Week 10

Continuous Markov processes

28 Definition and basic properties

Example 10.1. A gas station has a single pump and no space for vehicles to wait (if a vehicle arrives and the pump is not available, it leaves). Vehicles arrive to the gas station following a Poisson process with a rate of $\lambda = 3/20$ vehicles per minute. The refuelling time can be modelled with an exponential random variable with mean 8 minutes for each vehicle, that is, the services rate is $\mu = 1/8$ per minute.

Let $X_t, t \geq 0$, represents the state of the gas station at the time t. X_t is a continuous-time random process with discrete state space $S = 1$ and 2, where $1 =$ "station occupied" and $2 =$"empty".

Note that X_t is not a discrete MC, but it satisfied the Markov property: the behavior of the future of the process only depends upon the current state and not any of the rest of the past.

Continuous-time Markov chain

Let $\{X_t, t \geq 0\}$ be a continuous-time random process with discrete state space S. $\{X_t, t \geq 0\}$ is said to be a *continuous-time Markov chain* if for any $0 \leq s_1 < s_2 < ... < s_n < s < t$ and $i_1, i_2, ..., i_n, i, j \in S$,

$$P(X_t = j \mid X_{s_1} = i_1, ..., X_{s_n} = i_n, X_s = i) = P(X_t = j \mid X_s = i),$$

whenever these conditional probabilities are well defined.

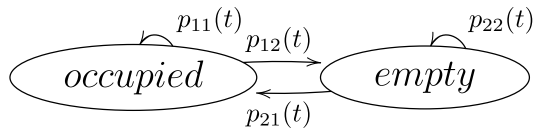
If $P(X_{t+s} = j \mid X_s = i)$ only depends on t for all $s, t \geq 0$, the process is said to be a **time-homogeneous MC** (or to have stationary transition probabilities). In this case, we set

$$p_{ij}(t) = P(X_{t+s} = j \mid X_s = i) = P(X_t = j \mid X_0 = i), \quad t \geq 0.$$

As in the discrete MC case, we have

- $p_{ij}(t) \geq 0$ and $\sum_{j \in S} p_{ij}(t) = 1$, for all $t \geq 0, i, j \in S$;
- $p_{ij}(s+t) = \sum_{k \in S} p_{ik}(s)\, p_{kj}(t)$, for all $s, t \geq 0, i, j \in S$;
- $P_j(s+t) = \sum_{k \in S} P_k(s)\, p_{kj}(t)$, for all $s, t \geq 0, j \in S$, where $P_j(s) = P(X_s = j)$.

Example 10.1 (continuous). The state transition diagram at time t:



Let Y denote the time of a vehicle refuelling the gas. Let $Z(t)$ denote the numbers of vehicles by time t. We have

$$Y \sim Exp(1/8), \qquad Z(t) \sim Poisson(\lambda t), \quad \lambda = 3/20.$$

For small $h > 0$, we have

$$\begin{aligned} p_{11}(h) &= P(X_{h+s} = 1 \mid X_s = 1) \\ &= P(\text{station is occupied during } (s, s+h]) \\ &= P(Y > h + s \mid Y > s) = P(Y > h) = e^{-h/8} \end{aligned}$$

$$\begin{aligned} p_{12}(h) &= P(X_{h+s} = 2 \mid X_s = 1) \\ &= P(\text{station is occupied at time } s\text{, but empty at time } s+t) \\ &= P(Y \leq h + s \mid Y > s) = 1 - P(Y > h) = 1 - e^{-h/8} \end{aligned}$$

$$\begin{aligned} p_{21}(h) &= P(X_{h+s} = 1 \mid X_s = 2) \\ &= P(\text{at least one vehicle has arrived during } (s, s+h]) \\ &= 1 - e^{-\lambda h} \end{aligned}$$

$$\begin{aligned} p_{22}(h) &= P(\text{no vehicle arrives during } (s, s+h]) \\ &= e^{-\lambda h}. \end{aligned}$$

Basic Assumption

Many important processes are continuous homogeneous MC satisfying the following assumption:

$$
\begin{aligned}
p_{ii}(h) &= P(X_h = i \mid X_0 = i) = 1 - \lambda_i\, h + o(h), && (1)\\
p_{ij}(h) &= P(X_h = j \mid X_0 = i) = q_{ij} h + o(h), \quad \text{for } i \neq j\ . && (2)
\end{aligned}
$$

In this situation, for each $i \in S$, we must have

$$
\lambda_i = \sum_{j \in S, j \neq i} q_{ij}, \quad q_{ij} \geq 0, \quad \lambda_i \geq 0.
$$

Note: λ_i is often called **the transition rate out of state** i, i.e., the rate at which the chain leaves state i. As a conseqeuce, $1/\lambda_i$ is the mean holding time for the chain being in state i.

