Lecture-20: Uniformization of Markov Processes

1 Alternative construction of continuous time Markov chain

Definition 1.1. Let $Z : \Omega \to \mathcal{X}^{\mathbb{N}}$ be a discrete time Markov chain with a countable state space $\mathcal{X}$, and the corresponding transition probability matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$. Further, we let $\nu \in \mathbb{R}_+^{\mathcal{X}}$ be the set of transition rates such that $p_{xx} = 0$ if $\nu_x > 0$. Let $S_0 \triangleq 0$ and $T : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ be a random sequence, where T_n is a random variable distributed exponentially with rate $\nu_{Z_{n-1}}$ and independent of $(S_0, (Z_0, T_1), \dots, (Z_{n-2}, T_{n-1}))$. We define the nth transition instant $S_n \triangleq \sum_{i=1}^n T_i$. For any initial state $Z_0 \in \mathcal{X}$, we can define a right continuous with left limits piece-wise constant stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ for each $t \in \mathbb{R}_+$ as

$$X_t \triangleq \sum_{n \in \mathbb{N}} Z_{n-1} \mathbb{1}_{[S_{n-1}, S_n)}(t).$$

We define the counting process $N : \Omega \to \mathbb{Z}_+^{\mathbb{R}_+}$ such that the number of transitions of process X in the duration $(0, t]$ for any time $t \in \mathbb{R}_+$ is denoted by $N_t \triangleq \sum_{n \in \mathbb{N}} \mathbb{1}_{\{S_n \leqslant t\}}$. The history of the process until time t is denoted by $\mathcal{F}_t \triangleq \sigma(S_0, (Z_0, T_1), \dots, (Z_{N_t}, A_t))$.

Remark 1. From the definition, any sample path of the process is right-continuous with left limits, and has countable state space $\mathcal{X}$.

Remark 2. A necessary condition for the process X to be defined on index set $\mathbb{R}_+$, is that for each $t \in \mathbb{R}_+$, there exists an n such that $S_n \leqslant t < S_{n+1}$. That is, $P\{N_t < \infty\} = P\{S_\infty > t\} = 1$ for all $t \in \mathbb{R}_+$. This is equivalent to $P\{S_\infty = \infty\} = 1$, or $P\{S_\infty < \infty\} = 0$. Let $\omega \in \{S_\infty < \infty\}$, then we can't define the process for $t > S_\infty$. If the process X is regular, i.e. $\sup_x \nu_x < \infty$, then we can show that $P\{N_t < \infty\} = P\{S_\infty > t\} = 1$ for all $t \in \mathbb{R}_+$.

Lemma 1.2. *Consider the process X defined in Definition 1.1 and $(s, t]$ a non empty interval of $\mathbb{R}_+$. Conditioned on X_s, the increment $N_t - N_s$ of the counting process N is independent of $\mathcal{F}_s$, and depends only on the duration $t - s$ of the increment. That is, for a historical event $H \in \mathcal{F}_s$ and state $x \in \mathcal{X}$,*

$$P(\{N_t - N_s = k\} \mid \{X_s = x\} \cap H) = P_x(\{N_{t-s} = k\}).$$

Proof. From the independence of inter-transition times, we know that T_{N_s+j} is independent of the history $\mathcal{F}_s$ for $j \geqslant 2$ conditioned on the process state $X_s = x$. Hence for any historical event $H \in \mathcal{F}_s$ and state $x \in \mathcal{X}$, we can write the conditional probability of increment $N_t - N_s$ for $t > s$, as

$$\begin{aligned} P(\{N_t - N_s = k\} \mid \{X_s = x\} \cap H) &= P\left(\left\{Y_s + \sum_{i=N_s+2}^{N_s+k} T_i \leqslant t - s < Y_s + \sum_{i=N_s+2}^{N_s+k+1} T_i\right\} \mid \{X_s = x\} \cap H\right) \\ &= P\left(\left\{Y_s + \sum_{i=N_s+2}^{N_s+k} T_i \leqslant t - s < Y_s + \sum_{i=N_s+2}^{N_s+k+1} T_i\right\} \mid \{X_s = x\}\right). \end{aligned}$$

