Lecture-20: Strong Markov Property

1 n-step transition

Definition 1.1. For a time homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, we can define **$n$-step transition probability matrix** $P^{(n)}$, with its (x, y) entry being the **n-step transition probability** for X_{m+n} to be in state y given the event $\{X_m = x\}$. That is, $p^{(n)}_{xy} \triangleq P(\{X_{n+m} = y\} \mid \{X_m = x\})$ for all $x, y \in \mathcal{X}$ and $m, n \in \mathbb{Z}_+$.

Remark 1. That is, the row $P^{(n)}_x = (p^{(n)}_{xy} : y \in \mathcal{X}) \in \mathcal{M}(\mathcal{X})$ is the conditional distribution of X_n given the initial state $\{X_0 = x\}$.

Theorem 1.2. *The n-step transition probabilities for a homogeneous Markov chain form a semi-group. That is, for all positive integers $m, n \in \mathbb{Z}_+$*

$$P^{(m+n)} = P^{(m)} P^{(n)}.$$

Proof. The events $\{\{X_m = z\} : z \in \mathcal{X}\}$ partition the sample space Ω, and hence we can express the event $\{X_{m+n} = y\}$ as the following disjoint union

$$\{X_{m+n} = y\} = \cup_{z \in \mathcal{X}} \{X_{m+n} = y, X_m = z\}.$$

It follows from the Markov property and law of total probability that for any states x, y and positive integers m, n

$$\begin{aligned} p^{(m+n)}_{xy} &= \sum_{z \in \mathcal{X}} P_x(\{X_{n+m} = y, X_m = z\}) = \sum_{z \in \mathcal{X}} P(\{X_{n+m} = y\} \mid \{X_m = z, X_0 = x\}) P_x(\{X_m = z\}) \\ &= \sum_{z \in \mathcal{X}} P(\{X_{n+m} = y\} \mid \{X_m = z\}) P_x(\{X_m = z\}) = \sum_{z \in \mathcal{X}} p^{(m)}_{xz} p^{(n)}_{zy} = (P^{(m)} P^{(n)})_{xy}. \end{aligned}$$

Since the choice of states $x, y \in \mathcal{X}$ were arbitrary, the result follows. □

Corollary 1.3. *The n-step transition probability matrix is given by $P^{(n)} = P^n$ for any positive integer n.*

Proof. In particular, we have $P^{(n+1)} = P^{(n)} P^{(1)} = P^{(1)} P^{(n)}$. Since $P^{(1)} = P$, we have $P^{(n)} = P^n$ by induction. □

Definition 1.4. For a time homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ we denote the probability mass function of Markov chain at step n by $\pi_n \in \mathcal{M}(\mathcal{X})$.

Lemma 1.5 (Chapman Kolmogorov). *The right multiplication of a probability vector with the transition matrix P transforms the probability distribution of current state to probability distribution of the next state. That is,*

$$\pi_{n+1} = \pi_n P, \text{ for all } n \in \mathbb{N}.$$

Proof. To see this, we fix $y \in \mathcal{X}$ and from the law of total probability and the definition conditional probability, we observe that

$$\pi_{n+1}(y) = P\{X_{n+1} = y\} = \sum_{x \in \mathcal{X}} P\{X_{n+1} = y, X_n = x\} = \sum_{x \in \mathcal{X}} P\{X_n = x\} p_{xy} = (\pi_n P)_y.$$

□

2 Strong Markov property (SMP)

We are interested in generalizing the Markov property to any random times. For a DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ and a random variable $\tau : \Omega \to \mathbb{N}$, we are interested in knowing whether for any historical event $H_{\tau-1} = \cap_{n=0}^{\tau-1} \{X_n = x_n\}$ and any state $x, y \in \mathcal{X}$, we have

$$P(\{X_{\tau+1} = y\} \mid H_{\tau-1} \cap \{X_\tau = x\}) = p_{xy}.$$

Example 2.1 (Two-state DTMC). Consider the two state Markov chain $X \in \{0,1\}^{\mathbb{Z}_+}$ such that $P_0\{X_1 = 1\} = q$ and $P_1\{X_1 = 0\} = p$ for $p, q \in [0,1]$. Let $\tau : \Omega \to \mathbb{N}$ be a random variable defined as

$$\tau \triangleq \sup\{n \in \mathbb{N} : X_i = 0, \text{ for all } i \leqslant n\}.$$

That is, $\{\tau = n\} = \{X_1 = 0, \dots, X_n = 0, X_{n+1} = 1\}$. Hence, for the historical event $H_{\tau-1} = \{X_1 = \dots, X_{\tau-1} = 0\}$, the conditional probability $P(\{X_{\tau+1} = 1\} \mid H_{\tau-1} \cap \{X_\tau = 0\}) = 1$, and not equal to q.

