Lecture-05: Reversibility

1 Introduction

Definition 1.1. A stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{R}}$ is **time reversible** if the vector $(X_{t_1}, \dots, X_{t_n})$ has the same distribution as $(X_{\tau - t_1}, \dots, X_{\tau - t_n})$ for all finite positive integers n, time instants $t_1 < t_2 < \dots < t_n$ and shifts $\tau \in \mathbb{R}$.

Lemma 1.2. *A time reversible process is stationary.*

Proof. It suffices to show that for any shift $s \in \mathbb{R}$, a finite $n \in \mathbb{N}$, time instants $t_1 < \dots < t_n$, and states $x_1, \dots, x_n \in \mathcal{X}$, we have

$$P\Big(\cap_{i \in [n]} \{X_{t_i} = x_i\} \Big) = P\Big(\cap_{i \in [n]} \{X_{s+t_i} = x_i\} \Big).$$

This follows from time reversibility of X, since both $(X_{t_1}, \dots, X_{t_n})$ and $(X_{s+t_1}, \dots, X_{s+t_n})$ have the same distribution as $(X_{-t_1}, \dots, X_{-t_n})$, by taking $\tau = 0$ and $\tau = -s$ respectively. ☐

Theorem 1.3. *A stationary Markov process $X : \Omega \to \mathcal{X}^{\mathbb{R}}$ with countable state space $\mathcal{X} \subseteq \mathbb{R}$ and probability transition kernel $P : \mathbb{R}_+ \to \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ is time reversible iff there exists a probability distribution $\pi \in \mathcal{M}(\mathcal{X})$, that satisfy the detailed balanced conditions*

$$\pi_x P_{xy}(t) = \pi_y P_{yx}(t) \text{ for all } x, y \in \mathcal{X} \text{ and times } t \in \mathbb{R}_+. \tag{1}$$

When such a distribution π exists, it is the invariant distribution of the process.

Proof. We assume that the process X is time reversible, and hence stationary. We denote the stationary distribution by π, and by time reversibility of X, we have

$$P_\pi \{X_{t_1} = x, X_{t_1+t} = y\} = P_\pi \{X_{t_1} = y, X_{t_1+t} = x\},$$

for $\tau = 2t_1 + t$. Hence, we obtain the detailed balanced conditions in Eq. (1).

Conversely, let π be the distribution that satisfies the detailed balanced conditions in Eq. (1), then summing up both sides over $y \in \mathcal{X}$, we see that π is the invariant distribution for X. Let $x \in \mathcal{X}^m$, then applying detailed balanced equations in Eq. (1) repeatedly, we can write

$$\pi(x_1) P_{x_1 x_2}(t_2 - t_1) \dots P_{x_{m-1} x_m}(t_m - t_{m-1}) = \pi(x_m) P_{x_m x_{m-1}}(t_m - t_{m-1}) \dots P_{x_2 x_1}(t_2 - t_1).$$

For the time homogeneous stationary Markov process X, it follows that for all $t_0 \in \mathbb{R}_+$

$$P_\pi \{X_{t_1} = x_1, \dots, X_{t_m} = x_m\} = P_\pi \{X_{t_0} = x_m, \dots, X_{t_0 + t_m - t_1} = x_1\}.$$

Since $m \in \mathbb{N}$ and $t_0, t_1 \dots, t_m$ were arbitrary, the time reversibility follows. ☐

1.1 Reversible Chains

Corollary 1.4. *A stationary homogeneous discrete time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}}$ with transition matrix $P \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ is time reversible iff there exists a probability distribution $\pi \in \mathcal{M}(\mathcal{X})$, that satisfies the detailed balanced conditions*

$$\pi_x P_{xy} = \pi_y P_{yx}, \quad x, y \in \mathcal{X}. \tag{2}$$

When such a distribution π exists, it is the invariant distribution of the process.

Example 1.5 (Random walks on edge-weighted graphs). Consider an undirected graph $G = (\mathcal{X}, E)$ with the vertex set $\mathcal{X}$ and the edge set $E = \{\{x,y\} : x, y \in \mathcal{X}\}$ being a subset of unordered pairs of elements from $\mathcal{X}$. We say that y is a neighbor of x (and x is a neighbor of y), if $e = \{x,y\} \in E$ and denote $x \sim y$. We assume a function $w : E \to \mathbb{R}_+$, such that w_e is a positive number associated with each edge $e = \{x,y\} \in E$. Let $X_n \in \mathcal{X}$ denote the location of a particle on one of the graph vertices at the nth time-step. Consider the following random discrete time movement of a particle on this graph from one vertex to another. If the particle is currently at vertex x then it will next move to vertex y with probability

$$P^G_{xy} \triangleq P(\{X_{n+1} = y\} \mid \{X_n = x\}) = \frac{w_e}{\sum_{f:x\in f} w_f} 1_{\{e=\{x,y\}\}}.$$

The Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ describing the sequence of vertices visited by the particle is a random walk on an undirected edge-weighted graph. Google's PageRank algorithm, to estimate the relative importance of webpages, is essentially a random walk on a graph!

