1 Continuous-time Markov chains

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\{X(t)\}_{t\geq 0}$ be a stochastic process taking values in a state space S and

$$\mathcal{F}_t = \mathcal{F}_t^X = \sigma(X(s) \,:\, 0 \leq s \leq t)$$

be its natural filtration.

Definition 1.1. $\{X(t)\}_{t\geq 0}$ is called a **continuous-time Markov chain (CTMC)** if

(1) the state space S is at most countable.

(2) The process satisfies the **Markov property**: for $s, t \geq 0$ and $i \in S$,

$$\mathbb{P}(X(t+s) = i \,|\, \mathcal{F}_t) = \mathbb{P}(X(t+s) = i \,|\, X(t))$$

To mirror the notion of the Markov property for DTMC, we see that Condition (2) is equivalent to

(2') for any $s, t \geq 0$, $n \in \mathbb{N}$, $0 \leq r_0 < \cdots < r_n \leq t$, and $i, j, x_0, \ldots, x_n \in S$,

$$\mathbb{P}\left(X(t+s) = j \middle| X(t) = i, X(r_n) = x_n, \ldots, X(r_0) = x_0 \right) = \mathbb{P}\left(X(t+s) = j \middle| X(t) = i \right).$$

Remark 1.2. The Poisson process is a CTMC. More generally, every continuous-time process that has independent increments and takes values in $\mathbb{Z}$ is a CTMC. However, recall that the Markov property does not necessarily imply independent increments (See Problem 1.1)

As in the case of a DTMC, from now on, we only consider homogeneous CTMCs unless otherwise stated. This homogeneous assumption is needed to state any results about the long time behavior of a Markov chain.

Definition 1.3. A CTMC is called **(time-)homogeneous** if, for any $s, t \geq 0$ and $i, j \in S$,

$$\mathbb{P}\left(X(t+s) = j \middle| X(t) = i \right) = \mathbb{P}\left(X(s) = j \middle| X(0) = i \right).$$

Example 1.4. The non-homogeneous Poisson process is a non-homogeneous CTMC. On the other hand, the (homogeneous) Poisson process becomes a homogeneous CTMC, if we define a **Poisson process with start in** $i \in \{0, 1, \ldots\}$ as

$$\widetilde{N}(t) := i + N(t),$$

where $N(t)$ is a Poisson process with start in $N(0) = 0$.

The key difference between a CTMC and a DTMC is that in discrete time, the Markov chain moves to a new state at times $t = 1, 2, 3, ...$ while in a CTMC, the Markov chain can move to a new state at any $t \geq 0$. Just like for the Poisson process, the times of the jumps are the **arrival** times.

1.1 The transition semigroup of a CTMC

We define the analogue of the transition matrix for DTMC, which will encode all the information needed to generate the entire CTMC.

Definition 1.5. The **transition probabilities** of a **homogeneous CTMC** are defined as

$$p_{ij}(t) = \mathbb{P}\left(X(t) = j \middle| X(0) = i \right) = \mathbb{P}\left(X(t+s) = j \middle| X(s) = i \right).$$

The **transition semigroup** is defined as

$$\boldsymbol{P}(t) = \left(p_{ij}(t)\right)_{i,j\in S}, \qquad t \geq 0,$$

with $P(0) = I$, the identity matrix. We assume from now on that

$$\boldsymbol{P}(h) \longrightarrow \boldsymbol{P}(0) = \boldsymbol{I},$$

that is, $p_{ij}(h) \to \delta_{ij}$ as $h \downarrow 0$.

Just like for the transition matrix, for each $t \geq 0$, the matrix $\boldsymbol{P}(t) = (p_{ij}(t))_{i,j\in S}$ satisfies

$$\begin{cases} p_{ij}(t) \geq 0, & \text{for all } i, j \in S, \\ \sum_{j\in S} p_{ij}(t) = 1, & \text{for all } i \in S. \end{cases}$$

The following result explains the term "semigroup".

