1.4 Strong Markov property

In Section 1.1 we proved the Markov property. This says that for each time m, conditional on $X_m = i$, the process after time m begins afresh from i. Suppose, instead of conditioning on $X_m = i$, we simply waited for the process to hit state i, at some random time H. What can one say about the process after time H? What if we replaced H by a more general random time, for example $H-1$? In this section we shall identify a class of random times at which a version of the Markov property does hold. This class will include H but not $H-1$; after all, the process after time $H-1$ jumps straight to i, so it does not simply begin afresh.

A random variable $T : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$ is called a *stopping time* if the event $\{T = n\}$ depends only on $X_0, X_1, \dots, X_n$ for $n = 0, 1, 2, \dots$. Intuitively, by watching the process, you know at the time when T occurs. If asked to stop at T, you know when to stop.

Examples 1.4.1

(a) The *first passage time*

$$T_j = \inf\{n \geq 1 : X_n = j\}$$

is a stopping time because

$$\{T_j = n\} = \{X_1 \neq j, \dots, X_{n-1} \neq j, X_n = j\}.$$

(b) The first hitting time H^A of Section 1.3 is a stopping time because

$$\{H^A = n\} = \{X_0 \notin A, \dots, X_{n-1} \notin A, X_n \in A\}.$$

(c) The *last exit time*

$$L^A = \sup\{n \geq 0 : X_n \in A\}$$

is not in general a stopping time because the event $\{L^A = n\}$ depends on whether $(X_{n+m})_{m \geq 1}$ visits A or not.

We shall show that the Markov property holds at stopping times. The crucial point is that, if T is a stopping time and $B \subseteq \Omega$ is determined by $X_0, X_1, \dots, X_T$, then $B \cap \{T = m\}$ is determined by $X_0, X_1, \dots, X_m$, for all $m = 0, 1, 2, \dots$.

Theorem 1.4.2 (Strong Markov property). *Let $(X_n)_{n \geq 0}$ be Markov(λ, P) and let T be a stopping time of $(X_n)_{n \geq 0}$. Then, conditional on $T < \infty$ and $X_T = i$, $(X_{T+n})_{n \geq 0}$ is Markov(δ_i, P) and independent of $X_0, X_1, \dots, X_T$.*

Proof. If B is an event determined by $X_0, X_1, \dots, X_T$, then $B \cap \{T = m\}$ is determined by $X_0, X_1, \dots, X_m$, so, by the Markov property at time m

$$\begin{aligned} &P(\{X_T = j_0, X_{T+1} = j_1, \dots, X_{T+n} = j_n\} \cap B \cap \{T = m\} \cap \{X_T = i\}) \\ &\quad = P_i(X_0 = j_0, X_1 = j_1, \dots, X_n = j_n) P(B \cap \{T = m\} \cap \{X_T = i\}) \end{aligned}$$

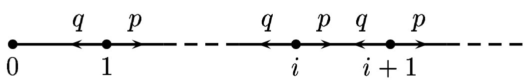
where we have used the condition $T = m$ to replace m by T. Now sum over $m = 0, 1, 2, \ldots$ and divide by $P(T < \infty, X_T = i)$ to obtain

$$\begin{aligned} &P(\{X_T = j_0, X_{T+1} = j_1, \ldots, X_{T+n} = j_n\} \cap B \mid T < \infty, X_T = i) \\ &\quad = P_i(X_0 = j_0, X_1 = j_1, \ldots, X_n = j_n) P(B \mid T < \infty, X_T = i). \qquad \square \end{aligned}$$

The following example uses the strong Markov property to get more information on the hitting times of the chain considered in Example 1.3.3.

Example 1.4.3

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



where $0 < p = 1 - q < 1$. We know from Example 1.3.3 the probability of hitting 0 starting from 1. Here we obtain the complete distribution of the time to hit 0 starting from 1 in terms of its probability generating function. Set

$$H_j = \inf\{n \geq 0 : X_n = j\}$$

and, for $0 \leq s < 1$

$$\phi(s) = E_1(s^{H_0}) = \sum_{n < \infty} s^n P_1(H_0 = n).$$

Suppose we start at 2. Apply the strong Markov property at H_1 to see that under P_2, conditional on $H_1 < \infty$, we have $H_0 = H_1 + \widetilde{H}_0$, where $\widetilde{H}_0$, the time taken after H_1 to get to 0, is independent of H_1 and has the (unconditioned) distribution of H_1. So

$$\begin{aligned} E_2(s^{H_0}) &= E_2(s^{H_1} \mid H_1 < \infty) E_2(s^{\widetilde{H}_0} \mid H_1 < \infty) P_2(H_1 < \infty) \\ &= E_2(s^{H_1} 1_{H_1 < \infty}) E_2(s^{\widetilde{H}_0} \mid H_1 < \infty) \\ &= E_2(s^{H_1})^2 = \phi(s)^2. \end{aligned}$$

Then, by the Markov property at time 1, conditional on $X_1 = 2$, we have $H_0 = 1 + \overline{H}_0$, where $\overline{H}_0$, the time taken after time 1 to get to 0, has the same distribution as H_0 does under P_2. So

$$\begin{aligned} \phi(s) = E_1(s^{H_0}) &= p E_1(s^{H_0} \mid X_1 = 2) + q E_1(s^{H_0} \mid X_1 = 0) \\ &= p E_1(s^{1 + \overline{H}_0} \mid X_1 = 2) + q E_1(s \mid X_1 = 0) \\ &= p s E_2(s^{H_0}) + qs \\ &= p s \phi(s)^2 + qs. \end{aligned}$$

Thus $\phi = \phi(s)$ satisfies

$$ps\phi^2 - \phi + qs = 0 \tag{1.5}$$

and

$$\phi = (1 \pm \sqrt{1 - 4pqs^2})/2ps.$$

Since $\phi(0) \le 1$ and ϕ is continuous we are forced to take the negative root at $s = 0$ and stick with it for all $0 \le s < 1$.

