What's Up With π?

Lemma (π1): If a stationary distribution π exists, then all states j that have $\pi_j > 0$ are recurrent.

Proof: We want to show that $g_j = 1$.

Let N_j be the number of visits to state j at times ≥ 1. Then

$$N_j = \sum_{n=1}^{\infty} I_{\{X_n=j\}}.$$

So,

$$\begin{aligned} \mathsf{E}_i[N_j] &= \mathsf{E}_i[\textstyle\sum_{n=1}^{\infty} I_{\{X_n=j\}}] = \sum_{n=1}^{\infty} \mathsf{E}_i[I_{\{X_n=j\}}] \\ &= \textstyle\sum_{n=1}^{\infty} P_i(X_n = j) \sum_{n=1}^{\infty} P(X_n = j | X_0 = i) \\ &= \textstyle\sum_{n=1}^{\infty} p_{ij}^{(n)}. \end{aligned}$$

Consider the quantity $\sum_i \pi_i \mathsf{E}_i[N_j]$:

$$\begin{aligned} \textstyle\sum_i \pi_i \mathsf{E}_i[N_j] &= \textstyle\sum_i \pi_i \sum_{n=1}^{\infty} p_{ij}^{(n)} = \sum_{n=1}^{\infty} \sum_i \pi_i p_{ij}^{(n)} \\ &\overset{station.}{=} \textstyle\sum_{n=1}^{\infty} \pi_j = \infty \end{aligned}$$

since $\pi_j > 0$.

On the other hand, since $\mathsf{E}_i[N_j] = g_{ij}/(1-g_j)$, we have

$$\infty = \textstyle\sum_i \pi_i \mathsf{E}_i[N_j] = \sum_i \pi_i \frac{g_{ij}}{1-g_j} \leq \sum_i \pi_i \frac{1}{1-g_j}$$

since $g_{ij} \leq 1$. So,

$$\infty = \tfrac{1}{1-g_j} \textstyle\sum_i \pi_i = \frac{1}{1-g_j} \cdot 1 = \frac{1}{1-g_j}$$

which implies that $g_j = 1$, as desired.

Define, for any state i, the set

$$A_i = \{n : p_{ii}^{(n)} > 0\}.$$

Note then that the period of i, which we have denoted by d_i is the greatest common divisor of all of the elements in A_i.

Claim: If i has period 1, then there is a number K such that, for all $n \geq K$, n is in A_i.

Partial Proof: We will quote, without proof, a result from number theory that says

> "If the greatest common divisor of a set A_i is 1, then there are integers $i_1, i_2, \ldots, i_m$ in A_i and positive or negative integer coefficients $c_1, c_2, \ldots, c_m$ such that $c_1 i_1 + c_2 i_2 + \cdots + i_m c_m = 1$."

in order to show that A_i contains two consecutive integers. Then we will show that A_i containing two consecutive integers gives us the result we want.

Proof Step 1:

If i has period 1, then, by definition, the greatest common divisor of the elements in A_i is 1. So, by the claim quoted above, there are integers $i_1, i_2, \ldots, i_m$ in A_i and positive or negative integer coefficients $c_1, c_2, \ldots, c_m$ such that $c_1 i_1 + c_2 i_2 + \cdots + i_m c_m = 1$.

Let j be the number of positive coefficients and let k be the number of negative coefficients. (So $j + k = m$.) Let $s_1, s_2, \ldots, s_j$ be the subscripts of the i's in $i_1, i_2, \ldots, i_m$ with positive coefficients and let $t_1, t_2, \ldots, t_k$ be the subscripts of the i's with negative coefficients. Define $a_l = c_{s_l}$ for $l = 1, 2, \ldots, j$ and $b_l = -c_{t_l}$ for $l = 1, 2, \ldots, k$.

Then, we have

$$a_1 i_{s_1} + \cdots + a_j i_{s_j} = b_1 i_{t_1} + \ldots + b_k i_{t_k} + 1. \tag{1}$$

Note that any positive linear combination of elements in A_i is also in A_i. For example, if you can, with positive probability, go from state i to state i in 3 steps and you can also go in 7 steps, then 3 and 7 are numbers in A_i. Furthermore, $4 \cdot 3 + 2 \cdot 7$, for example, is also in A_i since you can go from i to i in 3 steps, then another 3 steps, then another 3 steps, then another 3 steps, then in 7 steps, and then another 7 steps.

So, (1) shows us two consecutive integers in A_i.

Proof Step 2:

We have now shown that there are two consecutive integers, say k and $k+1$ in A_i. We now wish to conclude that this implies that all integers, after some point are in A_i.

