Lecture-21: DTMC: Representation

1 n-step transition

Consider a homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with countable state space $\mathcal{X}$ and transition matrix P.

Definition 1.1. We would respectively denote the conditional probability of events and conditional expectation of random variables, conditioned on the event $X_0 = x$, by

$$P_x(A) = P(A \mid \{X_0 = x\}), \qquad \mathbb{E}_x[Y] = \mathbb{E}\left[A \mid \{X_0 = x\}\right].$$

Proposition 1.2. *Conditioned on the initial state, any finite dimensional distribution of a homogeneous Markov chain is stationary.*

Proof. To this end, we compute the transition probabilities for the path $(x_1, \dots, x_n)$ taken by the sample path $(X_1, \dots, X_n)$ when $X_0 = x_0$ and by the sample path $(X_{m+1}, \dots, X_{m+n})$ when $X_m = x_0$. For each $i \in \{0, \dots, n\}$, we can define events

$$H_i \triangleq \cap_{j=1}^{i} \{X_j = x_j\}, \qquad \tilde{H}_i \triangleq \cap_{j=0}^{i} \{X_j = x_j\} = h_i \cap \{X_0 = x_0\}.$$

We observe that $H_i = \{X_i = x_i\} \cap H_{i-1}$ and $H_i, \tilde{H}_i \in \mathcal{F}_i = \sigma(X_0, \dots, X_i)$ for all $i \in \mathbb{N}$. From the definition of $H_{n-1}, \tilde{H}_{n-1}$, and the conditional probability, we can write

$$P_{x_0}(H_n) = P_{x_0}(\{X_n = x_n\} \cap H_{n-1}) = P(\{X_n = x_n\} \mid \tilde{H}_{n-1}) P_{x_0}(H_{n-1}).$$

Using the fact that $\tilde{H}_{n-1} = \{X_{n-1} = x_{n-1}\} \cap \tilde{H}_{n-2}$, and the Markovity and homogeneity of the process X, we obtain

$$P(\{X_n = x_n\} \mid \tilde{H}_{n-1}) = P(\{X_n = x_n\} \mid \{X_{n-1} = x_{n-1}\} \cap \tilde{H}_{n-2}) = p_{x_{n-1}x_n}.$$

Inductively, we can write the conditional joint distribution of H_n given the event $\{X_0 = x_0\}$ as

$$P_{x_0}(H_n) = p_{x_0 x_1} \dots p_{x_{n-1}x_n}.$$

Similarly, we can write for the sample path $(X_{m+1}, \dots, X_{m+n})$ given $X_m = x_0$,

$$P(\{X_{m+1} = x_1, \dots, X_{m+n} = x_n\} \mid \{X_m = x_0\}) = \prod_{i=1}^{n} P(\{X_{m+i} = x_i)\} \mid \{X_{m+i-1} = x_{i-1}\}) = p_{x_0x_1} p_{x_1x_2} \dots p_{x_{n-1}x_n}.$$

□

Corollary 1.3. *The n-step transition probabilities are stationary for any homogeneous Markov chain. That is, for any states $x, y \in \mathcal{X}$ and $n, m \in \mathbb{N}$, we have*

$$P(\{X_{n+m} = y\} | \{X_m = x\}) = P(\{X_n = y\} | \{X_0 = x\}).$$

Proof. It follows from summing over intermediate steps. In particular, we can partition the outcome space Ω in terms of disjoint events $\left\{E_{n-1}(x_0, \dots, x_{n-1}) \triangleq \cap_{i=1}^{n-1} \{X_i = x_i\} : x_1, \dots, x_{n-1} \in \mathcal{X}\right\}$. Then, we can write

$$\{X_n = x_n\} = \cup_{x_1, \dots, x_{n-1} \in \mathcal{X}} \{X_n = x_n\} \cap E_{n-1}(x_0, \dots, x_{n-1}).$$

Using the law of total probability, we can write the conditional probability

$$P_{x_0}\{X_n = x_n\} = \sum_{x_1, \dots, x_{n-1} \in \mathcal{X}} P_{x_0}(\{X_n = x_n\} \cap E_{n-1}(x_0, \dots, x_{n-1})).$$

Similarly, we can write using the law of total probability for partition $\left\{F_{n-1}(x_1,\dots,x_{n-1}) \triangleq \cap_{i=1}^{n-1}\{X_{m+i}=x_i\} : x_1,\dots,x_{n-1} \in\right.$ we can write

$$P(\{X_{m+n}=x_n\} \mid \{X_m=x_0\}) = \sum_{x_1,\dots,x_{n-1}\in\mathcal{X}} P(\{X_{m+n}=x_n\}\cap F_{n-1}(x_0,\dots,x_{n-1}) \mid \{X_m=x_0\}).$$

