

# Math 525: Lecture 19

March 22, 2018

## 1 Strong Markov property

The memoryless/Markov property for a stationary Markov chain can be expressed as follows: “for each  $n$ , the process  $X_n, X_{n+1}, \dots$  is a Markov chain with the same transition probabilities as  $X_0, X_1, \dots$ .”

**Question:** is the same result true if we replace  $n$  by a stopping time  $\tau$ ?

Let  $(X_n)_{n \geq 0}$  be a Markov chain with transition matrix  $P$  and let  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ . The Markov property can be rephrased in terms of  $\mathcal{F}_n$  as follows:

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, \Lambda) = P_{ij} \quad \text{for } \Lambda \in \mathcal{F}_n.$$

Equivalently,

$$\mathbb{P}(\Lambda, X_n = i, X_{n+1} = j) = \mathbb{P}(\Lambda, X_n = i)P_{ij} \quad \text{for } \Lambda \in \mathcal{F}_n. \quad (1)$$

**Proposition 1.1** (Strong Markov property). *Let  $\tau$  be a finite stopping time (i.e.,  $\tau < \infty$  a.s.). Given that  $X_\tau = i$ , the sequence  $(X_{\tau+n})_{n \geq 0}$  is a Markov chain with transition matrix  $P$ .*

*Proof.* Let  $S$  be a finite stopping time and let  $\Lambda \in \mathcal{F}_S$ . Remember that

$$\mathcal{F}_S = \{\Lambda \in \mathcal{F} : \Lambda \cap \{S = n\} \in \mathcal{F}_n \text{ for all } n = 0, 1, 2, \dots\}.$$

Then, for states  $i$  and  $j$ ,

$$\mathbb{P}(\Lambda, X_S = i, X_{S+1} = j) = \sum_{k \geq 0} \mathbb{P}(\Lambda \cap \{S = k\}, X_k = i, X_{k+1} = j).$$

Since  $\Lambda \in \mathcal{F}_S$ , it follows that  $\Lambda \cap \{S = k\} \in \mathcal{F}_S$  for each  $k$ . By (1),

$$\mathbb{P}(\Lambda, X_S = i, X_{S+1} = j) = P_{ij} \sum_{k \geq 0} \mathbb{P}(\Lambda \cap \{S = k\}, X_k = i) = P_{ij} \mathbb{P}(\Lambda, X_S = i).$$

Divide both sides by  $\mathbb{P}(\Lambda, X_S = i)$  to get

$$\mathbb{P}(X_{S+1} = j \mid \Lambda, X_S = i) = \frac{\mathbb{P}(\Lambda, X_S = i, X_{S+1} = j)}{\mathbb{P}(\Lambda, X_S = i)} = P_{ij}. \quad (2)$$

Now, fix  $n$ . Take  $S = \tau + n$ . Since  $S$  is the sum of a stopping time and a constant,  $S$  is also a stopping time. Let  $\Lambda = \{X_\tau = i_0, \dots, X_{\tau+n} = i_n\}$ . Therefore, by (2),

$$\mathbb{P}(X_{\tau+n+1} = i_{n+1} \mid X_{\tau+n} = i_n, \dots, X_\tau = i_0) = P_{i_n i_{n+1}}.$$

That is,  $(X_{\tau+n})_{n \geq 0}$  is a Markov chain with transition matrix  $P$ . □

An immediate corollary of the above is the following:

**Corollary 1.2.** *Let  $X_0, X_1, \dots$  be i.i.d. discrete random variables (so that  $(X_n)_{n \geq 0}$  is a stationary Markov chain). Let  $\tau$  be a finite stopping time. Then,  $X_{\tau+1}$  has the same distribution as  $X_0$ .*

Before we prove this, let's try to understand the intuition.

**Example 1.3.** A gambler plays roulette and chooses a time to place a bet. Let  $X_n$  be the outcome of the  $n$ -th spin and  $\tau$  be the stopping time at which the bet is placed. The above says that  $X_{\tau+1}$  has the same distribution as  $X_0$ . In other words, assuming that the gambler eventually places a bet (i.e.,  $\tau < \infty$ ), they are no better off than they would have been had they placed the bet at time zero (i.e.,  $\tau = 0$ ).

