Lecture-17: Continuous Time Markov Chains

1 Markov Process

Definition 1.1. Consider a real-valued stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ indexed by positive reals and state space $\mathcal{X}$, adapted to its natural filtration $\mathcal{F}_\bullet = (\mathcal{F}_t : t \in \mathbb{R}_+)$ where $\mathcal{F}_t \triangleq \sigma(X_s, s \leqslant t)$ for all $t \in \mathbb{R}_+$. Then, X is a **Markov process** if it satisfies the Markov property. That is, for any Borel measurable set $A \in \mathcal{B}(\mathcal{X})$, the distribution of the future states conditioned on the present, is independent of the past, and

$$P(\{X_t \in A\} \mid \mathcal{F}_s) = P(\{X_t \in A\} \mid \sigma(X_s)), \text{ for all } s \leqslant t \in \mathbb{R}_+.$$

Definition 1.2. A Markov process $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with countable state space $\mathcal{X}$ is called **continuous time Markov chain (CTMC)**.

Remark 1. The Markov property for the CTMCs can be interpreted as follows. For all times $0 < t_1 < \cdots < t_m < t$ and states $x_1, \ldots, x_m, y \in \mathcal{X}$, we have

$$P(\{X_t = y\} \mid \cap_{k=1}^m \{X_{t_k} = x_k\}) = P(\{X_t = y\} \mid \{X_{t_m} = x_m\}).$$

Example 1.3 (Counting process). Any simple counting process $N : \Omega \to \mathbb{Z}_+^{\mathbb{R}_+}$ with independent increments is a CTMC. This implies any (possibly time-inhomogeneous) Poisson process is a CTMC. Countability of the state space is clear from the definition of the counting process. For Markov property, we observe that for $t > s$, the increment $N_t - N_s$ is independent of $\mathcal{F}_s$. Let $\mathcal{F}_\bullet = (\mathcal{F}_t : t \in \mathbb{R}_+)$ be the natural filtration for process N, then

$$\mathbb{E}[\mathbb{1}_{\{N_t=j\}} \mid \mathcal{F}_s] = \sum_{i \in \mathbb{Z}_+} \mathbb{E}[\mathbb{1}_{\{N_t=j, N_s=i\}} \mid \mathcal{F}_s] = \sum_{i \in \mathbb{Z}_+} \mathbb{1}_{\{N_s=i\}} \mathbb{E}\mathbb{1}_{\{N_t - N_s = j-i\}} = \mathbb{E}[\mathbb{1}_{\{N_t=j\}} \mid \sigma(N_s)].$$

1.1 Transition probability kernel

Definition 1.4. We define the **transition probability** from state x at time s to state y at time $t+s$ as

$$P_{xy}(s, s+t) \triangleq P(\{X_{s+t} = y\} \mid \{X_s = x\}).$$

Definition 1.5. The Markov process has **homogeneous** transitions for all states $x, y \in \mathcal{X}$ and all times $s, t \in \mathbb{R}_+$, if

$$P_{xy}(t) \triangleq P_{xy}(0,t) = P_{xy}(s, s+t).$$

We denote the **transition probability kernel/function** at time t by $P(t) \triangleq (P_{xy}(t) : x, y \in \mathcal{X})$.

Remark 2. We will mainly be interested in continuous time Markov chains with homogeneous jump transition probabilities. We will assume that the sample path of the process X is right continuous with left limits at each time $t \in \mathbb{R}_+$.

Remark 3. Conditioned on the initial state of the process is x, we denote the conditional probability for any event $A \in \mathcal{F}$ as $P_x(A) \triangleq P(A \mid \{X_0 = x\})$ and the conditional expectation for any random variable $Y : \Omega \to \mathbb{R}$ as $\mathbb{E}_x Y \triangleq \mathbb{E}[Y \mid \{X_0 = x\}]$.