Pure birth Process

Let $\{X_t, t \geq 0\}$ be a random process with state space $S = \{0, 1, 2, ...\}$. $\{X_t, t \geq 0\}$ is said a (pure) **birth process** if $\{X_t, t \geq 0\}$ is a time-homogeneous MC satisfying

(i). $P(X_{t+h} = i \mid X_t = i) = 1 - \lambda_i h + o(h)$, as $h \to 0$;

(ii). $P(X_{t+h} = i + 1 \mid X_t = i) = \lambda_i h + o(h)$, as $h \to 0$;

(iii). $P(X_{t+h} \geq i + 2 \mid X_t = i) = o(h)$, as $h \to 0$,

where $\lambda_i > 0, i \in S$, are called the **birth rate** of the process.

A **Poisson process** with rate $\lambda > 0$ is a birth process with $\lambda = \lambda_i$ for all $i \in S$.

Note: Birth (Poisson) processes arise in situations where one is interested in the number of "customers" which arrive up to time $t \geq 0$. Note that

$$P(X_{t+h} = i + 1 \mid X_t = i) = P(X_{t+h} - X_t = 1 \mid X_t = i).$$

The general birth processes allow the chance of an event occurring at a given instant of time to depend upon the number of events which have already occurred. An example of this phenomenon is the reproduction of living organisms such as a growth of a colony bacteria and the spread of epidemics.

Birth and Death Process

Let $\{X_t, t \geq 0\}$ be a time-homogeneous MC with state space $S = \{0, 1, 2, ...\}$. $\{X_t, t \geq 0\}$ is said to be a **Birth and Death process** if, as $h \to 0$,

$$P(X_{t+h} = k \mid X_t = i) = \begin{cases} 1 - \lambda_i h + o(h) & \text{for } k = i, \\ b_i h + o(h) & \text{for } k = i+1, \\ \mu_i h + o(h) & \text{for } k = i-1, \\ o(h) & \text{for } |k-i| \geq 2, \end{cases}$$

where $\lambda_i = b_i + \mu_i$ and positive b_i's and μ_i's ($\mu_0 = 0$) are called the **birth rates** and **death rates**, respectively.

Note: Birth and death processes play a key role in queueing theory. Queues or Queueing Systems arise in the situations where "customers" arrive to seek some kind of services, such as customers queueing up before the m-cashiers in a supermarket, telephone callers waiting for one of the lines of an exchange to become available, and so on.

The **holding time** of a continuous MC

When the process enters state i, the time it spends there before it leaves state i is called the **holding time** in state i. Let T_i denote the holding time in state i.

Th10.1 T_i is exponential distributed. In particular, if the transition probability $p_{ij}(t)$ satisfy (1) and (2), i.e.,

$$\begin{aligned} p_{ii}(h) &= P(X_h = i \mid X_0 = i) = 1 - \lambda_i h + o(h), \\ p_{ij}(h) &= P(X_h = j \mid X_0 = i) = q_{ij}h + o(h), \quad \text{for } i \neq j\ . \end{aligned}$$

then $T_i \sim Exp(\lambda_i), \quad i \in S$.

Proof.

$$\begin{aligned} P(T_i > s + t \mid T_i > s) &= P(X_{s+t} = i \mid X_s = i) \\ &= P(X_t = i \mid X_0 = i) = P(T_i > t). \qquad (3) \end{aligned}$$

Therefore, the distribution of T_i has the memoryless property, which implies that it is exponential.

