Lecture-16 : Continuous Time Markov Chains

1 Markov Process

Definition 1.1. For any stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ indexed by positive reals and taking values in $\mathcal{X} \subseteq \mathbb{R}$, the history of the process until time $t > 0$ by is the collection of all the events that can be determined by the realization of the process X until time t. We denote the process history by

$$\mathcal{F}_t \triangleq \sigma(X_u,\ u \leqslant t).$$

Definition 1.2. A real-valued stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ indexed by positive reals, and with state space $\mathcal{X} \subseteq \mathbb{R}$, is a **Markov process** if it satisfies the Markov property. That is for any Borel measurable set $A \in \mathcal{B}(\mathbb{R})$, the distribution of the future states conditioned on the present, is independent of the past, and

$$P(\{X_{t+s} \in A\} \,|\, \mathcal{F}_s) = P(\{X_{t+s} \in A\} \,|\, \sigma(X_s)), \text{ for all } s,t \geqslant 0.$$

A Markov process with countable state space $\mathcal{X}$ is referred to as **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 $x_1, \dots, x_m, y \in \mathcal{X}$, we have

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

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 $\mathcal{F}_s$. Hence for the natural filtration $\mathcal{F}_\bullet$,

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

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

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

Definition 1.5. The Markov process has **homogeneous** transitions for all states $x, y \in \mathcal{X}$ and all times $s, t \geqslant 0$, 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 $P(t) = (P_{xy}(t) : x, y \in \mathcal{X})$.

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 $t \in \mathbb{R}_+$.

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 get $1 = P(\{X(t) \in \mathcal{X}\} \,|\, \{X(0) = x\}) = \sum_{y \in \mathcal{X}} P_{xy}(t)$ for any $x \in \mathcal{X}$. □

Lemma 1.7 (semigroup). *Transition kernel satisfies the semigroup property, i.e. $P(s+t) = P(s)P(t),\ s,t \in \mathbb{R}_+$.*

Proof. From the Markov property and homogeneity of CTMC, and law of total probability, we can write the (x,y)th entry of kernel matrix $P(s+t)$ as

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

Result follows since $x,y \in \mathcal{X}$ were chosen arbitrarily. $\square$

Lemma 1.8 (continuity). *Transition 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. That is, $P_{xx}(0) = 1$ and $P_{xy}(0) = 0$ for all $y \neq x \in \mathcal{X}$.*

Proof. From the continuity of probability functions and right continuity of the process at time $t = 0$, we get that $\lim_{t\downarrow 0} P(t) = I$. Using the semigroup property of the transition kernel, we can write $P(t+h) - P(t) = P(t)(P(h) - I)$. The continuity of transition kernel at time $t = 0$, and boundedness of $P(t)$ implies continuity of $P(t)$ at all times $t > 0$. $\square$

Lemma 1.9 (continuity). *Transition kernel P for a homogeneous CTMC is continuous in time $t \in \mathbb{R}_+$,* $\lim_{t\downarrow 0} P(t) = I$.

Proof. From the semi-group property of probability kernel, we have $P(t+h) - P(t) = P(t)(P(h) - I)$. Since probability is a bounded function, it suffices to show continuity at $t = 0$. The continuity at $t = 0$ follows from the continuity of probability functions and alternate characterization of homogeneous CTMC. $\square$

Lemma 1.10 (continuity). *Transition kernel P for a homogeneous CTMC is continuous in time $t \in \mathbb{R}_+$,* $\lim_{t\downarrow 0} P(t) = I$.

Proof. From the semi-group property of probability kernel, we have $P(t+h) - P(t) = P(t)(P(h) - I)$. Since probability is a bounded function, it suffices to show continuity at $t = 0$. The continuity at $t = 0$ follows from the continuity of probability functions and alternate characterization of homogeneous CTMC. $\square$

Remark 3. Since each entry of transition kernel $P(t)$ is a probability, semigroup property leads to characterization of the kernel $P(t)$ completely.

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 $P(t) = e^{tQ}$ 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 continuous function, it follows that $P(t) = P(1)^t$ for all $t \in \mathbb{R}$ and the result follows from definition of $e^Q = P(1)$. $\square$

Proposition 1.13. *For a time-homogeneous CTMC $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$, with transition kernel P, for all times $0 < t_1 < \dots < t_m$ and states $x_0, x_1, \dots, x_m \in \mathcal{X}$, we have*

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

Corollary 1.14. *All finite dimensional distributions of the CTMC $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ is governed by the initial distribution.*

Proof. Let ν_0 be the initial distribution of the CTMC X, such that $\nu_0(x_0) = P\{X_0 = x_0\}$ for each $x_0 \in \mathcal{X}$. For all finite index sets $F \subset \mathbb{R}_+, |F| = m$ and states $(x_j \in \mathcal{X} : j \in [m])$, we have

$$P(\cap_{t_j\in F} \{X_{t_j} = x_j\}) = \sum_{x_0\in\mathcal{X}} \nu_0(x_0)P_{x_0x_1}(t_1)\dots P_{x_{m-1}x_m}(t_m - t_{m-1}).$$

$\square$

1.2 Excess time in a state

Definition 1.15. From the definition of excess time as the time until next transition, we can write the excess time at time $t \in \mathbb{R}_+$ for the CTMC X as

$$Y_t \triangleq \inf\{s > 0 : X_{t+s} \neq X_t\}.$$

Remark 4. We observe that Y_t is the excess remaining time at time t, the process spends in state X_t. That is, $X_{t+Y_t} \neq X_t$.

