THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH4240 - Stochastic Processes - 2023/24 Term 2

Chapter III Markov Jump Process

1 Introduction to Markov jump process

• From now on we consider the continuous-time stochastic process.

• **Jump process**: It is a continuous-time stochastic process $\{X_t\}_{t\geq 0}$ taking values in a countable set S. A sample path of X_t is described as follows:

- Let $X_0 = x_0 \in S$. It stays at x_0 until time $\tau_1 > 0$ and jump to another state $x_1 (\neq x_0) \in S$. Assume τ_1 is finite.
- Repeat the previous procedure in a similar way: it stays at x_1 until time $\tau_2 (> \tau_1)$ and jump to another state $x_2 (\neq x_1) \in S$. Assume τ_2 is finite.
- $\cdots$

We also assume $\lim_{n\to\infty} \tau_n = \infty$ (*No blow up*). Example: The model for customer arrival.

• Probability structure: A state x is absorbing if once it is reached the process remains there forever. For a non-absorbing state x, we need two things

- $F_x(t)$ to describe the distribution of the waiting time τ_1 to jump
- Q_{xy} to describe the transition probability to jump from x to $y(\neq x)$:

$$Q_{xx} = 0, \quad \sum_y Q_{xy} = 1. \tag{1}$$

We also assume that "τ_1 (*the waiting time to jump*)" and "X_{τ_1} (*where to jump*)" are independent:

$$P_x(\tau_1 \leq t, X_{\tau_1} = y) = F_x(t) Q_{xy}. \tag{2}$$

Then, the continuous-time jump process with such probability structure is described by

$$P_{xy}(t) := P_x(X_t = y), \tag{3}$$

that is the probability that the process starting in state x will be in state y at time $t \geq 0$. $P_{xy}(t)$ is called the **transition function**.

• **Markov property**:

$$P\big(X_t = y | X_{s_1} = x_1, \cdots X_{s_n} = x_n, X_s = x\big) = P(X_t = y | X_s = x), \tag{4}$$

for all $0 \le s_1 \le \cdots \le s_x \le s \le t$ and for all states $x_1, \cdots, x_n, x, y$.

In this course we always assume that the process is **time-homogeneous**, meaning that for any $0 \le s < t$,

$$P(X_t = y | X_s = x) = P(X_{t-s} = y | X_0 = x) = P_x(X_{t-s} = y). \tag{5}$$

A **Markov jump process** (MJP) means a continuous-time jump process satisfying the above Markov property. Note that it is NOT obvious that such MJP exists. We will first look at it by the Poisson process (Model: customer arrival), and more examples of MJPs will be provided later on.

From now on, we always consider the MJP which is time-homogeneous.

• Considering a non-absorbing state x, the waiting to jump τ_1 turns out to be an exponential r.v. Indeed,

(a) One can show that for $X_0 = x$, the r.v.

$$\tau_x := \inf\{t > 0 : X_t \neq x\} \tag{6}$$

(*the first time to jump*) is memoryless, meaning

$$P(\tau_x > s + t | \tau_x > s) = P(\tau_x > t), \quad \forall\, s, t \ge 0. \tag{7}$$

(Think about the model of waiting for an unreliable bus driver: *If we have been waiting for s units of time then the probability we must wait for t more units of time is the same as if we have not waited at all!*) See the lecture for the proof (*Use the Markov property*).

(b) One can further show that any memoryless r.v. must be exponential; see the lecture for the proof. For instance, for τ_x,

$$P(\tau_x > t) = e^{-q_x t}, \quad q_x = \frac{1}{E(\tau_x)}. \tag{8}$$

Here $q_x (> 0)$ represents the rate leaving x. Thus, the density function of τ_1 is $q_x e^{-q_x t}$, and

$$F_x(t) = P_x(\tau_x \le t) = 1 - e^{-q_x t}. \tag{9}$$

2 Poisson process

• There are several ways to define the Poisson process. Here, we would use the waiting time to do it.

• The PP is introduced as follows:

(a) We start from $\xi_n (\sim \xi)$, $n = 1, 2, \cdots$, which are i.i.d. exp. r.v. with parameter $\lambda > 0$:

$$P(\xi > t) = e^{-\lambda t}, \quad \lambda = \frac{1}{E(\xi)}, \tag{10}$$

where **ξ is regarded as the waiting time for the next arrival**, and λ is understood to be the arrival rate.