In particular, if (1) and (2) are satisfied, we have $T_i \sim Exp(\lambda_i), \quad i \in S$. Indeed, by letting $g(t) = P(T_i > t) = P(X_t = i \mid X_0 = i)$, it follows from (3) that

$$g(s + t) = g(s)g(t), \quad s, t \geq 0,$$

indicating

$$\begin{aligned} g(t) &= g(t/n + ... + t/n) = g^n(t/n) \\ &= \left[1 - \lambda_i t/n + o(t/n)\right]^n \sim e^{-\lambda_i t}, \end{aligned}$$

i.e., $T_i \sim Exp(\lambda_i)$.

29 The Embedded Markov Chain and the generator matrix

For a continuous MC satisfying (1) and (2), i.e.,

$$\begin{aligned} p_{ii}(h) &= P(X_h = i \mid X_0 = i) = 1 - \lambda_i h + o(h), \\ p_{ij}(h) &= P(X_h = j \mid X_0 = i) = q_{ij} h + o(h), \quad \text{for } i \neq j \ , \end{aligned}$$

the chain will jump to the next state at time T_i, where $T_i \sim Exp(\lambda_i)$.

Where does it jump to? The probability jump from state i to $j \neq i$ is

$$\begin{aligned} p_{ij} &= \lim_{h \to 0} P\big(X_{t+h} = j \mid X_t = i, X_{t+h} \neq i\big) \\ &= \lim_{h \to 0} \frac{P\big(X_{t+h} = j \mid X_t = i\big)}{P\big(X_{t+h} \neq i \mid X_t = i\big)} \\ &= \lim_{h \to 0} \frac{p_{ij}(h)}{1 - p_{ii}(h)} = q_{ij}/\lambda_i. \end{aligned} \tag{4}$$

Therefore, the process acts like:

It remains in state i for a period with mean $1/\lambda_i$ ($T_i \sim Exp(\lambda_i)$), and then jump from state i to $j \neq i$ with probability $p_{ij} = q_{ij}/\lambda_i$; and so on.

The embedded Markov chain or jump chain

For the jump probability p_{ij}, if $\lambda_i \neq 0$, we have

$$p_{ii} = 0, \qquad \sum_{j \in S, j \neq i} p_{ij} = \lambda_i^{-1} \sum_{j \in S, j \neq i} q_{ij} = 1;$$

if $\lambda_i = 0$, then $p_{ii} = 1$ and $p_{ij} = 0$ when $i \neq j$.

The matrix P whose $(i,j)th$ entry is p_{ij} is the one-step transition probability matrix of a (discrete-time) Markov chain. We call this discrete-time chain **jump chain** or **the embedded Markov chain**.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} & \cdots \\ p_{21} & p_{22} & \cdots & p_{2r} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ p_{r1} & p_{r2} & \cdots & p_{rr} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$$

Every continuous-time Markov chain has an associated embedded discrete-time Markov chain.

Note: If $\lambda_i = 0$, i is an absorbing state.

Example 10.1 (continuous). The embedded Markov chain is

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Recall that

$$p_{11}(h) = e^{-h/8}, \quad p_{22}(h) = e^{-\lambda h}.$$

Example 10.2. Consider a Poisson process N_t with rate $\lambda > 0$. We have

$$\begin{aligned} p_{ii}(h) &= 1 - \lambda h + o(h) \\ p_{i,i+1}(h) &= \lambda h + o(h) \\ p_{i,i+j}(h) &= o(h), \quad j \geq 2 \text{ or } j < 0 \end{aligned}$$

which yields that $\lambda_i = \lambda$, $q_{i,i+1} = \lambda$ and $q_{i,i+j} = 0, j \geq 2$ or $j < 0$. Hence

$$p_{ii} = 0, \quad p_{i,i+1} = 1, \quad i = 0, 1, 2, ..$$

i.e.,

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & ... \\ 0 & 0 & 1 & 0 & .. \\ . & . & . & & \end{pmatrix}.$$

Remark

A continuous Markov chain can be understood to have two components. First, it has a discrete-time jump Markov chain that gives the transition probabilities p_{ij} from state i to j. Second, we have a holding time parameter λ_i that controls the amount of time spent in state i.

