} 1\{X_n = i\} \right)
\]
and verify that
\[ \gamma_i^0 = \inf_{\lambda} \lambda_i \quad \text{for all } i \]
where the infimum is taken over all invariant measures \( \lambda \) 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 \( \pi \) 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.