Further, from the memoryless property of an exponential random variable, the excess time Y_s distribution conditioned on $\mathcal{F}_s$ is exponential with rate $\nu_{Z_{N_s}}$, i.e. identically distributed to T_{N_s+1}. Therefore, the conditional distribution of $(Y_s, T_{N_s+2}, \dots, T_{N_s+k})$ given the current process state $X_s = x$, is identical to that of the conditional distribution of inter-transition times $(T_1, T_2, \dots, T_k)$ given initial state $X_0 = x$. It follows that

$$P\left(\left\{Y_s + \sum_{i=N_s+2}^{N_s+k} T_i \leqslant t - s < Y_s + \sum_{i=N_s+2}^{N_s+k+1} T_i\right\} \mid \{X_s = x\}\right) = P_x\{N_{t-s} = k\}.$$

□

Proposition 1.3. *The stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ constructed in Definition 1.1 is a time-homogeneous continuous time Markov chain.*

Proof. For states $x, y \in \mathcal{X}$, we can write the probability of process being in state y, conditioned on any historical events $H \in \mathcal{F}_s$ as

$$P(\{X_t = y\} \mid \{X_s = x\} \cap H) = \sum_{k \in \mathbb{Z}_+} P(\{X_t = y, N_t - N_s = k\} \mid \{X_s = x\} \cap H).$$

From the construction of process X in Definition 1.1, the conditional independence and time homogeneity of counting process from Lemma 1.2, and Markov property of Z, we can write the conditional probability for each $k \in \mathbb{N}$, as

$$P(\{X_t = y, N_t - N_s = k\} \mid \{X_s = x\} \cap H) = P(\{Z_{N_s + k} = y\} \mid \{Z_{N_s} = x\}) P_x \{N_{t-s} = k\} = P_x \{X_{t-s} = y, N_{t-s} = k\}.$$

Thus, we have shown the time homogeneity and Markov property for the process X. □

Theorem 1.4. *A stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ defined on countable state space $\mathcal{X} \subseteq \mathbb{R}$ and having right continuous sample paths with left limits, is a continuous time Markov chain iff for each transition $n \in \mathbb{N}$*
(a) sojourn time T_n is independent and exponentially distributed with rate ν_x where $X_{S_{n-1}} = x$, and
(b) jump transition probabilities $p_{xy} = P_{xy}(S_{n-1}, S_n)$ are independent of S_n and $\sum_{y \neq x} p_{xy} = 1$.

1.1 Generator Matrix

Theorem 1.5. *For a regular continuous time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$, the generator matrix exists and is defined for all $x, y \in \mathcal{X}$ in terms of transition rates $\nu \in \mathbb{R}_+^{\mathcal{X}}$ and jump transition matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$, as*

$$Q_{xx} = -\nu_x, \qquad\qquad Q_{xy} = \nu_x p_{xy}, \quad y \neq x.$$

Proof. Recall that $\lim_{t \downarrow 0} P(t) = I$ and $Q \triangleq \left.\frac{dP(t)}{dt}\right|_{t=0}$. Consider a fixed time $t \in \mathbb{R}_+$ and states $x, y \in \mathcal{X}$. For a regular continuous time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$, we have $\cup_{n \in \mathbb{Z}_+} \{N_t = n\}$ almost surely. Therefore, we can expand the (x, y)th entry of transition matrix in terms of disjoint events $\{N_t = n\}$ as $P_{xy}(t) = P_x \{X_t = y\} = \sum_{n \in \mathbb{Z}_+} P_x \{X_t = y, N_t = n\}$. We can write the upper and lower bound the transition probability $P_{xy}(t)$ as

$$\sum_{n=0}^{1} P_x \{X_t = y, N_t = n\} \leqslant P_{xy}(t) \leqslant \sum_{n=0}^{1} P_x \{X_t = y, N_t = n\} + P_x \{N_t \geqslant 2\}.$$