Theorem 1.6 (Chapman–Kolmogorov equations)

For $s, t \geq 0$,

$$\boldsymbol{P}(t+s) = \boldsymbol{P}(t)\boldsymbol{P}(s) = \boldsymbol{P}(s)\boldsymbol{P}(t).$$

The transition mechanism of a CTMC has two components:

1. The time spent at a given state i before leaving it.

2. The jump mechanism by which the next state is chosen.

Knowing the answers to these questions gives us a procedure to simulate a CTMC on a computer, first by generating the time the CTMC jumps to the next state, then generating the state the CTMC jumps to. We will see in the following sections that both of this information is determined completely by the transition semigroup.

1.2 The sojourn times of a CTMC

We begin by defining a notion of how long a CTMC stays at a state before leaving it.

Definition 1.7. Given that $X(0) = i$, the **sojourn time** of the CTMC at state i is the random time U_i defined as

$$U_i = \inf \{t \geq 0 : X(t) \neq i\}.$$

The sojourn time for the Poisson process is exponential. We will see that this holds more generally.

Remark 1.8. Note that $U_i > 0$ can only happen if $X(0) = i$. Therefore, we will consider U_i only under the probability measure $\mathbb{P}(\,\cdot\,|X(0) = i)$.

Proposition 1.9

Under $\mathbb{P}(\,\cdot\,|X(0) = i)$, the sojourn time U_i has an exponential distribution with rates α_i that can depend on the state i.

This result says that the time to leave a state is an exponential random variable. The natural question is what is the value of the parameter in the exponential distribution of U_i? We will answer these questions in the next sections, but we start by giving some brief motivation about how this information can be encoded by the transition semigroup.

We will see how the small time behavior of a CTMC is the key to this information. Notice that if U_i is exponential with rate α_i, then it satisfies

$$\begin{aligned} \mathbb{P}(U_i > \triangle t) = e^{-\alpha_i \triangle t} &= 1 - \alpha_i \triangle t + \frac{1}{2}(\alpha_i \triangle t)^2 - \frac{1}{3!}(\alpha_i \triangle t)^3 + \cdots \\ &= 1 - \alpha_i \triangle t + o(\triangle t) \end{aligned}$$

where the error term $o(\triangle t)$ is much smaller than $\triangle t$ for $\triangle t$ small,

$$\lim_{\triangle t \to 0} \frac{o(\triangle t)}{\triangle t} = 0.$$

That is,

$$\mathbb{P}(\text{no jumps occurred by time } \triangle t) = 1 - \alpha_i \triangle t + o(\triangle t).$$

Rearranging this equation, and using the fact that $p_{ii}(\triangle t)$ gives the result on the left hand side and $p_{ii}(0) = 1$ implies that

$$-\alpha_i = \lim_{\triangle t \to 0} \frac{p_{ii}(\triangle t) - p_{ii}(0)}{\triangle t} = p'_{ii}(0).$$

Therefore, the derivatives or rates of change of the transition semigroup contains the relevant information about the rates of the exponential distribution in the sojourn times.

Remark 1.10. Furthermore, by taking complements of the set $\{\text{no transition occurred by time } \triangle t\}$ we see that

$$\begin{aligned}
&\mathbb{P}(\text{one jump occurred by time } \triangle t) = \alpha_i \triangle t + o(\triangle t),\\
&\mathbb{P}(\text{at least two jumps occurred by time } \triangle t) = o(\triangle t).
\end{aligned}$$

The first implies that the rates of jumps α_i is proportional to the time interval. The second also removes the probability that two jumps happen simultaneously. Both of these facts mirror the Poisson process.

1.2.1 Infinitesimal generator matrix

We answer the first question in Section 1.1, namely what is the rate of the exponential clock that determines when the CTMC leaves its current state. From now on, we assume that

$$t \longmapsto p_{ij}(t) \quad \text{is differentiable at } t = 0 \text{ for all } i, j \in S.$$

This allows us to define a matrix that encodes the instantaneous rate at which a Markov chain transitions between states.