*Proof.*  $X_0, X_1, \dots$  is a Markov chain with transition probabilities  $P_{ij} = \mathbb{P}(X_{n+1} = x_j \mid X_n = x_i) \equiv p_j$ . In other words,  $P_{ij}$  does not depend on  $i$ . By the strong Markov property,  $X_\tau, X_{\tau+1}, \dots$  is once again a Markov chain with the same transition probabilities. That is, if  $\Lambda \in \mathcal{F}_\tau$ ,

$$\mathbb{P}(X_{\tau+1} = x_j \mid \Lambda, X_\tau = x_i) = P_{ij} = p_j = \mathbb{P}(X_0 = x_j). \quad \square$$

**Exercise 1.4.**  $X_\tau$  does not necessarily have the same distribution as  $X_0$ . Why?

## 2 Recurrence and transience

If we start a Markov chain at state  $i$ , will it ever return to  $i$ ? How many times will it return to  $i$ ? These are the questions we look to answer next.

**Definition 2.1.** For a discrete random variable  $Y$ , we define its expectation conditional on an event  $\Lambda$  with  $\mathbb{P}(\Lambda) > 0$  by

$$\mathbb{E}[Y \mid \Lambda] = \frac{\mathbb{E}[Y \cdot I_\Lambda]}{\mathbb{P}(\Lambda)}.$$

Actually, the conditional expectation for a general random variable is *much* harder to define. We might come back to it later. To simplify notation, let

$$\mathbb{P}^i(\Lambda) = \mathbb{P}(\Lambda \mid X_0 = i)$$

be the probability of  $\Lambda$  conditional on the initial state of the Markov chain being  $i$ . Similarly, let

$$\mathbb{E}^i(Y) = \mathbb{E}[Y \mid X_0 = i].$$

**Definition 2.2.** The *first hitting time* of  $i$  is  $T_i = \inf\{n \geq 1 : X_n = i\}$ .

If  $X_0 \neq i$ ,  $T_i$  is the first time the chain reaches  $i$ . If  $X_0 = i$ ,  $T_i$  is the first time the chain returns to  $i$ . In the above,  $\inf \emptyset = \infty$ , so that  $T_i = \infty$  corresponds to the chain never hitting/returning to  $i$ .

**Exercise 2.3.**  $T_i$  is a stopping time.

**Exercise 2.4.** Let  $i \neq j$ . Show that  $i \rightarrow j$  if and only if  $\mathbb{P}^i\{T_j < \infty\} > 0$ .

**Definition 2.5.** A state  $i$  is *recurrent* if  $\mathbb{P}^i\{T_i < \infty\} = 1$ . It is *transient* otherwise.

We can extend the definition of  $T_i$  as follows: let  $T_i^1 = T_i$  and

$$T_i^{n+1} = \inf \{k > T_i^n : X_k = i\}.$$

That is,  $T_i^n$  is the  $n$ -th time the Markov chain visits  $i$ .

**Exercise 2.6.**  $T_i^n$  is a stopping time for each  $n \geq 1$ .

**Definition 2.7.** The *total number of returns to  $i$*  is

$$N_i = |\{n : T_i^n < \infty\}| = |\{n \geq 1 : X_n = i\}|.$$

**Proposition 2.8.** Let  $i$  be a given state and define  $p = P^i\{T_i < \infty\}$ . Then,

$$\mathbb{P}^i\{N_i \geq n\} = p^n.$$

In particular, if  $X_0 = i$  then

- $N_i = \infty$  a.s. if  $i$  is recurrent and
- $\mathbb{E}^i[N_i] = \frac{p}{1-p}$  if  $i$  is transient.