Proposition 1.6. *Consider an irreducible homogeneous Markov chain that describes the random walk on an edge weighted graph with a finite number of vertices. In steady state, this Markov chain is time reversible with stationary probability of being in a state $x \in \mathcal{X}$ given by*

$$\pi_x = \frac{\sum_{f:x\in f} w_f}{2\sum_{g\in E} w_g}. \tag{3}$$

Proof. Using the definition of transition probabilities for this Markov chain and the given distribution π defined in (3), we notice that

$$\pi_x P^G_{xy} = \frac{w_e}{\sum_{f\in E} w_f} \mathbb{1}_{\{e=\{x,y\}\}}, \qquad \pi_y P^G_{yx} = \frac{w_e}{\sum_{f\in E} w_f} \mathbb{1}_{\{e=\{x,y\}\}}.$$

Hence, the detailed balance equation for each pair of states $x, y \in \mathcal{X}$ is satisfied, and the result follows. □

We can also show the following *dual* result.

Lemma 1.7. *Let $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ be a time reversible Markov chain on a finite state space $\mathcal{X}$ and transition probability matrix $P \in [0,1]^{\mathcal{X}\times\mathcal{X}}$. Then, there exists a random walk on a weighted, undirected graph G with the same transition probability matrix P.*

Proof. We create a graph $G = (\mathcal{X}, E)$, where $E = \{\{x,y\} : x, y \in \mathcal{X}, P_{xy} > 0\}$. For the stationary distribution $\pi : \mathcal{X} \to [0,1]$ for the Markov chain X, we set the edge weights

$$w_{\{x,y\}} \triangleq \pi_x P_{xy} = \pi_y P_{yx},$$

With this choice of weights, it is easy to check that $w_x = \sum_{f:x\in f} w_f = \pi_x$, and the transition matrix associated with a random walk on this graph is exactly P with $P^G_{xy} = \frac{w_{\{x,y\}}}{w_x} = P_{xy}$. □

Is every Markov chain time reversible?

1. If the process is not stationary, then no. To see this, we observe that

   $$P\{X_{t_1} = x_1, X_{t_2} = x_2\} = \nu_{t_1}(x_1) P_{x_1x_2}(t_2 - t_1), \quad P\{X_{\tau-t_2} = x_2, X_{\tau-t_1} = x_1\} = \nu_{\tau-t_2}(x_2) P_{x_2x_1}(t_2 - t_1).$$

   If the process is not stationary, the two probabilities can't be equal for all times τ, t_1, t_2 and states $x_1, x_2 \in \mathcal{X}$.

2. If the process is stationary, then it is still not true in general. Suppose we want to find a stationary distribution $\alpha \in \mathcal{M}(\mathcal{X})$ that satisfies the detailed balance equations $\alpha_x P_{xy} = \alpha_y P_{yx}$ for all states $x, y \in \mathcal{X}$. For any arbitrary Markov chain X, one may not end up getting any solution. To see this consider a state $z \in \mathcal{X}$ such that $P_{xy} P_{yz} > 0$. Time reversibility condition implies that $P_\alpha\{X_1 = x, X_2 = y, X_3 = z\} = P_\alpha\{X_1 = z, X_2 = y, X_3 = z\}$, and hence

   $$\frac{\alpha_x}{\alpha_z} = \frac{P_{zy} P_{yx}}{P_{xy} P_{yz}} \neq \frac{P_{zx}}{P_{xz}}.$$

Thus, we see that a necessary condition for time reversibility is $P_{xy}P_{yz}P_{zx} = P_{xz}P_{zy}P_{yx}$ for all $x, y, z \in \mathcal{X}$.