Definition 1.11. The matrix $\boldsymbol{Q} := (q_{ij})_{i,j\in S}$ with entries

$$q_{ij} = \frac{\mathrm{d}}{\mathrm{d}t} p_{ij}(t)\bigg|_{t=0} = \lim_{h \downarrow 0} \frac{p_{ij}(h) - \delta_{ij}}{h} = p'_{ij}(0). \tag{1}$$

is called the **infinitesimal generator** of the CTMC.

Remark 1.12. In matrix form, the matrix $\boldsymbol{Q}$ can be written as

$$\boldsymbol{Q} := (q_{ij})_{i,j\in S} = \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{P}(t)\bigg|_{t=0} = \lim_{h \downarrow 0} \frac{\boldsymbol{P}(h) - \boldsymbol{I}}{h} = \boldsymbol{P}'(0).$$

since $\boldsymbol{P}(0) = \boldsymbol{I}$ the identity matrix.

By rearranging the definition of $\boldsymbol{Q}$, we see that for $h = \triangle t$, we have that for $i \neq j$

$$p_{ij}(\triangle t) = (\triangle t) \times q_{ij}$$

so the probability that the chain moves from i to j in a short time is proportional to its rate q_{ij}. The values Q-matrix will encode the rates at which the Markov chain move from i to j. This is made precise with the following result.

Proposition 1.13

1. *The diagonal elements of* $\boldsymbol{Q}$ *are*

$$q_{ii} = -\alpha_i, \qquad i \in S,$$

where α_i *is the parameter of the exponential distribution of the sojourn time at state* i.

2. *The off-diagonal elements of* $\boldsymbol{Q}$ *satisfy*

$$q_{ij} \geq 0 \qquad \text{for } i \neq j,$$

and

$$\sum_{j \neq i} q_{ij} = \alpha_i \qquad \text{for all } i \in S.$$

This proposition allows us to read the parameters of the exponential distributions fo the sojourn times from the diagonal elements of the Q-matrix.

1.3 Embedded DTMC

We answer the second question in Section 1.1, namely when the CTMC leaves its current state, what is the probability it goes to each state. One might expect that this information will be encoded by a DTMC.

Definition 1.14. Consider the transition matrix $\widetilde{\boldsymbol{P}} := (\tilde{p}_{ij})_{i,j \in S}$ with entries

$$\widetilde{p}_{ij} := \begin{cases} \dfrac{q_{ij}}{-q_{ii}} & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases}$$

The DTMC with transition matrix $\widetilde{\boldsymbol{P}}$ is called the **embedded DTMC** of the CTMC.

Remark 1.15. The second part of Proposition 1.13 implies that

$$\begin{aligned} \widetilde{p}_{ij} &\geq 0 \qquad \text{for } i \neq j, \\ \sum_j \widetilde{p}_{ij} &= 1 \qquad \text{for all } i \in S, \end{aligned}$$

so $\widetilde{\boldsymbol{P}}$ is a stochastic matrix, and therefore a valid transition matrix.

To see why $\widetilde{\boldsymbol{P}}$ is the transition matrix of a DTMC, we can simply define $X_n = X(T_n)$, where T_n denotes the time of the nth jump, to be the location of the DTMC at the time nth jump. Since the Markov chain is homogeneous, we will see in the following result that the values $\tilde{p}_{ij}$ is the probability that the Markov chain jumps to state j given that it started at state i and a jump just occurred.

Proposition 1.16

For $i \neq j$, *we have*

$$\mathbb{P}(X(U_i) = j | X(0) = i) = \tilde{p}_{ij}, \qquad j \neq i.$$

The intuition of the above description of the CTMC is that $X(t)$ stays at a state i for a random time period U_i with distribution $\operatorname{Exp}(\alpha_i)$ and after that moves on to the next state j, which is chosen according to the transition matrix $\widetilde{\boldsymbol{P}}$.