If k and $k+1$ are in A_i, then so are

$$2k, \;\; 2k+1, \;\; \text{and} \;\; 2k+2$$

since we can go from i to i in k steps plus another k steps or in k steps plus another $k+1$ steps, or in $k+1$ steps plus another $k+1$ steps.

Since $2k, 2k+1, 2k+2$ are in A_i, then so are

$$4k, \;\; 4k+1, \;\; 4k+2, \;\; 4k+3, \;\; \text{and} \;\; 4k+4.$$

as these are all the distinct sums of pairs from $2k, 2k+1, 2k+2$.

Continuing, we get that k and $k+1$ in A_i implies that

$$jk, jk+1, \ldots, jk+j$$

are in A_i for any positive integer j.

For $j \geq k-1$, these blocks of numbers included in A_i will start to overlap, thereby leaving no gaps in the remaining sequence of integers included in A_i.

We will use this claim to help us prove the following Theorem.

Theorem (π1): Suppose that a Markov chain defined by the transition probabilities p_{ij} is irreducible, aperiodic, and has stationary distribution π. Then for all states i and j,

$$p_{ij}^{(n)} \to \pi_j, \qquad \text{as } n \to \infty.$$

Proof:

1. Let S denote the state space and let $\{X_n\}$ and $\{Y_n\}$ denote two independent copies of the Markov chain.

   Consider the bivariate Markov chain $\{(X_n, Y_n)\}$ on $S^2 = S \times S$ and let the transition probabilities be denoted by

   $$p_{(i_x,i_y),(j_x,j_y)}.$$

   Note that, by defintion of $\{X_n\}$ and $\{Y_n\}$ and their independence,

   $$p_{(i_x,i_y),(j_x,j_y)} = p_{i_x,j_x} \cdot p_{i_y,j_y}.$$

   We are going to show that

   $$|P(X_n = j) - P(Y_n = j)| \to 0, \qquad \text{as } n \to \infty \tag{2}$$

   regardless of the starting values of $\{X_n\}$ and $\{Y_n\}$. So, we are able to take $X_0 = i$ and Y_0 to be a random variable with distribution π, and then by (2) we will have

   $$|p_{ij}^{(n)} - \pi_j| \to 0, \qquad \text{as } n \to \infty.$$

2. Claim: This bivariate Markov chain is irreducible.

   Proof of Claim:

   We want to take any states (i_x, i_y) and (j_x, j_y) and find an integer l such that

   $$p_{(i_x,i_y),(j_x,j_y)}^{(l)} > 0.$$

   - Take any states (i_x, i_y) and (j_x, j_y) in S^2.

- Since the original chain is irreducible, $i_x \leftrightarrow j_x$ and $i_y \leftrightarrow j_y$. ie: There exist integers n and m such that $p^{(n)}_{i_x,j_x} > 0$ and $p^{(m)}_{i_y,j_y} > 0$. (There are also two more integers that reverse these transitions, but we don't care about them.)
- Since the original chain is aperiodic, states j_x and j_y have period 1. Hence, by the claim preceding this Theorem, there exists a K such that

$$p^{(k+K)}_{j_x,j_x} > 0 \qquad \text{and} \qquad p^{(k+K)}_{j_y,j_y} > 0$$

  for all k.

  Specifically,

$$p^{(m+K)}_{j_x,j_x} > 0 \qquad \text{and} \qquad p^{(n+K)}_{j_y,j_y} > 0.$$

- Therefore,

$$p^{(n+m+K)}_{(i_x,i_y),(j_x,j_y)} > 0$$

  since the components move independently.

3. Since the two coordinates are independent,

$$\pi_{(i_x,i_y)} = \pi_{i_x} \cdot \pi_{i_y}$$

defines a stationary distribution for the Markov chain.

Proof:

We need to show that

$$\pi_{(j_x,j_y)} = \sum_{i_x,i_y} \pi_{(i_x,i_y)} p_{(i_x,i_y),(j_x,j_y)}$$

Well,

$$\begin{aligned}
\textstyle\sum_{i_x,i_y} \pi_{(i_x,i_y)} p_{(i_x,i_y),(j_x,j_y)} &= \textstyle\sum_{i_x} \sum_{i_y} \pi_{(i_x,i_y)} p_{(i_x,i_y),(j_x,j_y)} \\
&= \textstyle\sum_{i_x} \sum_{i_y} \pi_{i_x} \pi_{i_y} p_{i_x,j_x} p_{i_y,j_y} \\
&= \textstyle\sum_{i_x} \pi_{i_x} p_{i_x,j_x} \sum_{i_y} \pi_{i_y} p_{i_y,j_y} \\
&= \textstyle\sum_{i_x} \pi_{i_x} p_{i_x,j_x} \pi_{j_y} \qquad (\pi \text{ stationary}) \\
&= \textstyle\pi_{j_y} \sum_{i_x} \pi_{i_x} p_{i_x,j_x} \\
&= \pi_{j_y} \pi_{j_x} \stackrel{def}{=} \pi_{(j_x,j_y)} \qquad (\pi \text{ stationary})
\end{aligned}$$