Theorem 1.8 (Kolmogorov's criterion for time reversibility of Markov chains). *A stationary Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}}$ is time reversible if and only if starting in state $x_0 \in \mathcal{X}$, any path back to state x_0 has the same probability as the time reversed path, for all initial states $x_0 \in \mathcal{X}$. That is, for any $n \in \mathbb{N}$ and state vector $x \in \mathcal{X}^n$*

$$P_{x_0x_1}P_{x_1x_2}\dots P_{x_nx_0} = P_{x_0x_n}P_{x_nx_{n-1}}\dots P_{x_1x_0}. \tag{4}$$

Proof. The detailed balance equation for a time reversible Markov process X implies that (4) holds for any finite set of states. Conversely, if (4) holds for any non-negative integer $n \in \mathbb{N}$, then for any states $x_0, y \in \mathcal{X}$, we have

$$(P^{n+1})_{x_0y}P_{yx_0} = \sum_{x_1,x_2,\dots x_n} P_{x_0x_1}\dots P_{x_ny}P_{yx_0} = \sum_{x_1,x_2,\dots x_n} P_{x_0y}P_{yx_n}\dots P_{x_1x_0} = P_{x_0y}(P^{n+1})_{yx_0}.$$

Taking the limit $n \to \infty$ and noticing that $\lim_{n\to\infty}(P^n)_{xy} = \pi_y$ for all $x, y \in \mathcal{X}$, we observe that X is a time-reversible process. □

1.2 Reversible Processes

Corollary 1.9. *A stationary Markov process $X : \Omega \to \mathcal{X}^{\mathbb{R}}$ with generator matrix $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ is time reversible iff there exists a probability distribution $\pi \in \mathcal{M}(\mathcal{X})$, that satisfies the detailed balanced conditions*

$$\pi_x Q_{xy} = \pi_y Q_{yx}, \quad x, y \in \mathcal{X}. \tag{5}$$

When such a distribution π exists, it is the invariant distribution of the process.

Definition 1.10. Let $X : \Omega \to \mathcal{X}^{\mathbb{R}}$ be a stationary homogeneous Markov process with stationary distribution $\pi \in \mathcal{M}(\mathcal{X})$ and the generator matrix $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$. The **probability flux** from state x to state y is defined as $\lim_{t\to\infty} \frac{N_t^{xy}}{t}$, where $N_t^{xy} \triangleq \sum_{n\in\mathbb{N}} \mathbb{1}_{\{S_n \leqslant t, Z_n = y, Z_{n-1} = x\}}$ denotes the total number of transitions from state x to state y in the time duration $(0, t]$.

Lemma 1.11. *For a time-homogeneous CTMC X, the probability flux from state x to state y is $\pi_x Q_{xy}$.*

Proof. Let $X_0 = x$ and $\tau_x^+(k)$ be the kth visiting time to state x. It follows that $\tau_x^+ : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ is a renewal sequence. We consider the reward process $(N_t^{xy}, t \geqslant 0)$ where N_t^{xy} is the number of transitions from state x to y in duration $(0, t]$. We denote the total number of transitions from state x to state y in the kth inter-renewal duration by

$$N^{xy}(k) \triangleq \sum_{n\in\mathbb{N}} \mathbb{1}_{\left\{\tau_x^+(k-1) < S_n \leqslant \tau_x^+(k)\right\}} \mathbb{1}_{\{Z_{n-1}=x, Z_n=y\}}.$$

From the renewal reward theorem for the embedded DTMC $Z : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, we can write the average number of transitions from x to y as

$$u_x p_{xy} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{\{Z_{n-1}=x, Z_n=y\}} = \frac{\mathbb{E}_x N^{xy}(k)}{\mathbb{E}_x N_x(k)} = u_x \mathbb{E}_x N^{xy}(k).$$

Then, we observe that $\mathbb{E}_x N^{xy}(k) = p_{xy}$ and recall that $\mathbb{E}_x \tau_x^+(1) = \frac{1}{\pi_x \nu_x}$. From the renewal reward process, we obtain

$$\lim_{t\to\infty} \frac{N_t^{xy}}{t} = \frac{\mathbb{E}_x N^{xy}(1)}{\mathbb{E}_x \tau_x^+(1)} = \pi_x \nu_x p_{xy} = \pi_x Q_{xy}.$$

□

Lemma 1.12. *For a stationary homogeneous Markov process $X : \Omega \to \mathcal{X}^{\mathbb{R}}$, probability flux balances across a cut $A \subseteq \mathcal{X}$, that is*

$$\sum_{y\notin A}\sum_{x\in A} \pi_x Q_{xy} = \sum_{x\in A}\sum_{y\notin A} \pi_y Q_{yx}.$$