1.4 The forward and backward Kolmogorov equations

We have seen that the $\boldsymbol{Q}$ matrix is easily computable given $\boldsymbol{P}$ and the $\boldsymbol{Q}$ matrix encodes all the information needed to generate the CTMC. In this section, we will see that the $\boldsymbol{P}$ matrix can also be computed from the $\boldsymbol{Q}$ matrix.

Proposition 1.17

The time derivative $\boldsymbol{P}'(t)$ of the transition semigroup satisfies the following two equations, where $\boldsymbol{Q}$ is the infinitesimal generator of the CTMC.

- ***Kolmogorov backward equation***

$$\boldsymbol{P}'(t) = \boldsymbol{Q}\boldsymbol{P}(t) \tag{2}$$

- ***Kolmogorov forward equation***

$$\boldsymbol{P}'(t) = \boldsymbol{P}(t)\boldsymbol{Q} \tag{3}$$

These equations give a statement about the rate of change in $\boldsymbol{P}$ in terms of itself. These equations remind us of the growth equations in ODEs, which states that

$$f'(t) = kf(t) \implies f(t) = f(0)e^{tf(t)}.$$

A similar result holds for matrices.

Theorem 1.18

Under some technical conditions to ensure that the terms below are well defined, the Kolmogorov forward and backward equations have the following solution, subject to the initial condition $\boldsymbol{P}(0) = \boldsymbol{I}$,

$$\boldsymbol{P}(t) = e^{t\boldsymbol{Q}}.$$

Where, $e^{\boldsymbol{M}}$ is the **matrix exponential** *of the square matrix $\boldsymbol{M}$, defined by*

$$e^{\boldsymbol{M}} = \sum_{n=0}^{\infty} \frac{1}{n!}\boldsymbol{M}^n.$$

Even when the state space is finite, i.e., $S = \{0, 1, 2, \ldots, N\}$, the matrix exponential $\boldsymbol{P}(t) = e^{t\boldsymbol{Q}}$ is still not easy to calculate unless $\boldsymbol{Q}$ is diagonalizable, that is, there exists an invertible matrix $\boldsymbol{A}$ such that

$$\boldsymbol{Q} = \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^{-1},$$

where

$$\boldsymbol{D} = \begin{pmatrix} d_0 & & & \\ & d_1 & & \\ & & \ddots & \\ & & & d_N \end{pmatrix}.$$

In this case,

$$\boldsymbol{Q}^2 = (\boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^{-1})^2 = \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^{-1} = \boldsymbol{A}\boldsymbol{D}\boldsymbol{I}\boldsymbol{D}\boldsymbol{A}^{-1} = \boldsymbol{A}\boldsymbol{D}^2\boldsymbol{A}^{-1}$$

and, in the same way,

$$\boldsymbol{Q}^k = (\boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^{-1})^k = \boldsymbol{A}\boldsymbol{D}^k\boldsymbol{A}^{-1}.$$

Since $\boldsymbol{D}^k$ is also a diagonal matrix, we have

$$\boldsymbol{P}(t) = e^{t\boldsymbol{Q}} = \sum_{n=0}^{\infty} \frac{t^n \boldsymbol{Q}^n}{n!} = \sum_{n=0}^{\infty} \frac{t^n \left(\boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^{-1}\right)^n}{n!} = \sum_{n=0}^{\infty} \frac{t^n \boldsymbol{A}\boldsymbol{D}^n\boldsymbol{A}^{-1}}{n!} = \boldsymbol{A}\left(\sum_{n=0}^{\infty} \frac{t^n \boldsymbol{D}^n}{n!}\right)\boldsymbol{A}^{-1} = \boldsymbol{A}e^{t\boldsymbol{D}}\boldsymbol{A}^{-1},$$

where

$$e^{t\boldsymbol{D}} = \begin{pmatrix} e^{td_0} & & & \\ & e^{td_1} & & \\ & & \ddots & \\ & & & e^{td_N} \end{pmatrix}.$$

1.5 Example Problems

1.5.1 Proofs of Results

Problem 1.1. Give an example of stochastic process that satisfies the Markov property, but does not have independent increments.