Lemma 1.6 (Stochasticity). *Transition kernel $P : \mathbb{R}_+ \to [0,1]^{\mathcal{X} \times \mathcal{X}}$ at each time $t \in \mathbb{R}_+$ is a stochastic matrix.*

Proof. From the countable partition of the state space $\mathcal{X}$, we can write $1 = P_x(\{X_t \in \mathcal{X}\}) = \sum_{y \in \mathcal{X}} P_{xy}(t)$ for any state $x \in \mathcal{X}$. □

Lemma 1.7 (Semigroup property). *Transition probability kernel* $P : \mathbb{R}_+ \to [0,1]^{\mathcal{X}\times\mathcal{X}}$ *satisfies the semigroup property, i.e.* $P(s+t) = P(s)P(t)$ *for all* $s, t \in \mathbb{R}_+$.

Proof. From the Markov property and homogeneity of CTMC, and law of total probability, we can write

$$P_{xy}(s+t) = P_{xy}(0, s+t) = \sum_{z\in\mathcal{X}} P_{xz}(0,s)P_{zy}(s,s+t) = \sum_{z\in\mathcal{X}} P_{xz}(0,s)P_{zy}(0,t) = [P(s)P(t)]_{xy}.$$

□

Lemma 1.8 (Continuity). *Transition probability kernel* $P : \mathbb{R}_+ \to [0,1]^{\mathcal{X}\times\mathcal{X}}$ *for a homogeneous CTMC* $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ *is a continuous function of time* $t \in \mathbb{R}_+$*, such that* $\lim_{t\downarrow 0} P(t) = I$*, the identity matrix.*

Proof. We will first show the continuity of transition kernel at time $t = 0$. From right continuity of sample paths for process X, we have $\lim_{t\downarrow 0} X_t = X_0$ and from continuity of probability functions we get $\lim_{t\downarrow 0} P_x \{X_t = y\} = P_x \{\lim_{t\downarrow 0} X_t = y\} = I_{xy}$.

Fix a time $t \in \mathbb{R}_+$, to write the difference $P(t+h) - P(t) = P(t)(P(h) - I)$ using the semigroup property of the transition kernel. The continuity of transition kernel at time $t = 0$, and boundedness of $P(t)$ implies continuity of $P(t)$ at all times $t \in \mathbb{R}_+$. □

Remark 4. Consider a time-homogeneous CTMC $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with transition kernel $P : \mathbb{R}_+ \to [0,1]^{\mathcal{X}\times\mathcal{X}}$. Then, for all times $0 < t_1 < \cdots < t_m$ and states $x_0, x_1, \dots, x_m \in \mathcal{X}$, we have

$$P(\cap_{k=1}^m \{X_{t_k} = x_k\} \mid \{X_0 = x_0\}) = P_{x_0x_1}(t_1)P_{x_1x_2}(t_2 - t_1)\dots P_{x_{m-1}x_m}(t_m - t_{m-1}).$$

If the initial distribution is $\nu_0 \in \mathcal{M}(\mathcal{X})$ such that $\nu_0(x) = P\{X_0 = x\}$ for each $x \in \mathcal{X}$, then we observe that all finite dimensional distributions of the CTMC X are governed by the initial distribution ν_0 and the transition probability kernel P. That is,

$$P\Big(\cap_{k=1}^m \{X_{t_k} = x_k\}\Big) = \sum_{x_0\in\mathcal{X}} \nu_0(x_0)P_{x_0x_1}(t_1)\dots P_{x_{m-1}x_m}(t_m - t_{m-1}).$$

1.2 Strong Markov property

Lemma 1.9. *A continuous time Markov chain* $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ *has the strong Markov property.*

Proof. It follows from the right continuity of the CTMC process X, and the fact that the map $t \mapsto \mathbb{E}[f(X_{t+s}) \mid \sigma(X_t)]$ is right-continuous for any bounded continuous function $f : \mathcal{X} \to \mathbb{R}$. To see the right continuity of the map, we observe that

$$\mathbb{E}[f(X_{t+s}) \mid \sigma(X_t)] = \sum_{x\in\mathcal{X}} \mathbb{1}_{\{X_t = x\}} \sum_{y\in\mathcal{X}} P_{xy}(s)f(y).$$

Right-continuity of the map follows from the right continuity of the sample paths of process X, right-continuity and boundedness of the kernel function, and boundedness and continuity of f, and bounded convergence theorem. □