To recover the distribution of H_0 we expand the square-root as a power series:

$$\begin{aligned}\phi(s) &= \frac{1}{2ps}\left\{1 - \left(1 + \tfrac{1}{2}(-4pqs^2) + \tfrac{1}{2}(-\tfrac{1}{2})(-4pqs^2)^2/2! + \dots\right)\right\}\\ &= qs + pq^2s^3 + \dots\\ &= sP_1(H_0 = 1) + s^2P_1(H_0 = 2) + s^3P_1(H_0 = 3) + \dots .\end{aligned}$$

The first few probabilities $P_1(H_0 = 1), P_1(H_0 = 2), \dots$ are readily checked from first principles.

On letting $s \uparrow 1$ we have $\phi(s) \to P_1(H_0 < \infty)$, so

$$P_1(H_0 < \infty) = \frac{1 - \sqrt{1 - 4pq}}{2p} = \begin{cases} 1 & \text{if } p \le q \\ q/p & \text{if } p > q. \end{cases}$$

(Remember that $q = 1 - p$, so

$$\sqrt{1 - 4pq} = \sqrt{1 - 4p + 4p^2} = |1 - 2p| = |2q - 1|. \quad)$$

We can also find the mean hitting time using

$$E_1(H_0) = \lim_{s \uparrow 1} \phi'(s).$$

It is only worth considering the case $p \le q$, where the mean hitting time has a chance of being finite. Differentiate (1.5) to obtain

$$2ps\phi\phi' + p\phi^2 - \phi' + q = 0$$

so

$$\phi'(s) = (p\phi(s)^2 + q)/(1 - 2ps\phi(s)) \to 1/(1 - 2p) = 1/(q - p) \quad \text{as } s \uparrow 1.$$

See Example 5.1.1 for a connection with branching processes.

Example 1.4.4

We now consider an application of the strong Markov property to a Markov chain $(X_n)_{n \ge 0}$ observed only at certain times. In the first instance suppose that J is some subset of the state-space I and that we observe the chain only when it takes values in J. The resulting process $(Y_m)_{m \ge 0}$ may be obtained formally by setting $Y_m = X_{T_m}$, where

$$T_0 = \inf\{n \ge 0 : X_n \in J\}$$

and, for $m = 0, 1, 2, \dots$

$$T_{m+1} = \inf\{n > T_m : X_n \in J\}.$$

Let us assume that $\mathbb{P}(T_m < \infty) = 1$ for all m. For each m we can check easily that T_m, the time of the mth visit to J, is a stopping time. So the strong Markov property applies to show, for $i_0, \dots, i_{m+1} \in J$, that

$$\begin{aligned} &P(Y_{m+1} = i_{m+1} \mid Y_0 = i_0, \dots, Y_m = i_m) \\ &= P(X_{T_{m+1}} = i_{m+1} \mid X_{T_0} = i_0, \dots, X_{T_m} = i_m) \\ &= P_{i_m}(X_{T_1} = i_{m+1}) = \overline{p}_{i_m i_{m+1}} \end{aligned}$$

where, for $i, j \in J$

$$\overline{p}_{ij} = h_i^j$$

and where, for $j \in J$, the vector $(h_i^j : i \in I)$ is the minimal non-negative solution to

$$h_i^j = p_{ij} + \sum_{k \notin J} p_{ik} h_k^j. \tag{1.6}$$

Thus $(Y_m)_{m \geq 0}$ is a Markov chain on J with transition matrix $\overline{P}$.

A second example of a similar type arises if we observe the original chain $(X_n)_{n \geq 0}$ only when it moves. The resulting process $(Z_m)_{m \geq 0}$ is given by $Z_m = X_{S_m}$ where $S_0 = 0$ and for $m = 0, 1, 2, \dots$

$$S_{m+1} = \inf\{n \geq S_m : X_n \neq X_{S_m}\}.$$

Let us assume there are no absorbing states. Again the random times S_m for $m \geq 0$ are stopping times and, by the strong Markov property

$$\begin{aligned} &P(Z_{m+1} = i_{m+1} \mid Z_0 = i_0, \dots, Z_m = i_m) \\ &= P(X_{S_{m+1}} = i_{m+1} \mid X_{S_0} = i_0, \dots, X_{S_m} = i_m) \\ &= P_{i_m}(X_{S_1} = i_{m+1}) = \widetilde{p}_{i_m i_{m+1}} \end{aligned}$$

where $\widetilde{p}_{ii} = 0$ and, for $i \neq j$

$$\widetilde{p}_{ij} = p_{ij} / \sum_{k \neq i} p_{ik}.$$

Thus $(Z_m)_{m \geq 0}$ is a Markov chain on I with transition matrix $\widetilde{P}$.

Exercises

1.4.1 Let $Y_1, Y_2, \dots$ be independent identically distributed random variables with $P(Y_1 = 1) = P(Y_1 = -1) = 1/2$ and set $X_0 = 1$, $X_n = X_0 + Y_1 + \dots + Y_n$ for $n \geq 1$. Define

$$H_0 = \inf\{n \geq 0 : X_n = 0\}.$$

Find the probability generating function $\phi(s) = E(s^{H_0})$.

Suppose the distribution of $Y_1, Y_2, \dots$ is changed to $P(Y_1 = 2) = P(Y_1 = -1) = 1/2$. Show that ϕ now satisfies

$$s\phi^3 - 2\phi + s = 0.$$

1.4.2 Deduce carefully from Theorem 1.3.2 the claim made at (1.6).