Since $I_{xy} = \mathbb{1}_{\{x \neq y\}}$, we can write the probability of zero transition in time $(0, t]$ as $P_x \{X_t = y, N_t = 0\} = I_{xy} e^{-\nu_x t}$. To compute the probability of single transition in time $(0, t]$, we apply the tower property of conditional expectation, to write

$$P_x \{X_t = y, N_t = 1\} = \mathbb{1}_{\{x \neq y\}} \mathbb{E}_x \mathbb{E}[\mathbb{1}_{\{X_t = y, T_2 > t - S_1, S_1 \leqslant t\}} | \mathcal{F}_{S_1}] = (1 - I_{xy}) p_{xy} \mathbb{E}_x \mathbb{1}_{\{S_1 \leqslant t\}} e^{-\nu_y (t - S_1)}.$$

Since $\{N_t \geqslant 2\}$ is of order $o(t)$ for small t, we can write

$$\frac{P_{xy}(t) - I_{xy}}{t} = -\nu_x I_{xy} \left(\frac{1 - e^{-\nu_x t}}{\nu_x t} \right) + \nu_x p_{xy} \frac{(e^{-\nu_y t} - e^{-\nu_x t})}{(\nu_x - \nu_y) t} (1 - I_{xy}) + o(t).$$

Taking limit as $t \downarrow 0$, we get the result. □

Corollary 1.6. *For each state $x \in \mathcal{X}$, the generator matrix $Q \in \mathbb{R}^{\mathcal{X} \times \mathcal{X}}$ for a pure jump continuous time Markov chain X, satisfies*

$$0 \leqslant -Q_{xx} < \infty, \qquad\qquad 0 \leqslant Q_{xy} < \infty, \; y \in \mathcal{X} \setminus \{x\} \qquad\qquad \sum_{y \in \mathcal{X}} Q_{xy} = 0.$$

Remark 3. For a homogeneous discrete time Markov chain with one-step transition probability matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$, we can write the n-step transition probability matrix $p^{(n)} = p^n$. That is, for any given stochastic matrix p, we can construct a discrete time Marko chain. We can generalize this notion to homogeneous continuous time Markov chains as well. Given a matrix $Q \in \mathbb{R}^{\mathcal{X} \times \mathcal{X}}$ that satisfies the properties of a generator matrix given in Corollary 1.6, we can construct a homogeneous continuous time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ by finding its transition kernel $P : \mathbb{R}_+ \to \mathcal{M}(\mathcal{X})^{\mathcal{X}}$, by defining $P(t) \triangleq e^{tQ}$ for all $t \in \mathbb{R}_+$. We observe that $P(1) = e^Q$ and we have $P(t) = P(1)^t$ for all $t \in \mathbb{R}_+$. We need to show that such a defined function is indeed a probability transition kernel. We will first show that such a function P satisfies some of the properties of the probability transition kernel, and then show that $P(t)$ is transition matrix at all times $t \in \mathbb{R}_+$.

Theorem 1.7. *Let $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ be a matrix that satisfies the properties of generator matrix given in Corollary 1.6. We define a function $P : \mathbb{R}_+ \to \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ by $P(t) \triangleq e^{tQ}$ for all $t \in \mathbb{R}_+$. Then the function P satisfies the following properties.*

(a) P has the semigroup property, i.e. $P(s+t) = P(s)P(t)$ for all $s,t \in \mathbb{R}_+$.

(b) P is the unique solution to the forward equation, $\frac{dP(t)}{dt} = P(t)Q$ with initial conditon $P(0) = I$.

(c) P is the unique solution to the backward equation, $\frac{dP(t)}{dt} = QP(t)$ with initial condition $P(0) = I$.