The generator matrix or Q-matrix

For a continuous MC $X_t, t \geq 0$, satisfying (1) and (2), i.e.,

$$\begin{aligned} p_{ii}(h) &= P(X_h = i \mid X_0 = i) = 1 - \lambda_i h + o(h), \\ p_{ij}(h) &= P(X_h = j \mid X_0 = i) = q_{ij} h + o(h), \quad \text{for } i \neq j \ , \end{aligned}$$

define a matrix $Q = (q_{ij})_{i,j \in S}$, where $q_{ii} = -\lambda_i$, i.e.,

$$Q = \begin{bmatrix} -\lambda_1 & q_{12} & \ldots & q_{1r} & \ldots \\ q_{21} & -\lambda_2 & \ldots & q_{2r} & \ldots \\ \vdots & \vdots & \vdots & \vdots & \\ q_{r1} & q_{r2} & \ldots & -\lambda_r & \ldots \\ \cdot & \cdot & \cdot & \cdot & \end{bmatrix}$$

Q is called **the generator matrix** or Q-matrix of $X_t, t \geq 0$, which is a fundamental quantity associated with the $X_t, t \geq 0$ containing all the information about the transition of the chain.

Recall it follows from (4) that

$$q_{ij} = \lambda_i \, p_{ij},$$

where p_{ij} is the transition probability from state i to j.

We say that q_{ij} is **the rate of going from state i to state j**. The rates q_{ij} taken all together contain more information about the process than the probabilities p_{ij} taken all together. This is because if we know all the q_{ij} we can calculate all the λ_i and then all the p_{ij}. But if we know all the p_{ij} we can not recover the q_{ij}. In many ways the q_{ij} are to continuous-time Markov chains what the p_{ij} are to discrete-time Markov chains.

There is **an important difference** between the q_{ij} in a continuous-time Markov chain and the p_{ij} in a discrete-time Markov chain. Namely, **the q_{ij} are rates, not probabilities** and, as such, while they must be nonnegative, they are not bounded by 1.

Note: if $\lambda_i = \lambda, i = 0, 1, 2, \ldots$, then $Q = \lambda \, (P - I)$, where

$$I \; = \; \begin{bmatrix} 1 & 0 & \ldots & 0 & \ldots \\ 0 & 1 & \ldots & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & \ldots & 1 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$$

Example 10.1 (continuous). The Q-matrix is

$$Q = \begin{pmatrix} -1/8 & 1/8 \\ \lambda & -\lambda \end{pmatrix}.$$

Example 10.2. Consider a Poisson process N_t with rate $\lambda > 0$. We have

$$\begin{aligned} p_{ii}(h) &= 1 - \lambda h + o(h) \\ p_{i,i+1}(h) &= \lambda h + o(h) \\ p_{i,i+j}(h) &= o(h), \quad j \geq 2 \text{ or } j < 0 \end{aligned}$$

which yields that $q_{ii} = -\lambda$, $q_{i,i+1} = \lambda$ and $q_{i,i+j} = 0, j \geq 2$ or $j < 0$. Hence

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \ldots \\ 0 & -\lambda & \lambda & 0 & .. \\ . & . & . & & \end{pmatrix}.$$

30 Forward and Backward Equations

For a continuous MC $X_t, t \geq 0$,

$$\begin{aligned} p_{ii}(h) &= P(X_h = i \mid X_0 = i) = 1 - \lambda_i h + o(h), \\ p_{ij}(h) &= P(X_h = j \mid X_0 = i) = q_{ij} h + o(h), \quad \text{for } i \neq j \ , \end{aligned}$$

Q-matrix is defined by $Q = (q_{ij})_{i,j \in S}$, where $q_{ii} = -\lambda_i$.