Solution 1.1. We first provide a discrete time example. Let $(\xi_n)_{n\geq 1}$ be independent Rademacher random variables,

$$\mathbb{P}(\xi = \pm 1) = \frac{1}{2}.$$

Notice that $(\xi_n)_{n\geq 0}$ is Markov because

$$\mathbb{P}(\xi_n = x_n \mid \xi_{n-1} = x_{n-1}, \dots, \xi_0 = x_0) = \mathbb{P}(\xi_n = x_n) = \mathbb{P}(\xi_n = x_n \mid \xi_{n-1} = x_{n-1})$$

but clearly the increments $X_n = \xi_n - \xi_{n-1}$ and $X_{n-1} = \xi_{n-1} - \xi_{n-2}$ are not independent because they depend on the same random variables. A continuous time version of this example can be constructed by consider the same process $(\xi_t)_{t\geq 0}$ but indexed by time.

Problem 1.2. Prove the Chapman–Kolmogorov Equations.

Solution 1.2. The proof is identical to the discrete case. By Markov property and homogeneity, we have by the chain rule for conditional probabilities

$$\begin{aligned}
p_{ij}(t+s) &= \mathbb{P}\left(X(t+s) = j \mid X(0) = i\right) \\
&= \sum_{k\in S} \mathbb{P}\left(X(t+s) = j \mid X(t) = k, X(0) = i\right) \mathbb{P}\left(X(t) = k \mid X(0) = i\right) \\
&= \sum_{k\in S} \mathbb{P}\left(X(t+s) = j \mid X(t) = k\right) \mathbb{P}\left(X(t) = k \mid X(0) = i\right) \\
&= \sum_{k\in S} \mathbb{P}\left(X(s) = j \mid X(0) = k\right) \mathbb{P}\left(X(t) = k \mid X(0) = i\right) \\
&= \sum_{k\in S} p_{ik}(t) p_{kj}(s).
\end{aligned}$$

Writing this identity in matrix notation finishes the proof.

Problem 1.3. Prove Proposition 1.9.

Solution 1.3. Recall that a random variable is memoryless if

$$\mathbb{P}(X > t + s \mid X > s) = \mathbb{P}(X > t)$$

which roughly says that previous elapsed time does not affect the future waiting time. The exponential distribution is the only continuous distribution with the memoryless property.

At an intuitive level, the memoryless property seems to share many connections with the Markov property in the sense that past information has a weak effect. We make this precise by considering the conditional probability

$$\begin{aligned}
\mathbb{P}\left(U_i > t+s \middle| U_i > s,\, X(0) = i\right) &= \mathbb{P}\left(X(u) = i \text{ for } u \in [0, t+s] \middle| X(u) = i \text{ for } u \in [0, s]\right) \\
&= \mathbb{P}\left(X(u) = i \text{ for } u \in (s, t+s] \middle| X(u) = i \text{ for } u \in [0, s]\right) \\
&= \mathbb{P}\left(X(u) = i \text{ for } u \in (s, t+s] \middle| X(s) = i\right), \text{ Markov property} \\
&= \mathbb{P}\left(X(u) = i \text{ for } u \in (0, t] \middle| X(0) = i\right), \text{ homogeneous} \\
&= \mathbb{P}\left(U_i > t | X(0) = i\right), \text{ definition}
\end{aligned}$$

Therefore, the distribution of the sojourn time U_i is **memoryless**. Since the exponential distribution is the only continuous distribution with the memoryless property, U_i follows an exponential distribution with some parameter α_i.

Problem 1.4. Prove Proposition 1.13.

Solution 1.4. This proof formalizes the computations at the end of Section 1.1.

