II Applied Probability

Ishan Nath, Lent 2023

Based on Lectures by Dr. Sourav Sarkar

February 22, 2023

Contents

1   Continuous-time Markov Chains

A sequence of random variables is called a *stochastic process*. The process $X = (X_n)_{n \in \mathbb{N}}$ is called a *discrete-time Markov chain* with state space I if for all $x_0, x_1, \ldots, x_n \in I$,

$$\mathbb{P}(X_n = x_n \mid X_{n-1} = x_{n-1}, \ldots, X_1 = x_1, X_0 = x_0) = \mathbb{P}(X_n = x_n \mid X_{n-1} = x_{n-1}).$$

It is *time-homogeneous* if

$$\mathbb{P}(X_{n+1} = y \mid X_n = x)$$

is independent of n. If it is time-homogeneous, we can write $P = (P_{x,y})_{x,y \in I}$ for the *transition matrix*, where

$$P_{x,y} = \mathbb{P}(X_1 = y \mid X_0 = x) = \mathbb{P}(X_{n+1} = y \mid X_0 = x).$$

The data associated to every time-homogeneous Markov chain is the transition matrix P and the initial distribution μ, where $\mathbb{P}(X_0 = x_0) = \mu(x_0)$ for all $x_0 \in I$.

From now on, we let I denote a countable state space, and $(\Omega, \mathcal{F}, \mathbb{P})$ denote the probability space on which all relevant random variables are defined.

Definition 1.1. $X = (X(t) \mid t \geq 0)$ is a (right-continuous) *continuous time random process* with values in I if

(a) For all $t \geq 0$, $X(t)$ (or X_t) is a random variable such that $X(t) : \Omega \to I$.

(b) For all $w \in \Omega$, $t \mapsto X_t(\omega)$ is right-continuous.

In our case, since I is countable, this means for all $w \in \Omega$, and for all $t \geq 0$, there exists $\varepsilon > 0$ (dependent of w and t) such that $X_t(w) = X_s(w)$, for all $s \in [t, t + \varepsilon]$.

Proposition 1.1. *A right-continuous random process is determined by its finite dimensional distribution*

$$\mathbb{P}(X_{t_0} = i_0, X_{t_i} = i_1, \ldots, X_{t_n} = i_n).$$

For every $w \in \Omega$, the path $t \mapsto X_t(\omega)$ of a right-continuous process is always constant for a while. There are three possibilities for its behaviour:

(i) The path may make infinitely many jumps, but only finitely many in any interval $[0, t]$.

(ii) The path makes finitely many jumps and gets absorbed in some state.

(iii) The path makes infinitely many jumps in a finite time interval. After the explosion time, the process starts up again.

We denote $J_0 = 0, J_1, J_2, \ldots$ as the *jump times*, and $S_1, S_2, \ldots$ the *holding times*. We can define J_i as $J_0 = 0$, and

$$J_{n+1} = \inf\{t \geq J_n \mid X_t \neq X_{J_n}\}.$$

Then we define

$$S_n = \begin{cases} J_n - J_{n-1} & J_n \neq \infty, \\ \infty & \text{otherwise.} \end{cases}$$

By right-continuity, $S_n > 0$ for all n. If $J_{n+1} = \infty$ for some n, then we define $X_\infty = X_{J_n}$ as the final value. Otherwise it is not defined.

The *explosion time* S is defined by

$$S = \sup_n J_n = \sum_{n=1}^{\infty} S_n.$$

We are not going to consider what happens to a chain after explosion. Instead, we will set $X_t = \infty$ for all $t > S$.

Definition 1.2. A right-continuous, continuous time random process $X = (X_t)_{t \geq 0}$ has the *Markov property* and is called a *continuous time Markov chain* if for all $i_1, \ldots, i_n \in I$ and $0 < t_1 < \cdots < t_n$,

$$\mathbb{P}(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_1} = i_1) = \mathbb{P}(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1}).$$

For all $k > 0$, $X_n = X(k_n)$ defines a discrete Markov chain.

Definition 1.3. We define the *jump chain* of $(X_t)_{t \geq 0}$ by setting $Y_n = X_{J_n}$ for all n.

Definition 1.4. The transition probabilities are defined as

$$P_{ij}(s, t) = \mathbb{P}(X_t = j \mid X_s = i).$$

The continuous times Markov chain is called *time-homogeneous* if the transition probabilities depend only on $t - s$, or $P_{ij}(s, t) = P_{ij}(0, t - s)$.

As in the discrete time case, a (time-homogeneous) Markov chain is characterised by its initial distribution λ, and its family of transition matrices

$$(P(t))_{t \geq 0} = (P_{ij}(t))_{\substack{i,j \\ t \geq 0}}.$$

These form a semigroup (a group without inverses).

It is easy to see that:

- $P(0)$ is the identity.

- $P(t)$ is a stochastic matrix for all t.

- $P(t+s) = P(t)P(s)$ for all s, t (Chapman-Kolmogorov equation).

