Lecture-20: Strong Markov Property

1 Random mapping theorem

We saw some example of Markov processes where $X_n = X_{n-1} + Z_n$, and $(Z_n : n \in \mathbb{N})$ is an iid sequence, independent of the initial state X_0. We will show that any discrete time Markov chain is of this form, where the sum is replaced by arbitrary functions.

Theorem 1.1 (Random mapping theorem). *For any DTMC X, there exists an* i.i.d. *sequence $Z \in \Lambda^{\mathbb{N}}$ and a function $f : \mathcal{X} \times \Lambda \to \mathcal{X}$ such that $X_n = f(X_{n-1}, Z_n)$ for all $n \in \mathbb{N}$.*

Remark 1. A **random mapping representation** of a transition matrix P on state space $\mathcal{X}$ is a function $f : \mathcal{X} \times \Lambda \to \mathcal{X}$, along with a Λ-valued random variable Y, satisfying

$$P\{f(x,Y) = y\} = p_{xy}, \text{ for all } x,y \in \mathcal{X}.$$

Proof. It suffices to show that every transition matrix P has a random mapping representation. Then for the mapping f and the *i.i.d* sequence $Z = (Z_n : n \in \mathbb{N})$ with the same distribution as random variable Y, we would have $X_n = f(X_{n-1}, Z_n)$ for all $n \in \mathbb{N}$.

Let $\Lambda = [0,1]$, and we choose the *i.i.d.* sequence Z, uniformly at random from this interval. Since $\mathcal{X}$ is countable, it can be ordered. We let $\mathcal{X} = \mathbb{N}$ without any loss of generality. We set $F_{xy} \triangleq \sum_{w \leqslant y} p_{xw}$ and define

$$f(x,z) = \sum_{y \in \mathbb{N}} y \mathbb{1}_{\left\{F_{x,y-1} < z \leqslant F_{x,y}\right\}} = \inf\left\{y \in \mathcal{X} : z \leqslant F_{x,y}\right\}.$$

Since $f(x, Z_n)$ is a discrete random variable taking value $y \in \mathcal{X}$, iff the uniform random variable Z_n lies in the interval $(F_{x,y-1}, F_{x,y}]$. That is, the event $\{f(x,Z_n) = y\} = \left\{Z_n \in (F_{x,y-1}, F_{x,y}]\right\}$ for all $y \in \mathcal{X}$. It follows that

$$P\{f(x,Z) = y\} = P\left\{F_{x,y-1} < Z \leqslant F_{x,y}\right\} = F_{x,y} - F_{x,y-1} = p_{xy}.$$

□

2 Strong Markov property (SMP)

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

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

Example 2.1 (Two-state DTMC). For 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 $T : \Omega \to \mathbb{N}$ be an integer random variable defined as

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

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

Definition 2.2. Let T be an integer valued stopping time with respect to a random sequence X. Then for all states $x, y \in \mathcal{X}$ and the event $H_{T-1} = \cap_{n=0}^{T-1} \{X_n = x_n\}$, the process X satisfies the **strong Markov property** if

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

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

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

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

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

Example 2.4 (For a non stopping time T). As an exercise, if we try to use the Markov property on arbitrary random variable T, the SMP may not hold. For example, define a non-stopping time $T \triangleq \inf\{n \in \mathbb{Z}_+ : X_{n+1} = y\}$ for $y \in \mathcal{X}$. In this case, we have

$$P(\{X_{T+1} = y\} \mid \{X_T = x, \dots, X_0 = x_0\}) = \mathbb{1}_{\{p_{xy}>0\}} \neq P(\{X_1 = y\} \mid \{X_0 = x\}) = p_{xy}.$$

Remark 2. A useful application of the strong Markov property is as follows. Let $x_0 \in \mathcal{X}$ be a fixed state and $\tau_0 = 0$. Let τ_n denote the stopping times at which the Markov chain visits x_0 for the nth time. That is,

$$\tau_n \triangleq \inf\{n > \tau_{n-1} : X_n = x_0\}.$$

Then $(X_{\tau_n+m} \in \mathcal{X}^\Omega : m \in \mathbb{Z}_+)$ is a stochastic replica of $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with $X_0 = x_0$.

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 $\tau_X^{\{y\},k-1}$ is almost sure finite, then $\tau_X^{\{y\},k-1}$ is a stopping time for process X. Therefore, from the strong Markov property for X and the fact that $\left\{H_X^{\{y\},k} = n\right\} \in \sigma(X_{\tau_X^{\{y\},k-1}+j} : j \in [n])$ for all $n \in \mathbb{N}$, we observe that $H_X^{\{y\},k}$ given $X_{\tau_X^{\{y\},k-1}}$ is independent of the random past $\sigma(X_0, \dots, X_{\tau_X^{\{y\},k-1}})$. Since $X_{\tau_X^{\{y\},k-1}} = y$ deterministically, it follows that $H_X^{\{y\},k}$ is independent of the random past $\sigma(X_0, \dots, X_{\tau_X^{\{y\},k-1}})$. It follows that $(H_X^{\{y\},1}, \dots, H_X^{\{y\},k})$ are independent random variables.

Remark 5. If $\tau_X^{\{y\},k-1}$ is almost sure finite, then from strong Markov property of X, we observe that $(X_{\tau_X^{\{y\},k-1}+j} : j \in \mathbb{N})$ is distributed identically to $(X_{\tau_X^{\{y\},\ell}+j} : j \in \mathbb{N})$. That is, $(H_X^{\{y\},k} : k \geqslant 2)$ are distributed identically.

Lemma 3.1. *If $H_X^{\{y\},1}$ and $H_X^{\{y\},2}$ are almost surely finite, then the random sequence $(H_X^{\{y\},k} : k \geqslant 2)$ is* i.i.d. .

Proof. From above two remarks, it suffices to show that $(\tau_X^{\{y\},k} : k \in \mathbb{N})$ are almost surely finite. We will show this by induction. Since $\tau_X^{\{y\},\ell} = H_X^{\{y\},1}$ is almost surely finite, $\tau_X^{\{y\},\ell}$ is stopping time. Since $\tau_X^{\{y\},2} = \tau_X^{\{y\},\ell} + H_X^{\{y\},2}$ is almost surely finite, it follows that $\tau_X^{\{y\},2}$ is a stopping time. By inductive hypothesis $\tau_X^{\{y\},k-1}$ is almost surely finite, and hence $H_X^{\{y\},k}$ is independent of $(H_X^{\{y\},1}, \ldots, H_X^{\{y\},k})$ and identically distributed to $H_X^{\{y\},2}$ and is almost surely finite. It follows that $\tau_X^{\{y\},k} = \tau_X^{\{y\},k-1} + H_X^{\{y\},k}$ is almost surely finite, and the result follows. □