Using the Q-matrix, we may calculate the transition probability $p_{ij}(t)$ by the following: write

$$P(t) \ = \ \begin{bmatrix} p_{11}(t) & p_{12}(t) & ... & p_{1r}(t) & ... \\ p_{21}(t) & p_{22}(t) & ... & p_{2r}(t) & ... \\ \vdots & \vdots & \vdots & \vdots & \\ p_{r1}(t) & p_{r2}(t) & ... & p_{rr}(t) & ... \\ \cdot & \cdot & \cdot & \cdot & \end{bmatrix}$$

Forward Equations:

$$P'(t) = P(t)\, Q,$$

which is equivalent to

$$p'_{ij}(t) = \sum_{k \in S} p_{ik}(t) q_{kj}, \text{ for all } i, j \in S. \tag{5}$$

Backward Equations:

$$P'(t) = Q\, P(t),$$

which is equivalent to

$$p'_{ij}(t) = \sum_{k \in S} q_{ik}\, p_{kj}(t), \text{ for all } i, j \in S. \tag{6}$$

Proof. Using the Chapman-Kolmogorov equation, we can write

$$\begin{aligned}
p_{ij}(t+\delta) &= \sum_{k\in S} p_{ik}(t)p_{kj}(\delta) \\
&= p_{ij}(t)p_{jj}(\delta) + \sum_{k\neq j} p_{ik}(t)p_{kj}(\delta) \\
&\approx p_{ij}(t)(1+q_{jj}\delta) + \sum_{k\neq j} p_{ik}(t)\delta q_{kj} \\
&= p_{ij}(t) + \delta p_{ij}(t)q_{jj} + \delta \sum_{k\neq j} p_{ik}(t)q_{kj} \\
&= p_{ij}(t) + \delta \sum_{k\in S} p_{ik}(t)q_{kj}.
\end{aligned}$$

Thus,

$$\frac{p_{ij}(t+\delta)-p_{ij}(t)}{\delta} \approx \sum_{k\in S} p_{ik}(t)q_{kj},$$

which is the (i,j)th element of $P(t)\,Q$. The above argument can be made rigorous.

Let $P_j(t) = P(X_t = j), j \in S$, and $\widetilde{P}(t) = (P_1(t), P_2(t), ...)$.

The following **Forward Equations** for unconditional probability $P_j(t)$ are also true:

$$\widetilde{P}'(t) = \widetilde{P}(t)\, Q,$$

which is equivalent to

$$P_j'(t) = \sum_{k \in S} P_k(t) q_{kj}, \text{ for all } j \in S. \tag{7}$$

Example 10.3 Consider a continuous Markov chain with two states $S = \{0, 1\}$ and the transition matrix for any $t \geq 0$ is given by

$$P(t) = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{bmatrix}.$$

- Find the generator matrix Q and the transition matrix P for the corresponding jump chain.
- Show that for any $t \geq 0$, we have

  $$P'(t) = P(t)Q = QP(t),$$

  where $P'(t)$ is the derivative of $P(t)$.

Solution. We have

$$\begin{aligned} p_{00}(h) &= \frac{1}{2} + \frac{1}{2}(1 - 2\lambda h) + o(h) = 1 - \lambda h + o(h), \\ p_{01}(h) &= \frac{1}{2} - \frac{1}{2}(1 - 2\lambda h) + o(h) = \lambda h + o(h), \\ p_{10}(h) &= \frac{1}{2} - \frac{1}{2}(1 - 2\lambda h) + o(h) = \lambda h + o(h), \\ p_{11}(h) &= \frac{1}{2} + \frac{1}{2}(1 - 2\lambda h) + o(h) = 1 - \lambda h + o(h). \end{aligned}$$

Hence

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}.$$

The transition matrix for the corresponding jump chain P is given by

$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{if } \lambda > 0.$$

We have

$$P'(t) = \begin{bmatrix} -\lambda e^{-2\lambda t} & \lambda e^{-2\lambda t} \\ \lambda e^{-2\lambda t} & -\lambda e^{-2\lambda t} \end{bmatrix},$$

where $P'(t)$ is the derivative of $P(t)$. We also have

$$P(t)Q = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{bmatrix} \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix} = \begin{bmatrix} -\lambda e^{-2\lambda t} & \lambda e^{-2\lambda t} \\ \lambda e^{-2\lambda t} & -\lambda e^{-2\lambda t} \end{bmatrix},$$

$$QP(t) = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix} \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{bmatrix} = \begin{bmatrix} -\lambda e^{-2\lambda t} & \lambda e^{-2\lambda t} \\ \lambda e^{-2\lambda t} & -\lambda e^{-2\lambda t} \end{bmatrix}.$$

We conclude

$$P'(t) = P(t)Q = QP(t).$$