Proof. From the row sum of generator matrix being zero, we get $\sum_{y\in\mathcal{X}} Q_{xy} = 0$ for all $x \in \mathcal{X}$. In particular, we get $\sum_{y\in\mathcal{X}}\sum_{x\in A}\pi_x Q_{xy} = 0$. Further, the global balance condition is $\pi Q = 0$, i.e. $\sum_{y\in\mathcal{X}}\pi_y Q_{yx} = 0$ for all $x \in \mathcal{X}$. In particular, we get $\sum_{x\in A}\sum_{y\in\mathcal{X}}\pi_y Q_{yx} = 0$. Further, we have the following identity from change of variables, $\sum_{y\in A}\sum_{x\in A}\pi_x Q_{xy} = \sum_{y\in A}\sum_{x\in A}\pi_y Q_{yx}$. Subtracting the second identity from the first, we get the result. □

Corollary 1.13. *For $A = \{x\}$, the above equation reduces to the full balance equation for state x, i.e.,*

$$\sum_{y\neq x}\pi_x Q_{xy} = \sum_{y\neq x}\pi_y Q_{yx}.$$

Example 1.14. We define two non-negative sequences birth and death rates denoted by $\lambda \in \mathbb{R}_+^{\mathbb{Z}_+}$ and $\mu \in \mathbb{R}_+^{\mathbb{N}}$. A Markov process $X : \Omega \to \mathbb{Z}_+^{\mathbb{R}_+}$ is called a *birth-death process* if its infinitesimal transition probabilities satisfy

$$P_{n,n+m}(h) = (1 - \lambda_n h - \mu_n h \mathbb{1}_{\{n\neq 0\}} - o(h))\mathbb{1}_{\{m=0\}} + \lambda_n h \mathbb{1}_{\{m=1\}} + \mu_n h \mathbb{1}_{\{m=-1\}}\mathbb{1}_{\{n\neq 0\}} + o(h).$$

We say $f(h) = o(h)$ if $\lim_{h\to 0}\frac{f(h)}{h} = 0$. In other words, a birth-death process is a CTMC with generator of the form

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Proposition 1.15. *An ergodic birth-death process in steady-state is time-reversible.*

Proof. Since the process is stationary, the probability flux must balance across any cut of the form $A = \{0, 1, 2, \dots, n\}$, for $n \in \mathbb{Z}_+$. But, this is precisely the equation $\pi_n\lambda_n = \pi_{n+1}\mu_{n+1}$ since there are no other transitions possible across the cut. So the process is time-reversible. □

In fact, the following, more general, statement can be proven using similar ideas.

Proposition 1.16. *Consider an irreducible ergodic CTMC $X : \Omega \to \mathcal{X}^{\mathbb{R}}$ on a countable state space $\mathcal{X}$ with generator matrix $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ having the following property. For any pair of states $x \neq y \in \mathcal{X}$, the transition graph has a* unique *path $x = x_0 \to x_1 \to \cdots \to x_{n(x,y)} = y$ and $y = x_{n(x,y)} \to x_{n-1} \to \cdots \to x_0 = x$ of distinct states. Then the CTMC in steady-state is time reversible.*

Proof. Let the stationary distribution of X be $\pi \in \mathcal{M}(\mathcal{X})$, such that $\pi Q = 0$. Let $x \neq y \in \mathcal{X}$, then either $Q_{xy} = Q_{yx} = 0$ or $Q_{xy}Q_{yx} > 0$. In the former case, the detailed balance equation is satisfied trivially for the pair (x, y). In the latter case, we define a set

$$A_x \triangleq \{z \in \mathcal{X} : z \text{ connected to } x \text{ via } y\}.$$

Clearly, $x \in A_x$ and $y \notin A_x$, and $Q_{zw} = Q_{wz} = 0$ for all $w \in A_x^c \setminus \{y\}$ and $z \in A_x \setminus \{x\}$. If not, then z is connected to x via y and w is connected to x not via y. Then, $x \to y \to z \to w$ and $x \to w$ there are two paths between $x, w \in \mathcal{X}$ and that contradicts the hypothesis. This implies that there are no paths between $A_x \setminus x$ and $A_x^c\, \{y\}$. From the probability flux balance across cuts, we obtain the detailed balance equation

$$\pi_x Q_{xy} = \pi_y Q_{yx}.$$

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

Exercise 1.17. Prove Corollary 1.4 and Corollary 1.9 from Theorem 1.3.