(d) For all $k \in \mathbb{N}$, we have $\left.\frac{d^kP(t)}{d^k(t)}\right|_{t=0} = Q^k$.

Proof. Given the definition of P and properties of Q, one can easily check these properties. □

Theorem 1.8. *A finite matrix $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ satisfies the properties of a generator matrix given in Corollary 1.6 iff the function $P : \mathbb{R}_+ \to \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ defined by $P(t) \triangleq e^{tQ}$ is a stochastic matrix for all $t \in \mathbb{R}_+$.*

Proof. Sufficiency has already been seen before, and hence we will focus only on necessity. Accordingly, we assume that $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ satisfies the properties of a generator matrix given in Corollary 1.6, then we will show that $P(t) = e^{tQ}$ is a stochastic matrix. Recall that $Q\mathbf{1}^T = 0$ for all ones column vector $\mathbf{1}^T$, and hence $Q^n\mathbf{1}^T = 0$ for all $n \in \mathbb{N}$. Expanding $P(t)$ in terms of expression for matrix exponentiation, we write $P(t) = I + \sum_{k\in\mathbb{N}} \frac{t^n}{n!}Q^n$. This implies that $P(t)\mathbf{1}^T = \mathbf{1}^T$. □

1.2 Transition graph

The weighted directed transition graph (V, E, w) consists of vertex set $V = \mathcal{X}$ and the edges being

$$E = \left\{(x,y) \in \mathcal{X} \times \mathcal{X} : Q_{xy} > 0, y \neq x\right\}.$$

The weights $w : E \to \mathbb{R}_+$ of the directed edges are given by $w_{xy} = Q_{xy} = \nu_x p_{xy}$.

2 Uniformization

Consider a homogeneous continuous time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ in which the mean time spent in a state is identical for all states, i.e. $\nu_x = \nu$ uniformly for all states $x \in \mathcal{X}$. Since the random amount of time spent in each state is *i.i.d.* with common exponential distribution of rate ν, the associated counting process $N : \Omega \to \mathbb{Z}_+^{\mathbb{R}_+}$ is a Poisson process with rate ν. In this case, we can explicitly characterize the probability transition kernel $P : \mathbb{R}_+ \to \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ for this continuous time Markov chain X in terms of the jump transition probability matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ and uniform transition rate ν. To this end, we use the law of total probability over countable partitions $(\{N_t = n\} : n \in \mathbb{Z}_+)$ to get

$$P_{xy}(t) = \sum_{n\in\mathbb{Z}_+} P_x\{N(t) = n\}\, P(\{X_t = y\} \mid \{X_0 = x, N_t = n\}) = \sum_{n\in\mathbb{Z}_+} p_{xy}^{(n)} e^{-\nu t}\frac{(\nu t)^n}{n!}.$$

This equation could also have been derived by observing that $Q = -\nu(I-p)$ and hence using the exponentiation of matrix, we can write

$$P(t) = e^{-\nu t(I-p)} = e^{-\nu t}e^{\nu t p} = e^{-\nu t}\sum_{n\in\mathbb{Z}_+} p^n \frac{(\nu t)^n}{n!}. \tag{1}$$

Eq. (1) gives a closed form expression for $P(t)$ and also suggests an approximate computation by an appropriate partial sum. However, its application is limited as the transition rates for all states are all assumed to be equal. It turns out that any regular Markov chain can be transformed in this form by allowing hypothetical transitions from a state to itself.

2.1 Uniformization step

Consider a regular continuous time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with bounded transition rates, with finite rate ν such that $\nu_x \leqslant \nu$ for all states $x \in \mathcal{X}$. Since from each state $x \in \mathcal{X}$, the Markov chain leaves at rate ν_x, we could equivalently assume that the transitions occur at a rate ν but only $\frac{\nu_x}{\nu}$ are real transitions and the remaining transitions are fictitious self-transitions.