(b) Then, we define

$$\tau_n := \xi_1 + \cdots + \xi_n, \quad n = 1, 2, \cdots, \tag{11}$$

and $\tau_0 := 0$. **τ_n is regarded as the time for the n^{th} arrival.**

(c) Now, for $t \geq 0$, we define

$$X_t := \max\{n \geq 0 : \tau_n \leq t\}, \tag{12}$$

regarded as the **NO of arrivals in $[0, t]$.** Note $X_0 = 0$, i.e., no arrival at initial time.

• From the construction of $\{X_t\}_{t \geq 0}$, one can show that *for any given $t > 0$, X_t has Poisson distribution with mean λt*:

$$P(X_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \cdots. \tag{13}$$

Recall that

$$E(X_t) = \lambda t$$

is the expected number of arrivals in $[0, t]$, and hence $\lambda = \frac{\lambda t}{t - 0}$ means the arrival rate. The proof of (13) is based on the identity

$$\{X_t = n\} = \{\tau_n \leq t < \tau_{n+1}\}; \tag{14}$$

see the lecture for additional details.

• One can further conclude:

(i) $X_0 = 0$.

(ii) For $0 < s < t$, $X_t - X_s$ has Poisson distribution with mean $\lambda(t-s)$ and is independent of X_s.

(iii) For any increment $0 \leq t_1 < \cdots < t_n$, $X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}}$ are independent.

These three properties are also often used as the definition of Poisson process.

• Moreover, one can show that $\{X_t\}_{t \geq 0}$ satisfies the Markov and time-homogenous property, and hence is a MJP, usually called the **Poisson process**.

3 Basic properties of Markov jump process

• Consider a general MJP $\{X_t\}_{t\geq 0}$ with countable state space S. Recall $P_{xy}(t) = P(X_t = y|X_0 = x)$. In general it is convenient to write it as the matrix form

$$P(t) = [P_{xy}(t)]. \tag{15}$$

• One can show

$$P(t+s) = P(t)P(s), \tag{16}$$

that is,

$$P_{xy}(t+s) = \sum_{z\in S} P_{xz}(t)P_{zy}(s). \tag{17}$$

This is the so-called **Chapman-Kolmogorov equation**.

• Recall the probability structure of MJP introduced before. We expect to bridge a relation between $P(t)$ and q_x (*the leaving rate*; the parameter of the exponential distribution for the waiting time to jump away from x) as well as Q_{xy} (*the probability for where to jump*).

Heuristically (of course it can be made rigorous; see the lecture for detailed proof), one has

(a) $P(t)$ is differentiable. Set $D := P'(0)$, and denote $D = [q_{xy}]$. D is called the **rate matrix**. For D, one is able to show

$$\sum_y q_{xy} = 0 \quad \text{(row sum is zero)}, \tag{18}$$

$$q_{xx} = -q_x \leq 0 \quad \text{(the rate to jump away from } x\text{)}, \tag{19}$$

and

$$q_{xy} \geq 0 \quad \text{for } y \neq x \quad \text{(the rate to jump away from } x \text{ to being in } y\text{)}. \tag{20}$$

Note

$$\sum_{y\neq x} q_{xy} = -q_{xx} = q_x, \tag{21}$$

or if $q_x \neq 0$ (thus > 0),

$$\sum_{y\neq x} \frac{q_{xy}}{q_x} = 1. \tag{22}$$

Hence, q_{xy}/q_x is understood to be the probability that the process jumps to y from x.

(b) Recall $Q = [Q_{xy}]$ is the Markov matrix associated with the process. One is able to show

$$Q_{xy} = \begin{cases} \dfrac{q_{xy}}{q_x} & \text{if } x \neq y \text{ and } q_x \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{23}$$

From these properties, we see it is more convenient to first have D so as to have q_x (equivalently $F_x(t)$) and Q and then derive $P(t)$. See the lecture for examples.