Part 1: We have as $h \downarrow 0$, the law of total probability implies that

$$p_{ii}(h) = \mathbb{P}\left(X(h) = i \middle| X(0) = i\right) = \mathbb{P}\left(U_i > h | X(0) = i\right) + \mathbb{P}\left(X(h) = i, U_i < h \middle| X(0) = i\right).$$

Notice that

$$\begin{aligned}
\mathbb{P}\left(X(h) = i, U_i < h \middle| X(0) = i\right) &\le \mathbb{P}\left(\text{at least two transitions occurred by time } h\right) \\
&= o(h).
\end{aligned}$$

Therefore,

$$p_{ii}(h) = \mathbb{P}\left(U_i > h | X(0) = i\right) + o(h) = e^{-\alpha_i h} + o(h)$$

which implies

$$q_{ii} = \lim_{h \downarrow 0} \frac{p_{ii}(h) - \delta_{ii}}{h} = \lim_{h \downarrow 0} \frac{e^{-\alpha_i h} + o(h) - 1}{h} = -\alpha_i.$$

Part 2: For $i \neq j$, we have

$$q_{ij} = \lim_{h \downarrow 0} \frac{p_{ij}(h) - 0}{h}$$

Since $p_{ij}(h)/h \ge 0$ for all $h > 0$, so must be q_{ij}. Next, we have

$$\sum_{j \neq i} p_{ij}(h) = 1 - p_{ii}(h),$$

and therefore

$$\sum_{j \neq i} q_{ij} = -p_{ii}'(0) = \alpha_i.$$

(The preceding argument is correct if S is finite and needs some additional care if S is infinite, because then the interchange of limit and an infinite sum needs extra justification).

Problem 1.5. Prove Proposition 1.16

Solution 1.5. For $i \neq j$, we define

$$r_{ij}(\triangle t) = \mathbb{P}(X(t+\triangle t) = j \mid X(t) = i, X(t+\triangle t) \neq i) = \mathbb{P}(X(h) = j \mid X(0) = i, X(\triangle t) \neq i)$$

by time homogeneity. When $\triangle t$ is very small, there will only be one transition with very high probability, so it roughly represents the probability that the chain jumps to state j from state i given that there is a transition at time $\triangle t$. We have

$$\mathbb{P}(X(U_i) = j | X(0) = i) = \lim_{\triangle t \to 0} R_{ij}(\triangle t).$$

By the definition of conditional probability,

$$r_{ij}(\triangle t) = \frac{\mathbb{P}(X(\triangle t) = j \mid X(0) = i)}{\mathbb{P}(X(\triangle t) \neq i \mid X(0) = i)} = \frac{p_{ij}(\triangle t)}{1 - p_{ii}(\triangle t)} = \frac{\frac{p_{ij}(\triangle t) - 0}{\triangle t}}{-\frac{p_{ii}(\triangle t) - 1(\triangle t)}{\triangle t}}$$

so taking $\triangle t \to 0$ implies that

$$\lim_{\triangle t \to 0} R_{ij}(\triangle t) = \frac{p'_{ij}(0)}{-p'_{ii}(0)} = \frac{q_{ij}}{-q_{ii}},$$

by the definition of the Q-matrix.

Problem 1.6. Prove the forward and backward equations in Proposition 1.17.

Solution 1.6.

Forward Equations: The Chapman-Kolmogorov equation implies

$$\boldsymbol{P}(t+h) - \boldsymbol{P}(t) = \boldsymbol{P}(h)\boldsymbol{P}(t) - \boldsymbol{P}(t) = (\boldsymbol{P}(h) - \boldsymbol{I})\,\boldsymbol{P}(t) = (\boldsymbol{P}(h) - \boldsymbol{P}(0))\,\boldsymbol{P}(t).$$

Thus,

$$\boldsymbol{P}'(t) = \lim_{h \downarrow 0} \frac{\boldsymbol{P}(t+h) - \boldsymbol{P}(t)}{h} = \lim_{h \downarrow 0} \frac{\boldsymbol{P}(h) - \boldsymbol{P}(0)}{h} \boldsymbol{P}(t) = \boldsymbol{Q}\boldsymbol{P}(t).$$