From the stationarity in joint distribution conditioned on initial state for the homogeneous Markov chain X, we have

$$P(\{X_{m+n}=x_n\}\cap F_{n-1}(x_0,\dots,x_{n-1}) \mid \{X_m=x_0\}) = P(\{X_n=x_n\}\cap E_{n-1}(x_0,\dots,x_{n-1}) \mid \{X_0=x_0\}).$$

The result follows. □

Hence, it follows that for a homogeneous Markov chain, we can define n-step transition probabilities for $x,y\in\mathcal{X}$ and $m,n\in\mathbb{N}$

$$p_{xy}^{(n)} \triangleq P(\{X_{n+m}=y\}|\{X_m=x\}).$$

That is, the row $P_x^{(n)} = (p_{xy}^{(n)} : y\in\mathcal{X})$ is the conditional distribution of X_n given $X_0=x$.

Theorem 1.4. *The n-step transition probabilities form a semi-group. That is, for all positive integers m, n*

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

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

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

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

Corollary 1.5. *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. □

Remark 1. That is, for all states x,y and non-negative integers $n\in\mathbb{Z}_+$, $p_{xy}^{(n)}=P_{xy}^n$.

2 Representation

2.1 Chapman Kolmogorov equations

We denote by $\pi_0\in\mathbb{R}_+^{\mathcal{X}}$ the initial distribution of the Markov chain, that is $\pi_0(x)=P\{X_0=x\}$. The distribution of X_n is given by $\pi_n\in\mathbb{R}_+^{\mathcal{X}}$, such that for any state $x\in\mathcal{X}$.

$$\pi_n(x) = P\{X_n=x\} = \sum_{z\in\mathcal{X}} p_{zx}^{(n)}\pi_0(z) = (\pi_0P^n)_x.$$

We can write this succinctly in terms of transition probability matrix P as $\mu_n=\mu_0P^n$. We can alternatively derive this result by the following Lemma.

Lemma 2.1. *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,*

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

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

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

□

2.2 Transition graph

We can define a collection E of possible one-step transitions indicated by the initial and the final state, as

$$E \triangleq \{[x,y\rangle \in \mathcal{X} \times \mathcal{X} : p_{xy} > 0\}.$$

A transition matrix P is sometimes represented by a directed weighted graph $G = (\mathcal{X}, E, W)$, where the set of nodes in the graph G is the state space $\mathcal{X}$, and the set of directed edges is the set of possible transitions. In addition, this graph has a weight $w_e = p_{xy}$ on each edge $e = [x,y\rangle \in E$.

Example 2.2 (Integer random walk). For an integer random walk $X = (X_n \in \mathbb{Z} : n \in \mathbb{N})$ with *i.i.d.* step-size sequence $Z = (Z_n \in \{-1,1\}, n \in \mathbb{N})$, we have and infinite graph $G = (\mathbb{Z}, E)$, where the edge set is

$$E = \{(n, n+1) : n \in \mathbb{Z}\} \cup \{(n, n-1) : n \in \mathbb{Z}\}.$$

We have plotted the sub-graph of the entire transition graph for states $\{-1,0,1\}$ in Figure 1.

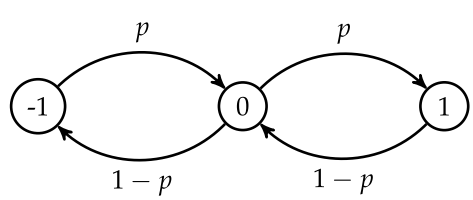

Figure 1: Sub-graph of the entire transition graph for an integer random walk with *i.i.d.* step-sizes in $\{-1,1\}$ with probability p for the positive step.

Example 2.3 (Sequence of experiments). Consider the sequence of experiments with the set of outcomes $\mathcal{X} = \{0,1\}$ with the transition matrix

$$P = \begin{bmatrix} 1-q & q \\ p & 1-p \end{bmatrix}.$$

We have plotted the corresponding transition graph for this two-state Markov chain in Figure 2.

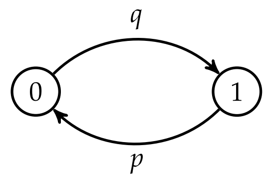

Figure 2: Markov chain for the sequence of experiments with two outcomes.

2.3 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 2.4 (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 2. 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

$$f(x,z) = \sum_{y \in \mathbb{N}} y 1_{\{F_{x,y-1} < z \leqslant F_{x,y}\}}.$$

It follows that $P\{f(x,Z) = y\} = P\{F_{x,y-1} < Z \leqslant F_{x,y}\} = p_{xy}$. □