Definition 2.2. Let $\tau : \Omega \to \mathbb{N}$ be a stopping time with respect to a random sequence $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$. Then for all states $x, y \in \mathcal{X}$ and the event $H_{\tau-1} = \cap_{n=0}^{\tau-1} \{X_n = x_n\}$, the process X satisfies the **strong Markov property** if

$$P(\{X_{\tau+1} = y\} \mid \{X_\tau = x\} \cap H_{\tau-1}) = P(\{X_{\tau+1} = y\} \mid \{X_\tau = x\}).$$

Lemma 2.3. *Homogeneous Markov chains satisfy the strong Markov property.*

Proof. Let $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ be a homogeneous DTMC with transition matrix P, and $\tau : \Omega \to \mathbb{N}$ be an associated stopping time. We take any historical event $H_{\tau-1} = \cap_{n=0}^{T-1} \{X_n = x_n\}$, and states $x, y \in \mathcal{X}$. From the definition of conditional probability, the law of total probability, and the Markovity of the process X, we have

$$\begin{aligned}
&P(\{X_{\tau+1} = y\} \mid H_{\tau-1} \cap \{X_\tau = x\}) = \frac{\sum_{n\in\mathbb{Z}_+} P(\{X_{\tau+1} = y, X_\tau = x\} \cap H_{\tau-1} \cap \{T = n\})}{P(\{X_\tau = x\} \cap H_{\tau-1})} \\
&= \sum_{n\in\mathbb{Z}_+} P(\{X_{n+1} = y\} \mid \{X_n = x\} \cap H_{n-1} \cap \{\tau = n\}) P(\{\tau = n\} \mid \{X_T = x\} \cap H_{\tau-1}) \\
&= p_{xy} \sum_{n\in\mathbb{Z}_+} P(\{\tau = n\} \mid \{X_T = x\} \cap H_{\tau-1}) = p_{xy}.
\end{aligned}$$

This equality follows from the fact that the event $\{\tau = n\}$ is completely determined by $(X_0, \dots, X_n)$. ☐

Remark 2. Consider a homogeneous DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ and the first instant $\tau_k \triangleq \tau_X^{\{y\},k}$ for the process X to hit k times, a state $y \in \mathcal{X}$. Recall that $\tau_0 \triangleq 0$ and recurrence time $H_k \triangleq \tau_k - \tau_{k-1} = \inf\{n \in \mathbb{N} : X_{\tau_{k-1}+n} = y\}$ for all $k \in \mathbb{N}$. We define a process $Y : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ where $Y_m \triangleq X_{\tau_k+m}$ for all $m \in \mathbb{Z}_+$. If τ_k is almost surely finite, then it is a stopping time with respect to process X. Using strong Markov property of DTMC X, we will show that Y is a stochastic replica of X with $X_0 = y$.

3 Hitting and Recurrence Times

We will consider a time-homogeneous discrete time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ on countable state space $\mathcal{X}$ with transition probability matrix $P : \mathcal{X} \times \mathcal{X} \to [0,1]$, and initial state $X_0 = x \in \mathcal{X}$. We denote the natural filtration generated by the process X as $\mathcal{F}_\bullet$, where $\mathcal{F}_n \triangleq \sigma(X_0, \dots, X_n)$ for all $n \in \mathbb{N}$.

Remark 3. Starting from state x, the mean number of visits to state y in n steps is $\mathbb{E}_x N_y(n) = \sum_{k=1}^n p_{xy}^{(k)}$. From the monotone convergence theorem, we also get that $E_x N_y(\infty) = \sum_{k\in\mathbb{N}} p_{xy}^{(k)}$.

Remark 4. If τ_{k-1} is almost sure finite, then τ_{k-1} is a stopping time for process X. From the strong Markov property of homogeneous DTMC X applied to stopping time τ_{k-1}, it follows that the future $\sigma(X_{\tau_{k-1}+j} : j \in \mathbb{N})$ is independent of the past $\sigma(X_0, \dots, X_{\tau_{k-1}})$ given the present $\sigma(X_{\tau_{k-1}})$. Since $X_{\tau_{k-1}} = y$ for $k \geqslant 2$ deterministically, it follows that $\sigma(X_{\tau_{k-1}})$ is a trivial event space and the future $\sigma(X_{\tau_{k-1}+j} : j \in \mathbb{N})$ is independent of the random past $\sigma(X_0, \dots, X_{\tau_{k-1}})$. We further observe that the distribution of $\sigma(X_{\tau_{k-1}+j} : j \in \mathbb{N})$ is identical to distribution of X given $X_0 = y$. Thus, the process $(X_{\tau_{k-1}+j} : j \in \mathbb{N})$ is distributed identically for all $k \geqslant 2$.

Remark 5. We observe that the recurrence time satisfies $\{H_k = n\} \in \sigma(X_{\tau_{k-1}+j} : j \in [n])$ for all $n \in \mathbb{N}$, and hence the recurrence time H_k is independent of the random past $\sigma(X_0, \dots, X_{\tau_{k-1}})$. Recursively applying this fact, we can conclude that $(H_1, \dots, H_k)$ are independent random variables. Further, since $(X_{\tau_{k-1}+j} : j \in \mathbb{N})$ is distributed identically for all $k \geqslant 2$, it follows that $(H_k : k \geqslant 2)$ are distributed identically.

Lemma 3.1. *If H_1 and H_2 are almost surely finite, then the random sequence $(H_k : k \geqslant 2)$ is* i.i.d. .

Proof. From the above two remarks, it suffices to show that each term of the random sequence $\tau : \Omega \to \mathbb{N}^{\mathbb{N}}$ is almost surely finite. We will show this by induction. Since $\tau_1 = H_1$ is almost surely finite, it follows that τ_1 is stopping time. Since $\tau_2 = \tau_1 + H_2$ is almost surely finite, it follows that τ_2 is a stopping time. By inductive hypothesis τ_{k-1} is almost surely finite, and hence H_k is independent of $(H_1, \dots, H_k)$ and identically distributed to H_2 and is almost surely finite. It follows that $\tau_k = \tau_{k-1} + H_k$ is almost surely finite, and the result follows. $\square$