Corollary 1.10. *A pure jump CTMC* X *satisfies the following strong Markov property. For any stopping time* τ *adapted to the natural filtration of* X*, finite* $m \in \mathbb{N}$*, finite times* $0 < t_1 < \cdots < t_m$*, any event* $H \in \mathcal{F}_\tau$*, and states* $x_0, x_1, \dots, x_m \in \mathcal{X}$*, we have*

$$P(\cap_{k=1}^m \{X_{t_k+\tau} = x_k\} \mid H \cap \{X_\tau = x_0\}) = P_{x_0}(\cap_{k=1}^m \{X_{t_k} = x_k\}).$$

Remark 5. In particular, for a pure-jump time-homogeneous CTMC X, stopping time τ, and event $H \in \mathcal{F}_\tau$, we have

$$P(\{X_{\tau+s} = y\} \mid \{X_\tau = x\} \cap H) = P_{xy}(s).$$

1.3 Generator Matrix

Definition 1.11 (Exponentiation of a matrix). For a matrix A with spectral radius less than unity, we can define $e^A \triangleq I + \sum_{n\in\mathbb{N}} \frac{A^n}{n!}$.

Lemma 1.12. *For a homogeneous CTMC, we can write the transition kernel* $P(t) = e^{tQ}$ *for all* $t \in \mathbb{R}_+$ *in terms of a constant matrix* $e^Q \triangleq P(1)$.

Proof. This follows from the semigroup property and the continuity of transition kernel $P(t)$. In particular, we notice that $P(n) = P(1)^n$ and $P(\frac{1}{m}) = P(1)^{\frac{1}{m}}$ for all $m,n \in \mathbb{N}$. Since, any rational number $q \in \mathbb{Q}$ can be expressed as a ratio of integers with no common divisor, we get

$$P(q) = P(1)^q, \quad q \in \mathbb{Q}.$$

Since the rationals are dense in reals and P is a continuous function, it follows that $P(t) = P(1)^t$ for all $t \in \mathbb{R}_+$ and the result follows from the definition of Q such that $e^Q = P(1)$. □

Remark 6. From Lemma **??** for a homogeneous CTMC $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$, we can write the probability transition kernel function $t \mapsto P(t) = e^{tQ}$, where $e^Q = P(1)$. The matrix $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ is called the generator matrix for the homogeneous CTMC X. From the Definition **??** for the exponentiation of matrix, this implies that

$$P(t) = I + \sum_{n\in\mathbb{N}} \frac{t^n}{n!} Q^n, \quad t \in \mathbb{R}_+. \tag{1}$$

This relation implies that the probability transition kernel can be written in terms of this fundamental generator matrix Q.

Definition 1.13 (Generator matrix). For a homogeneous continuous time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with transition kernel function $P : \mathbb{R}_+ \to [0,1]^{\mathcal{X}\times\mathcal{X}}$, the **generator matrix** $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ is defined as the following limit when it exists

$$Q \triangleq \lim_{t\downarrow 0} \frac{P(t) - I}{t}.$$

Remark 7. From Eq. (**??**), it is clear that the generator matrix is the limit defined above.

Remark 8. From the semigroup property of probability kernel function and definition of generator matrix, we get the backward equation

$$\frac{dP(t)}{dt} = \lim_{s\downarrow 0} \frac{P(s+t) - P(t)}{s} = \lim_{s\downarrow 0} \frac{(P(s) - I)}{s} P(t) = QP(t), \quad t \in \mathbb{R}_+.$$

Similarly, we can also get the forward equation

$$\frac{dP(t)}{dt} = \lim_{s\downarrow 0} \frac{P(s+t) - P(t)}{s} = P(t) \lim_{s\downarrow 0} \frac{(P(s) - I)}{s} = P(t)Q, \quad t \in \mathbb{R}_+.$$

Both these results need a formal justification of exchange of limits and summation, and we next present a formal proof for these two equations.

