Lecture-21: DTMC: Representation

1 Representation

Consider a Markov chain $X = (X_n \in \mathcal{X} : n \in \mathbb{Z}_+)$ with countable state space $\mathcal{X}$ and transition matrix P.

1.1 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 $x \in \mathcal{X}$.

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

We can write this succinctly in terms of transition probability matrix P as $\mu_n = \mu_0 P^n$. We can alternatively derive this result by the following Lemma.

Lemma 1.1. *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,*

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

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

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

□

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

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

Lemma 1.2. *The left multiplication of any vector of a function evaluation at all states, with the transition matrix P results in a vector of the expected value of the function at the next state given the current state. That is, for any function $f : \mathcal{X} \to \mathbb{R}$, we have $(Pf)_x = \mathbb{E}_x[f(X_1)]$.*

Proof. From the definition of transition probability matrix, we have

$$(Pf)_x = \sum_{y \in \mathcal{X}} p_{xy} f(y) = \sum_{y \in \mathcal{X}} f(y) \mathbb{E}_x 1_{\{X_1 = y\}} = \mathbb{E}_x[f(X_1)].$$

□

1.2 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 1.3 (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.

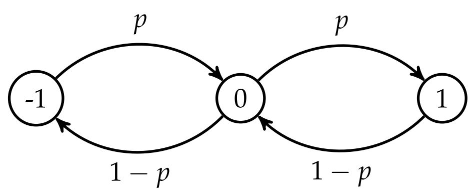

Figure 1: Sub-graph of the entire transition graph for an integer random walk with i.i.d. step-sizes in $\{-1,1\}$ with probability p for the positive step.

Example 1.4 (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.

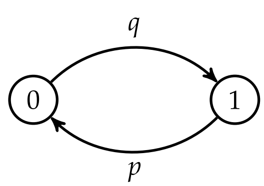

Figure 2: Markov chain for the sequence of experiments with two outcomes.

1.3 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 1.5 (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

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

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