For a homogeneous CTMC X, the distribution of excess time Y_t conditioned on the current state X_t, doesn't depend on time t. Hence, we can define the following conditional complementary distribution of excess time as

$$\bar{F}_x(u) \triangleq P(\{Y_t > u\} \mid \{X_t = x\}).$$

Lemma 1.16. *For a homogeneous CTMC X, there exists a positive sequence* $\nu : \mathcal{X} \to \mathbb{R}_+$*, such that*

$$\bar{F}_x(u) \triangleq P(\{Y_t > u\} \mid \{X_t = x\}) = e^{-u\nu_x}, \quad x \in \mathcal{X}.$$

Proof. We fix a state $x \in \mathcal{X}$, and observe that the function $\bar{F}_x \in [0,1]$ is non-negative, non-increasing, and right-continuous in u. Using the Markov property and the time-homogeneity, we can show that $\bar{F}_x$ satisfies the semigroup property. In particular,

$$\bar{F}_x(u+v) = P(\{Y_t > u+v\} \mid \{X_t = x\}) = P(\{Y_t > u, X_{t+u} = x, Y_{t+u} > v\} \mid \{X_t = x\}) = \bar{F}_x(u)\bar{F}_x(v).$$

The only continuous function $\bar{F}_x \in [0,1]$ that satisfies semigroup property is an exponential function with a negative exponent. □

Definition 1.17. For a CTMC X, a state $x \in \mathcal{X}$ is called

(i) **absorbing** if $\nu_x = 0$,

(ii) **stable** if $\nu_x \in (0, \infty)$, and

(iii) **instantaneous** if $\nu_x = \infty$.

Remark 5. The sojourn time in an absorbing state is ∞, zero in an instantaneous state, and almost surely finite and non-zero in a stable state.

Definition 1.18. A homogeneous CTMC with no instantaneous states is called a **pure jump** CTMC. A pure jump CTMC with

(i) all stable states and $\inf_{x \in \mathcal{X}} \nu_x \geqslant \nu > 0$ is called **stable**, and

(ii) $\sup_{x \in \mathcal{X}} \nu_x \leqslant \nu < \infty$ is called **regular**.

We will focus on pure jump CTMCs only.

Example 1.19 (Poisson process). Consider the counting process $N : \Omega \to \mathbb{Z}_+^{\mathbb{R}_+}$ for a Poisson point process with homogeneous rate λ. Using the stationary independent increment property, we have for all $u \geqslant 0$

$$\bar{F}_i(u) = P(\{N_{t+u} = i\} \mid \{N_t = i\}) = P\{N_{t+u} - N_t = 0\} = P\{Y_t > u\} = e^{-\lambda u}.$$

A Poisson process with finite non-zero rate is a pure-jump CTMC with stable states.

1.3 Strong Markov property

Consider a probability space $(\Omega, \mathcal{F}, P)$ and a continuous filtration $\mathcal{F}_\bullet = (\mathcal{F}_t \subseteq \mathcal{F} : t \in \mathbb{R}_+)$ defined on this space.

Definition 1.20. A random variable τ is a **stopping time** if $\{\tau \leqslant t\} \in \mathcal{F}_t$ for each $t \in \mathbb{R}_+$. That is, a random variable τ is a stopping time if the event $\{\tau \leqslant t\}$ can be determined completely by the history $\mathcal{F}_t = \sigma(X_u,\, u \leqslant t)$. An almost surely finite stopping time τ is called **proper**.

Definition 1.21. A stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ has **strong Markov property** if for any proper stopping time τ, and set $A \in \mathcal{B}(\mathcal{X})$, we have

$$P(\{X_{\tau+s} \in A\} \bigm| \mathcal{F}_\tau) = P(\{X_{\tau+s} \in A\} \bigm| \sigma(X_\tau)).$$

Lemma 1.22. *A continuous time Markov chain X has the strong Markov property.*

Proof. It follows from the right continuity of the CTMC process, and the fact that the map $t \mapsto \mathbb{E}[f(X_{s+t}) \mid \sigma(X_t)]$ 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

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

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. □

Theorem 1.23. *A pure jump CTMC X satisfies the following strong Markov property. For any proper stopping time τ, finite $m \in \mathbb{N}$, finite times $0 < t_1 < \cdots < t_m$, any event $H \in \mathcal{F}_\tau$ and states $x_0, x_1, \ldots, x_m \in \mathcal{X}$, we have*

$$P(\cap_{k=1}^m \{X_{t_k+\tau} = x_k\} \mid H \cap \{X_\tau = x_0\}) = P(\cap_{k=1}^m \{X_{t_k} = x_k\} \mid \{X_0 = x_0\}).$$

Remark 6. In particular for a pure-jump time-homogeneous CTMC X, proper stopping time τ, and event $H \in \mathcal{F}_\tau$, we have

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