Construction 2.1 (uniformization). For any regular continuous time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with transition rates $\nu \in \mathbb{R}_+^{\mathcal{X}}$ and jump probability transition matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$, we can find a finite rate $\nu \geqslant \sup_{x \in \mathcal{X}} \nu_x$. We construct a continuous time Markov chain $Y : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with uniform transition rates ν for all states $x \in \mathcal{X}$, and jump probability transition matrix $q \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ defined for all states $x, y \in \mathcal{X}$, as

$$q_{xy} = \frac{\nu_x}{\nu} p_{xy} \mathbb{1}_{\{y \neq x\}} + \left(1 - \frac{\nu_x}{\nu}\right) \mathbb{1}_{\{y = x\}}.$$

The process Y is called the *uniformized* version of process X. This technique of uniformizing the rate in which a transition occurs from each state to any other state by introducing self transitions is called *uniformization*.

Theorem 2.2. *A regular continuous time Markov chain X and its uniformized version Y are identical in distribution.*

Proof. We consider the *i.i.d.* sequence of inter transition times $T : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ with the common exponential distribution of rate ν for the Markov process Y. Assuming the initial state x for the Markov process Y, we define a random sequence of indicators $\xi : \Omega \to \{0,1\}^{\mathbb{N}}$, defined as $\xi_n \triangleq \mathbb{1}_{\{Y_{S_n} \neq x\}}$ for each $n \in \mathbb{N}$. From the definition of uniformized process Y, we know that $P_x \{\xi_1 = \xi_2 = \cdots = \xi_n = 0\} = q_{xx}^n = (1 - \frac{\nu_x}{\nu})^n$, and ξ is an *i.i.d.* sequence. We can define the number of transitions to exit state x, as a stopping time

$$\tau \triangleq \inf \{n \in \mathbb{N} : \xi_n = 1\}.$$

Since ξ is *i.i.d.* Bernoulli, τ is a geometric random variable with success probability $1 - q_{xx} = \frac{\nu_x}{\nu}$. To show that the two Markov processes Y and X have identical distribution, it suffices to show that
(a) $U \triangleq S_\tau = \sum_{n=1}^{\tau} T_n$ is distributed exponentially with rate ν_x, and
(b) $P(\{Y_U = y\} \mid \{Y_0 = x\}) = p_{xy}$.
To see (a), we observe that random sequence T and random variable τ are independent, and hence we can compute the moment generating function of U as

$$M_U(\theta) = \mathbb{E}\left[\mathbb{E}[\prod_{n=1}^{\tau} e^{-\theta T_n} \mid \tau]\right] = \mathbb{E} M_{T_1}^{\tau}(\theta) = \sum_{n \in \mathbb{N}} \left(\frac{\nu}{\nu + \theta}\right)^n q_{xx}^{n-1}(1 - q_{xx}) = \frac{\nu_x}{\nu_x + \theta}.$$

To see (b), from the Markov property of process Y and its embedded jump transition matrix q, we observe that

$$P_x \{Y_U = y\} = \sum_{n \in \mathbb{N}} P_x \{Y_U = y, \tau = n\} = \sum_{n \in \mathbb{N}} P_x \{Y_{S_1} = \cdots = Y_{S_{n-1}} = x, Y_{S_n} = y\} = \sum_{n \in \mathbb{N}} q_{xy} q_{xx}^{n-1} = \frac{q_{xy}}{1 - q_{xx}} = p_{xy}.$$

□

Remark 4. Any regular continuous time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ can be thought of as being in a process that spends a random time in state $x \in \mathcal{X}$ distributed exponentially with rate ν, and then makes a transition to state $y \in \mathcal{X}$ with probability q_{xy}. Then, one can write the probability transition kernel as

$$P_{xy}(t) = \sum_{n=0}^{\infty} q_{xy}^{(n)} e^{-\nu t} \frac{(\nu t)^n}{n!}.$$