

Chapter 2. The strong Markov property

Hereafter, we always consider time-homogeneous transition probabilities $p_{x,y}$ on a finite or countable state space S (although most of the ideas generalize to a rather general state space S). $\mathbf{X} = (X_0, X_1, \dots)$ denotes the Markov chain, and \mathbb{P}_μ the probability measure when \mathbf{X} starts with distribution μ , $X_0 \sim \mu$. $\mathcal{B}(S)$ is the space of bounded measurable functions $f: S \rightarrow \mathbb{R}$ and $\mathcal{P}(S)$ is the space of probability measures on S .

We will think that the transition probabilities $p_{x,y}$ are fixed once and for all. We then define the space $D(S)$ of sequences (x_i) in S , such that $p_{x_i, x_{i+1}} > 0$ for all $i \geq 0$. Finally recall that P is the Markov operator $Pf(x) := \mathbb{E}_x[f(X_1)]$.

In this chapter we aim to formalize some intuitive properties of a Markov chain. For instance, conditioning on the intermediate times we have for $f \in \mathcal{B}(S)$

$$\mathbb{E}_x[f(X_t)] = (P^t f)(x)$$

and for $t \geq s$, regardless of $\mu \in \mathcal{P}(S)$

$$\mathbb{E}_\mu[f(X_t) \mid X_s = x] = (P^{t-s} f)(x)$$

or say, for $f \in \mathcal{B}(S)$

$$\mathbb{E}_\mu[f(X_t, X_{t+1}) \mid X_s = x] = \mathbb{E}_x[f(X_{t-s}, X_{t-s+1})]$$

In this chapter we see how to properly write these kinds of properties and many others in a general framework.

The (homogeneous) Markov property

First, we introduce the **time shift** operator θ_t , for $t \geq 0$

$$\begin{aligned} \theta_t: D(S) &\rightarrow D(S) \\ (\theta_t \mathbf{X})_s &:= \mathbf{X}_{t+s} \end{aligned} \tag{1}$$

Proposition 0.1 (Homogeneous Markov property). *Let \mathbf{X} be a homogeneous Markov chain. Then for $s \geq 0$ and for all bounded measurable functions $F: D(S) \rightarrow \mathbb{R}$*

$$\mathbb{E}_\mu[F(\theta_s \mathbf{X}) \mid X_s = x] = \mathbb{E}_\mu[F(\theta_s \mathbf{X}) \mid X_0 = x_0, \dots, X_{s-1} = x_{s-1}, X_s = x] = \mathbb{E}_x[F(\mathbf{X})]$$

In other words, conditional on $X_t = x$, the sequence (X_t, X_{t+1}, \dots) is a Markov chain starting at x , with the same transition probabilities as \mathbf{X} , and independent of (X_0, \dots, X_{t-1}) .

Exercise 0.1. Proposition 0.1 may sound abstract, but it is a rather obvious statement. Use the [previous example](#) about the random walker, to make a trivial example of this statement. Take the transition probabilities $p_{x,y}$ homogeneous, and let $f(x, y)$ be the price of the flight from town x to town y . Let $F(\mathbf{X})$ be the cost paid for the first 10 flights. What does the proposition say?

Proposition 0.1. With no loss of generality, to keep the notation explicit, we can assume that F depends on finitely many coordinates $F(\mathbf{X}) = F(X_0, X_1, X_2, \dots, X_t)$, since measurable functions of $D(S)$ are limit of such functions. By linearity in F , it is enough to prove the equality for indicator functions, namely for $F(X_0, X_1, X_2, \dots, X_t) = \mathbf{1}_{z_0}(X_0)\mathbf{1}_{z_1}(X_1) \cdots \mathbf{1}_{z_t}(X_t)$ for some $z_0, \dots, z_t \in S$. Then for such an F

$$\begin{aligned} \mathbb{E}_\mu[F(\theta_s \mathbf{X}) \mid X_0 = x_0, \dots, X_{s-1} = x_{s-1}, X_s = x] &= \mathbb{P}_\mu(X_s = z_0, X_{s+1} = z_1, \dots, X_{s+t} = z_t \mid X_0 = x_0, \dots, X_{s-1} = x_{s-1}, X_s = x) \\ &= \mathbf{1}_{z_0}(x) \frac{p_{x_0, x_1} \cdots p_{x_{s-1}, x} p_{z_0, z_1} \cdots p_{z_{t-1}, z_t}}{p_{x_0, x_1} \cdots p_{x_{s-1}, x}} = \mathbf{1}_{z_0}(x) p_{z_0, z_1} \cdots p_{z_{t-1}, z_t} = \end{aligned}$$