*Proof.* Let  $p^{(n)} = \mathbb{P}^i\{N_i \geq n\}$ . Note that

$$N_i \geq n \iff T_i^n < \infty.$$

Therefore,  $p^{(1)} = p$  by definition. Now, suppose  $p^{(k)} = p^k$  for  $k = 1, 2, \dots, n$ . If  $T_i^n < \infty$ , then  $X_{T_i^n} = i$ . Therefore, by the strong Markov property,  $(X_{T_i^n + n})_{n \geq 0}$  is a Markov chain with the same transition matrix as  $(X_n)_{n \geq 0}$ . Therefore,

$$\mathbb{P}^i(N_i \geq n+1) = \mathbb{P}^i(N_i \geq n)p.$$

Equivalently,  $p^{(n+1)} = p^{(n)}p = p^n p = p^{n+1}$ . This establishes that

$$p^{(n)} = p^n$$

for all  $n \geq 1$ . Now, note that

$$\mathbb{E}^i[N_i] = \sum_{k \geq 1} \mathbb{P}^i\{N_i \geq k\} = \sum_{k \geq 1} p^k = \left( \sum_{k \geq 0} p^k \right) - 1 = \frac{1}{1-p} - 1 = \frac{p}{1-p}. \quad \square$$

*Remark 2.9.*

1. If the chain starts at a recurrent state, it returns to that state infinitely often.
2. The claim

$$\mathbb{E}^i [N_i] = \frac{p}{1-p}$$

is actually true even when  $i$  is recurrent under the interpretation  $1/(1-p) = \infty$ . Therefore,

$$\begin{aligned} i \text{ is transient} &\iff \mathbb{E}^i [N_i] < \infty \\ i \text{ is recurrent} &\iff \mathbb{E}^i [N_i] = \infty. \end{aligned}$$

We can express the above just using the transition matrix  $P$ . Note that the claims below also hold for non-finite state spaces, if we interpret  $P$  to be a *denumerable transition matrix*. For our purposes, an understanding of the finite case is sufficient.

**Proposition 2.10.** *A state  $i$  is recurrent if and only if  $\sum_{n \geq 0} (P^n)_{ii} = \infty$ .*

*Proof.* This is just a consequence of  $(P^n)_{ii} = \mathbb{P}(X_n = i \mid X_0 = i)$ . Since  $N_i = \sum_{n \geq 1} I_{\{X_n = i\}}$ , it follows that

$$\mathbb{E}^i [N_i] = \mathbb{E}^i \left[ \sum_{n \geq 1} I_{\{X_n = i\}} \right] = \sum_{n \geq 1} \mathbb{E}^i [I_{\{X_n = i\}}] = \sum_{n \geq 1} (P^n)_{ii}. \quad \square$$

**Corollary 2.11.** *A state  $i$  is transient if and only if  $\sum_{n \geq 0} (P^n)_{ii} < \infty$ .*

**Corollary 2.12.** *If  $i \rightarrow j$  and  $i$  is recurrent, then  $i \leftrightarrow j$ ,  $\mathbb{P}^j \{T_i < \infty\} = 1$ , and  $j$  is recurrent.*

*Proof.* Since  $i \rightarrow j$ ,  $\mathbb{P}^i \{T_j < \infty\} > 0$ . Therefore, if  $\mathbb{P}^j \{T_i < \infty\} < 1$ , then we obtain a contradiction to  $\mathbb{E}^i [N_i] = \infty$ .

Therefore, we can find  $r$  and  $s$  such that  $(P^r)_{ij} > 0$  and  $(P^s)_{ji} > 0$  (since  $i \leftrightarrow j$ ). Then, for any  $n$ ,

$$(P^{s+n+r})_{jj} = \sum_{k,\ell} (P^s)_{jk} (P^n)_{k\ell} (P^r)_{\ell j} \geq (P^s)_{ji} (P^n)_{ii} (P^r)_{ij}.$$

Therefore,

$$\sum_n (P^n)_{jj} \geq \sum_n (P^{s+n+r})_{jj} = (P^s)_{ji} (P^r)_{ij} \sum_n (P^n)_{ii}$$

Now, since  $(P^s)_{ji} (P^r)_{ij} > 0$ , then  $j$  is recurrent since  $\sum_n (P^n)_{ii} = \infty$ . □