Backward Equations: Similarly, since

$$\boldsymbol{P}(t+h) - \boldsymbol{P}(t) = \boldsymbol{P}(t)\,(\boldsymbol{P}(h) - \boldsymbol{P}(0))\,,$$

we have

$$\begin{aligned} \boldsymbol{P}'(t) &= \lim_{h \downarrow 0} \frac{\boldsymbol{P}(t+h) - \boldsymbol{P}(t)}{h} = \lim_{h \downarrow 0} \boldsymbol{P}(t) \frac{\boldsymbol{P}(h) - \boldsymbol{P}(0)}{h} \\ &= \boldsymbol{P}(t) \lim_{h \downarrow 0} \frac{\boldsymbol{P}(h) - \boldsymbol{P}(0)}{h} \qquad (4) \\ &= \boldsymbol{P}(t)\boldsymbol{Q}. \end{aligned}$$

Remark 1.19. If the state space S is **infinite**, the step (4) interchanges limit and infinite sum, which requires additional assumptions. For a **finite** state space S, this is always okay. No such interchange is needed for the derivation of (2).

Problem 1.7. Prove the matrix exponential satisfies the Kolmogorov equations as in Theorem 1.18

Solution 1.7. We first have $e^{0\cdot \boldsymbol{Q}} = \boldsymbol{Q}^0 = \boldsymbol{I} = \boldsymbol{P}(0)$. Next, assuming enough conditions on $\boldsymbol{P}$ so that we may interchange differentiation and summation,

$$\begin{aligned}\boldsymbol{P}'(t) = \frac{\mathrm{d}}{\mathrm{d}t} e^{t\cdot \boldsymbol{Q}} = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{n=0}^{\infty} \frac{t^n}{n!} \boldsymbol{Q}^n = \sum_{n=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}t} \frac{t^n}{n!} \boldsymbol{Q}^n = \sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} \boldsymbol{Q}^n \\ = \boldsymbol{Q} \left(\sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} \boldsymbol{Q}^{n-1} \right) = \boldsymbol{Q} \left(\sum_{m=0}^{\infty} \frac{t^m}{m!} \boldsymbol{Q}^m \right) = \boldsymbol{Q}\boldsymbol{P}(t) \\ = \left(\sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} \boldsymbol{Q}^{n-1} \right) \boldsymbol{Q} = \left(\sum_{m=0}^{\infty} \frac{t^m}{m!} \boldsymbol{Q}^m \right) \boldsymbol{Q} = \boldsymbol{P}(t)\boldsymbol{Q}.\end{aligned}$$

Note that we use the fact $\boldsymbol{Q}^n = \boldsymbol{Q}\boldsymbol{Q}^{n-1} = \boldsymbol{Q}^{n-1}\boldsymbol{Q}$ above.

1.5.2 Applications

Problem 1.8. What is the transition semigroup of a Poisson process with intensity $\lambda > 0$?

Solution 1.8. Recall that a Poisson process with intensity $\lambda > 0$ satisfies

$$\mathbb{P}\left(N(t+h) - N(t) = n\right) = e^{-\lambda h} \frac{(\lambda h)^n}{n!}, \quad n = 0, 1, 2, \ldots$$

Therefore,

$$p_{ij}(t) = \mathbb{P}(N(t) = j \mid N(0) = i) = \begin{cases} e^{-\lambda h} \frac{(\lambda h)^{j-i}}{(j-i)!} & j \geq i \\ 0 & j < i. \end{cases}$$

Problem 1.9. Find the infinitesimal generator of a Poisson process with intensity λ.

Solution 1.9. For the Poisson process with intensity $\lambda > 0$, we have

$$p_{ij}(t) = \begin{cases} e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!} & \text{if } j \geq i, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, for $j = i$,

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{ii}(t) \bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} e^{-\lambda t} \bigg|_{t=0} = -\lambda,$$

for $j = i + 1$,

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{i,i+1}(t) \bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} e^{-\lambda t} \lambda t \bigg|_{t=0} = \lambda,$$

and for $j \geq i + 2$,

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{ij}(t) \bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!} \bigg|_{t=0} = 0$$

Therefore, condition (1) holds and the matrix $(q_{ij})_{i,j=0,1,\ldots}$ looks like this:

$$\boldsymbol{Q} = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \ddots \\ 0 & 0 & 0 & -\lambda & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Problem 1.10. Find the transition matrix of the embedded DTMC associated with the Poisson process with intensity λ.