• Assuming $P(t)$ is differentiable, it follows from the Chapman-Kolmogorov equation that

$$P'(t) = P(t)D \qquad \text{forward equation,} \tag{24}$$
$$P'(t) = DP(t) \qquad \text{backward equation.} \tag{25}$$

Note $P(0) = I$, i.e., $P_{xy}(0) = \delta_{xy}$. The **solution to the forward equation** is formally written as

$$P(t) = e^{tD} := \sum_{n=0}^{\infty} \frac{(tD)^n}{n!}. \tag{26}$$

To find the p.d.f. of X_t, set

$$p_y(t) := P(X_t = y), \quad \text{or} \quad \vec{p}(t) = [p_y(t)]_{y \in S} \text{ in the vector form.} \tag{27}$$

Similar to what we showed in the discrete-time Markov chain,

$$\vec{p}(t) = \vec{p}(0)P(t), \tag{28}$$

so that

$$\vec{p}(t) = \vec{p}(0)e^{tD}, \tag{29}$$

• Assume that S is finite and $D = G\,\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)G^{-1}$, where λ_i are the eigenvalues of D and $G = [\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n]$ with $\vec{e}_i$ being the (column) eigenvectors associated with λ_i. Then,

$$P(t) = e^{tD} = G\mathrm{diag}\left(e^{\lambda_1 t}, \cdots, e^{\lambda_n t}\right) G^{-1}. \tag{30}$$

This is a convenient way for finding $P(t)$ instead of directly solving ODEs.

4 Important examples: Birth and death processes

• The Poisson process defined before is a special **pure-birth** (with a constant birth rate) process. You need to know how to compute $P(t)$ in terms of a pure-birth process; see the lecture.

• **Branching process**: Each particle waits to either split into two particles with probability p or vanish with probability $1 - p$. The waiting time is an exponential r.v. with rate λ. Set up the model to describe the number of particles at time t, and find the rate matrix. What if we allow new particles to immigrate into the system at a rate α?

• **Queuing model**: The knowns are the arrival rate λ (arrivals are Poisson) and the service rate μ (exponential r.v.). Note that the arrival and service are independent. Set up the model in terms of the number of servers to describe the number of persons waiting for service at time t: $M/M/k$ $(k = 1, 2, \cdots)$ and $M/M/\infty$.

5 Limiting properties

• $\vec{\pi}$ is a **stationary distribution** if (i) $\vec{\pi}$ is a probability vector, and (ii) $\vec{\pi}$ is stationary, i.e. $\vec{\pi}P(t) = \vec{\pi}$ or equivalently

$$\sum_{x \in S} \pi(x) P_{xy}(t) = \pi(y), \quad \forall\, y,\ \forall\, t. \tag{31}$$

One can show that $\vec{\pi}$ is a stationary distribution if and only if $\vec{\pi}$ is a probability vector and satisfies

$$\vec{\pi} D = 0, \tag{32}$$

where $D = P'(0)$ is the rate matrix. Apply this to a general birth and death process, particularly, queue models, to check the condition that the process has a stationary distribution.

• **Recurrence/transience**: Define $Z_n = X_{\tau_n}$, $n = 0, 1, 2 \cdots$ with $\tau_0 := 0$, where τ_n means the time for the n^{th} jump. Given a general Markov jump process introduced before, one can show that $\{Z_n\}_{n \geq 0}$ *is a discrete-time Markov chain with Q as transition matrix*. To check recurrence/transience, it suffices to only consider Q. Similarly, A MJP is irreducible if $\rho_{xy} > 0$ for all x, y. Apply this to a general birth and death process to check the condition that the process is recurrent or not, and irreducible or not.

• **Mean return time** and relation with the stationary distribution: As introduced in the past chapter, $m_x := E_x(T_x)$ stands for the mean return time. A recurrent state x is **positive recurrent** if $m_x < \infty$; **null recurrent** if $m_x = \infty$. One can show that an irreducible positive recurrent MJP must admit a unique stationary distribution $\vec{\pi}$, which, unless S consists of a single necessarily absorbing state, is given by

$$\pi(x) = \frac{1}{q_x m_x}, \quad x \in S. \tag{33}$$

• Long-term behavior of the process: For an irreducible MJP, $P(t)$ (as a Markov matrix) is always aperiodic for any given $t > 0$. For an irreducible positive recurrent MJP having stationary distribution $\vec{\pi}$, we have

$$\lim_{t \to \infty} P_{xy}(t) = \pi(y), \quad \forall\, x, y \in S. \tag{34}$$

——The End, Updated on April 17——