□

Stopping times

Definition 0.1 (Stopping times). A map $\tau: \Omega \rightarrow \mathbb{N} \cup \{\infty\}$ is called a stopping time if, for each $t \geq 0$, the value $\{\tau = t\}$ is determined by (X_0, \dots, X_t) . That is if $\mathbf{1}_{\{\tau=t\}}$ is a (measurable) function of (X_0, X_1, \dots, X_t) .

Remark. Notice that, since $\{\tau = t\} = \{\tau \leq t\} \setminus \{\tau \leq t-1\}$, it is equivalent to say that τ is a stopping time if $\{\tau \leq t\}$ is determined by (X_0, \dots, X_t) .

For instance, for $A \subset S$, the **hitting time** (discussed more accurately in the next chapter) of the set A

$$\tau_A := \inf\{t \geq 0 : X_t \in A\}$$

is a stopping time: $\{\tau_A = t\} = \{X_0 \notin A, \dots, X_{t-1} \notin A, X_t \in A\}$. However $\tau_A - 1$ is not (in general) a stopping time, since $\tau_A - 1 = t$ depends on X_{t+1} .

Exercise 0.2. Let $\tau_A^+ := \inf\{t \geq 1 : X_t \in A\}$ be the first return time to A . Check that τ_A^+ is a stopping time.

Let $\xi_A := \sup\{t \geq 0 : X_t \in A\}$ be the time of the last visit in A . Check that ξ_A is **not**, in general, a stopping time.

Exercise 0.3. Let $\tau_A^{(0)} = 0$, and for $k \geq 1$

$$\tau_A^{(k)} := \inf\{t > \tau_A^{(k-1)} : X_t \in A\}$$

be the k -th time the process visits A . Is $\tau_A^{(k)}$ a stopping time?

The Strong Markov Property

Theorem 0.1 (Strong Markov property). Let σ be a finite stopping time: $\mathbb{P}(\sigma < \infty) = 1$. Then a stronger version of the **Markov property** holds: for all bounded measurable functions $F: D(S) \rightarrow \mathbb{R}$.

$$\mathbb{E}_\mu[F(\theta_\sigma \mathbf{X}) \mid (X_r)_{r \leq \sigma}] = \mathbb{E}_{X_\sigma}[F(\mathbf{X})]$$

In particular

$$\mathbb{E}_\mu[F(\theta_\sigma \mathbf{X}) \mid X_\sigma = x] = \mathbb{E}_x[F(\mathbf{X})]$$

In other words, conditional on $X_\sigma = x$, $(X_\sigma, X_{\sigma+1}, \dots)$ is a Markov chain starting at x , independent of $X_0, \dots, X_{\sigma-1}$, and with the same transition probabilities as \mathbf{X} .

Proof. For all $(x_0, x_1, \dots) \in D(S)$ and $x \in S$

$$\begin{aligned} \mathbb{E}_\mu[F(\theta_\sigma \mathbf{X}) \mid X_0 = x_0, \dots, X_\sigma = x] &= \frac{\sum_{t=0}^{\infty} \mathbb{E}_\mu[F(\theta_t \mathbf{X}) \mathbf{1}_{\sigma=t} \mathbf{1}_{X_0=x_0, \dots, X_t=x}]}{\sum_{t=0}^{\infty} \mathbb{E}_\mu[\mathbf{1}_{\sigma=t} \mathbf{1}_{X_0=x_0, \dots, X_t=x}]} \\ &= \frac{\sum_{t=0}^{\infty} \mathbb{E}_\mu[F(\theta_t \mathbf{X}) \mid \sigma = t, X_0 = x_0, \dots, X_t = x] \mathbb{P}_\mu(\sigma = t, X_0 = x_0, \dots, X_t = x)}{\sum_{t=0}^{\infty} \mathbb{P}_\mu(\sigma = t, X_0 = x_0, \dots, X_t = x)} \end{aligned}$$