Theorem 1.14 (backward equation). *For a homogeneous CTMC $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with transition kernel function $P : \mathbb{R}_+ \to [0,1]^{\mathcal{X}\times\mathcal{X}}$ and generator matrix $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$, we have*

$$\frac{dP(t)}{dt} = QP(t), \quad t \in \mathbb{R}_+.$$

Proof. Fix states $x, y \in \mathcal{X}$ and we consider the lim inf and lim sup of (x,y)th term of $\frac{(P(s)-I)}{s} P(t)$. For any finite subset $F \subseteq \mathcal{X}$ containing x, we obtain

$$\liminf_{s\downarrow 0} \sum_{z\in\mathcal{X}} \frac{(P_{xz}(s) - I_{xz})}{s} P_{zy}(t) \geqslant \sum_{z\in F} \liminf_{s\downarrow 0} \frac{(P_{xz}(s) - I_{xz})}{s} P_{zy}(t) = \sum_{z\in F} Q_{xz} P_{zy}(t).$$

The above inequality holds for any finite set $F \subseteq \mathcal{X}$, and thus taking supremum over increasing sets F, we get the lower bound. For the upper bound, we observe for any finite subset $F \subseteq \mathcal{X}$ containing state x, we have $\sum_{z\notin F}(P_{xz}(s) - I_{xz}) + \sum_{z\in F}(P_{xz}(s) - I_{xz}) = 0$. Therefore,

$$\limsup_{s\downarrow 0} \sum_{z\in\mathcal{X}} \frac{(P_{xz}(s) - I_{xz})}{s} P_{zy}(t) \leqslant \limsup_{s\downarrow 0} \left(\sum_{z\in F} \frac{(P_{xz}(s) - I_{xz})}{s} P_{zy}(t) - \sum_{z\in F} \frac{(P_{xz}(s) - I_{xz})}{s} \right) = \sum_{z\in F} Q_{xz} P_{zy}(t) - \sum_{z\in F} Q_{xz}.$$

The above inequality holds for any finite set $F \subseteq \mathcal{X}$, and thus taking infimum over increasing sets F and recognizing that $\sum_{z\in\mathcal{X}} Q_{xz} = 0$, we get the upper bound. □

Theorem 1.15 (forward equation). *For a homogeneous CTMC $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with transition kernel function $P : \mathbb{R}_+ \to [0,1]^{\mathcal{X}\times\mathcal{X}}$ and generator matrix $Q \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$, we have*

$$\frac{dP(t)}{dt} = P(t)Q, \quad t \in \mathbb{R}_+.$$

Proof. Fix states $x, y \in \mathcal{X}$ and we consider the lim inf and lim sup of (x,y)th term of $P(t)\frac{(P(s)-I)}{s}$. We take a finite set $F \subseteq \mathcal{X}$ containing state y, to obtain the lower bound

$$\liminf_{s\downarrow 0} \sum_{z\in\mathcal{X}} P_{xz}(t)\frac{(P_{zy}(s)-I_{zy})}{s} \geqslant \sum_{z\in F} \liminf_{s\downarrow 0} P_{xz}(t)\frac{(P_{zy}(s)-I_{zy})}{s} \geqslant \sum_{z\in F} P_{xz}(t)Q_{zy}.$$

By taking limiting value for increasing sequence of finite sets $F \subseteq \mathcal{X}$, we obtain the lower bound. To obtain the upper bound, we observe for any finite subset $F \subseteq \mathcal{X}$ containing state y, we have

$$\limsup_{s\downarrow 0} \sum_{z\in\mathcal{X}} P_{xz}(t)\frac{(P_{zy}(s)-I_{zy})}{s} \leqslant \limsup_{s\downarrow 0} \left(\sum_{z\in F} P_{xz}(t)\frac{(P_{zy}(s)-I_{zy})}{s} + \sum_{z\notin F} \frac{P_{zy}(s)}{s} \right) = \sum_{z\in F} P_{xz}(t)Q_{zy} + \sum_{z\notin F} Q_{zy}.$$

The second equality follows from monotone convergence theorem. Taking infimum over increasing sets F and from the fact that $\sum_{z\notin\mathcal{X}} Q_{zy} = 0$, we get the upper bound. □