Solution 1.10. The entries can be computed using the formula and the formula for the Q-matrix in Problem 1.9

$$\tilde{p}_{n,n+1} = \frac{q_{n,n+1}}{-q_{nn}} = \frac{\lambda}{-(-\lambda)} = 1.$$

Since the sums along the rows must be 1, we have that all other entries are zero (which can be seen by also applying the formulas for to the other entries).

Remark 1.20. This is a very intuitive result. For instance, if $T_1, T_2, \ldots,$ are the arrival times, then we define $X_n = X(T_n)$, then we have that

$$\tilde{p}_{n,n+1} = 1$$

and $\tilde{p}_{n,j} = 0$ for all $j \neq n+1$, since we know that the Poisson process increases by 1 at each arrival time. The matrix is given below

$$\widetilde{\boldsymbol{P}} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \ddots \\ 0 & 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Problem 1.11. Consider a CTMC with state space $\{0, 1\}$ and generator

$$\boldsymbol{Q} = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}.$$

What is probability transition matrix $\boldsymbol{P}(t)$ corresponding to $\boldsymbol{Q}$?

Solution 1.11. For this CTMC, the probability transition matrix $\boldsymbol{P}(t)$ can be obtained explicitly. By the Kolmogorov backward equation $\boldsymbol{P}'(t) = \boldsymbol{Q}\boldsymbol{P}(t)$, we have

$$\begin{cases} p'_{00}(t) = -\alpha p_{00}(t) + \alpha p_{10}(t) \\ p'_{10}(t) = \beta p_{00}(t) - \beta p_{10}(t) \end{cases} \tag{5}$$

It suffices to compute these entries because $p_{01} = 1 - p_{00}$ and $p_{11} = 1 - p_{10}$. This is a system of coupled ordinary differential equations (ODEs). Multiplying the first equation by β, the second equation by α, and adding these two equations, we obtain

$$\beta p'_{00}(t) + \alpha p'_{10}(t) = 0.$$

Integrating this equation with the initial conditions $p_{00}(0) = 1$ and $p_{10}(0) = 0$, we have

$$\beta p_{00}(t) = \beta - \alpha p_{10}(t). \tag{6}$$

Substituting this into the second equation of (5) yields an ODE for p_{10},

$$p'_{10}(t) = -(\alpha + \beta) p_{10}(t) + \beta.$$

Multiplying both sides of the equation with $e^{(\alpha+\beta)t}$ yields

$$\left(e^{(\alpha+\beta)t} p_{10}(t)\right)' = \beta e^{(\alpha+\beta)t}.$$

Using the boundary condition $p_{10}(0) = 0$ again, we have

$$p_{10}(t) = e^{-(\alpha+\beta)t} \int_0^t \beta e^{(\alpha+\beta)s} \mathrm{d}s = \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta} e^{-(\alpha+\beta)t}.$$

Further, from (6),

$$p_{00}(t) = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t}.$$

Finally, we note that

$$p_{01}(t) = 1 - p_{00}(t) = 1 - \frac{\beta}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t} = \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t}$$

and

$$p_{11}(t) = 1 - p_{10}(t) = 1 - \frac{\beta}{\alpha+\beta} + \frac{\beta}{\alpha+\beta} e^{-(\alpha+\beta)t} = \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta} e^{-(\alpha+\beta)t}.$$

In summary,

$$\boldsymbol{P}(t) = \begin{pmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t} & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t} \\ \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta} e^{-(\alpha+\beta)t} & \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta} e^{-(\alpha+\beta)t} \end{pmatrix}.$$