4. Since $\pi_{(i_x,i_y)}$ is stationary for the bivariate chain, if we can show that $\pi_{(i_x,i_y)} > 0$ for all states (i_x, i_y), we will have, by Lemma (π1), that all states in the bivariate chain are recurrent.

Proof of $\pi_{(i_x,i_y)} > 0$:

- $$\pi_{(i_x,i_y)} = \pi_{i_x} \cdot \pi_{i_y}$$
  So, we need to show that $\pi_{i_x} > 0$ and $\pi_{i_y} > 0$ for all states i_x, i_y in S.
  ie: We need to show $\pi_j > 0$ for all $j \in S$.
- Since π is a distribution, $\sum_i \pi_i = 1$ implies that there is at least one state i^* such that $\pi_{i^*} > 0$.
  Since π is a stationary distribution, we have that
  $$\pi_j = \sum_i \pi_i p_{i,j}^{(n)}$$
  for any fixed time point n.
- Choose n so that $p_{i^*,j}^{(n)} > 0$. We can do this since the original Markov chain is irreducible.
- Then we have that
  $$\pi_j = \sum_i \pi_i p_{i,j}^{(n)} \geq \pi_{i^*} p_{i^*,j}^{(n)} > 0.$$

5. Let $T = \min\{n \geq 0 : X_n = Y_n\}$ and let $T_{(x)} = \min\{n \geq 0 : X_n = Y_n = x\}$.
   Since the bivariate chain $\{(X_n, Y_n)\}$ is irreducible (can get to (x,x)) and recurrent (will get to (x,x)), we have that
   $$T_{(x)} < \infty \qquad \text{with probability 1.}$$
   So,
   $$T < T_{(x)} < \infty \qquad \text{with probability 1.}$$

6. Claim: $P(X_n = j, T \leq n) = P(Y_n = j, T \leq n)$.
   ("On $\{T \leq n\}$, X_n and Y_n have the same distribution.")
   Proof:
   $$\begin{aligned} P(X_n = j, T \leq n) &= \textstyle\sum_{u=0}^n P(X_n = j, T = u) \\ &= \textstyle\sum_{u=0}^n \sum_i P(X_n = j, X_u = i, T = u) \\ &= \textstyle\sum_{u=0}^n \sum_i P(X_n = j | X_u = i, T = u) \cdot P(X_u = i, T = u) \\ &\overset{M.P.}{=} \textstyle\sum_{u=0}^n \sum_i P(X_n = j | X_u = i) \cdot P(X_u = i, T = u) \\ &= \textstyle\sum_{u=0}^n \sum_i P(Y_n = j | Y_u = i) \cdot P(Y_u = i, T = u) \\ &= P(Y_n = j, T \leq n) \end{aligned}$$
   In the second to last equality, the first factor came from the fact that $\{X_n\}$ and $\{Y_n\}$ have the same transition law. The second factor came from the fact that at time T, $X_n = Y_n$.

7. Note that

$$\begin{aligned} P(X_n = j) &= P(X_n = j, T \le n) + P(X_n = j, T > n) \\ &\overset{\textit{Step 6}}{=} P(Y_n = j, T \le n) + P(X_n = j, T > n) \\ &\le P(Y_n = j) + P(X_n = j, T > n). \end{aligned}$$

Similarly, we have that

$$P(Y_n = j) \le P(X_n = j) + P(X_n = j, T > n).$$

So,

$$|P(X_n = j) - P(Y_n = j)| \le P(X_n = j, T > n) + P(Y_n = j, T > n)$$

8. Summing over j, we get

$$\sum_j |P(X_n = j) - P(Y_n = j)| \le 2 \cdot P(T > n)$$

regardless of the intial values for $\{X_n\}$ and $\{Y_n\}$.

9. Therefore, if we let $X_0 = i$ and $Y_0 \sim \pi$, we get

$$\sum_j |p_{i,j}^{(n)} - \pi_j| \le 2 \cdot P(T > n) \to 0 \qquad \text{as } n \to \infty$$

since $T < \infty$ with probability 1.

Therefore

$$p_{i,j}^n \to \pi_j \qquad \text{as } n \to \infty$$

for all $i, j \in S$.