where we understand $\mathbb{E}[F|A]\mathbb{P}(A) = 0$ whenever $\mathbb{P}(A) = 0$. Since σ is a stopping time, $\{\sigma = t\}$ only depends on (X_0, \dots, X_t) . Thus by the [Markov Property](#)

$$\mathbb{E}_\mu[F(\theta_t \mathbf{X}) \mid \sigma = t, X_0 = x_0, \dots, X_t = x] = \mathbb{E}_x[F(\mathbf{X})]$$

from which we easily conclude, plugging the last equality in the above formula. □

The trace of a Markov chain

Let $A \subset S$. Suppose that we can only observe the Markov Chain inside A , namely we define

$$Y_t := X_{\tau_A^{(t+1)}} \tag{2}$$

with $\tau_A^{(t)}$ defined in [Exercise 0.3](#).

The Markov chain \mathbf{Y} is called the **trace of \mathbf{X} on A** .

Proposition 0.2 (Trace of a Markov Chain). *If $\mathbb{P}(\tau_A^{(t)} < \infty) = 1$ for all $t \geq 1$, the process $\mathbf{Y} = (Y_0, Y_1, \dots)$ defined in [Equation 2](#) is a Markov chain with transition probabilities $q_{x,y} = \mathbb{P}_x(X_{\tau_A^+} = y)$, where $\tau_A^+ := \inf\{t \geq 1 : X_t \in A\}$ is the stopping time defined in [Exercise 0.2](#).*

Proof. For $t \geq 0, y, y_0, \dots, y_t \in A$

$$\mathbb{P}(Y_{t+1} = y \mid Y_0 = y_0, \dots, Y_t = y_t) = \mathbb{P}(X_{\tau_A^{(t+2)}} = y \mid X_{\tau_A^{(1)}} = y_0, \dots, X_{\tau_A^{(t+1)}} = y_t)$$

However:

- Since the $\tau_A^{(s)}$ are stopping times, $\mathbf{1}_{X_{\tau_A^{(1)}}=y_0, \dots, X_{\tau_A^{(t+1)}}=y_t}$ is a function of $(X_0, \dots, X_{\tau_A^{(t+1)}})$.
- We have $\mathbf{1}_{X_{\tau_A^{(t+2)}}=y} = \theta_{\tau_A^{(t+1)}} \mathbf{1}_{X_{\tau_A^+}=y}$.

We can therefore apply the [Strong Markov property](#) to the function $F(\mathbf{X}) = \mathbf{1}_{X_{\tau_A^+}=y}$ to obtain

$$\mathbb{P}(Y_{t+1} = y \mid Y_0 = y_0, \dots, Y_t = y_t) = \mathbb{E}[\theta_{\tau_A^{(t+1)}} \mathbf{1}_{X_{\tau_A^+}=y} \mid X_{\tau_A^{(1)}} = y_0, \dots, X_{\tau_A^{(t+1)}} = y_t] = \mathbb{E}_{y_t}[\mathbf{1}_{X_{\tau_A^+}=y}]$$

which is the claim. □

Example 0.1. Each 30 minutes, the cat can either eat, sleep, or play outside. Each action influences the probability of the action to execute next: $p_{e,s} = 0.7, p_{e,p} = 0.3, p_{s,s} = 0.9, p_{s,e} = 0.1$ and $p_{p,e} = 1$ since the cat is tired after playing.

We take note of what the cat is doing every 30 minutes, but we forget about it if we do not see the cat (namely if it's playing outside). So our notes will only have two states $A = \{s, e\}$. The random sequence of our notes, e.g. $(s, s, s, e, e, s, \dots)$ is still a Markov chain with $p_{e,e} = 0.3, p_{e,s} = 0.7, p_{s,s} = 0.9, p_{s,e} = 0.1$.