To prove this, note

$$
\begin{aligned}
P_{xz}(t+s) &= \mathbb{P}(X_{t+s} = z \mid X_0 = x) \\
&= \sum_{y \in I} \mathbb{P}(X_{t+s} = z \mid X_0 = x, X_t = y)\mathbb{P}(X_t = y \mid X_0 = x) \\
&= \sum_{y \in I} \mathbb{P}(X_{t+s} = z \mid X_t = y)\mathbb{P}(X_t = y \mid X_0 = x) \\
&= \sum_{y \in I} \mathbb{P}(X_s = z \mid X_0 = y)\mathbb{P}(X_t = y \mid X_0 = x) \\
&= \sum_{y \in I} P_{yz}(s)P_{xy}(t) = P_x(t)P_z(s).
\end{aligned}
$$

1.1   Holding Times

Let X be a (right-cont. cont. time time-homogeneous) Markov chain on a countable state-space I, and suppose it starts from $x \in I$. Then how long does it stay in the state x?

We call S_x the *holding time* at state x. Since X is right continuous, $S_x > 0$. Let $s, t \geq 0$. Then,

$$
\begin{aligned}
\mathbb{P}(S_x > t + s \mid S_x > s) &= \mathbb{P}(X_u = x \,\forall u \in [0, t+s] \mid X_u = x \,\forall u \in [0, s]) \\
&= \mathbb{P}(X_u = x \,\forall u \in [s, t+s] \mid X_u = x \,\forall u \in [0, s]) \\
&= \mathbb{P}(X_u = x \,\forall u \in [s, t+s] \mid X_s = x) \\
&= \mathbb{P}(X_u = x \,\forall u \in [0, t] \mid X_0 = x) = \mathbb{P}(S_x > t).
\end{aligned}
$$

Thus S_x has the *memoryless property*. The same holds for the holding time at any state.

By the next theorem, we will get that S_x has the exponential distribution, say with parameter q_x.

Theorem 1.1 (Memoryless Property). *Let S be a positive random variable. Then S has the* memoryless property, *i.e.* $\mathbb{P}(S > t + s \mid S > s) = \mathbb{P}(S > t)$ *for all $s, t \geq 0$ if and only if S has the exponential distribution.*

Recall that $X \sim \text{Exp}(\lambda)$ if $\mathbb{P}(X > x) = e^{-\lambda x}$, and $X \geq 0$.

Proof: We show a memoryless random variable is exponential. Set $F(t) = \mathbb{P}(S > t)$. Then $F(s + t) = F(s)F(t)$ for all $s, t \geq 0$. Since S is a positive random variable, there exists $n \in \mathbb{N}$ large enough such that

$$F(1/n) = \mathbb{P}(S > 1/n) > 0.$$

Then, we get $F(1) = F(1/n + 1/n + \cdots + 1/n) = (F(1/n))^n > 0$. So we can set $F(1) = e^{-\lambda}$, for some $\lambda \geq 0$.

For $k \in \mathbb{N}$, we get $F(k) = F(1 + \cdots + 1) = (F(1))^k = e^{-\lambda k}$. Then, for all rationals, we get

$$F(p/q) = (F(1/q))^p = [(F(1/q))^q]^{p/q} = [F(1/q + \cdots + 1/q)]^{p/q} = e^{-\lambda p/q}.$$

For any $t \geq 0$, and for any $r, s \in \mathbb{Q}$ such that $r \leq t \leq s$, since F is decreasing,

$$e^{-\lambda s} = F(s) \leq F(t) \leq F(r) = e^{-\lambda r}.$$

Letting $r, s \to t$, we get $F(t) = e^{-\lambda t}$.

2   Poisson Process

We are now going to look at the simplest (and most important) example of a continuous time Markov chain, the *Poisson Process*.

Suppose $S_1, S_2, \ldots$ are iid random variables, with $S_1 \sim \text{Exp}(\lambda)$. Define the jump times to be $J_0 = 0$, $J_1 = S_1$, and in general, $J_n = S_1 + \cdots + S_n$, for all n.

We set $X_t = i$ if $J_i \leq t < J_{i+1}$, for all $i = 0, 1, 2, \ldots$. Here, $I = \{0, 1, 2, \ldots\}$, and note the Poisson process is right-continuous and increasing.

Then X is called a Poisson process of parameter/intensity λ, which we denote $PP(\lambda)$. We sometimes refer to the jump times (J_i) as the *points* of the Poisson process. Then X_t counts the number of points in $[0, t]$.

Theorem 2.1 (Markov Property). *Let $(X_t)_{t \geq 0}$ be $PP(\lambda)$. Then for all $s \geq 0$, the process $(X_{s+t} - X_s)_{t \geq 0}$ is also $PP(\lambda)$, and is independent of $(X_\gamma)_{0 \leq \gamma \leq s}$.*

> **Proof:**   Set $Y_t = X_{t+s} - X_s$ for all $t \geq 0$, and let $i = \{0, 1, 2, \ldots\}$. We condition on $\{X_s = i\}$. Then the jump times for the process Y are $J_{i+1} - s, J_{i+2} - s, \ldots$, and the holding times are
>
> $$T_1 = J_{i+1} - s = S_{i+1} - (s - J_i),$$
>
> and $T_2 = S_{i+2}, T_3 = S_{i+3}$, and so on. Here J and S are the jump times and holding times for X. Since
>
> $$\begin{aligned} \{X_s = i\} &= \{J_i \leq s\} \cap \{s < J_{i+1}\} \\ &= \{J_i \leq s\} \cap \{S_{i+1} > s - J_i\}, \end{aligned}$$
>
> conditional on $\{X_s = i\}$, by the memoryless property of the exponential distribution, we can see that $T_1 \sim \text{Exp}(\lambda)$. Moreover, the times T_j for $j \geq 2$ are independent of S_k for $k \leq i + 1$, and hence independent of $(X_\gamma)_{\gamma \leq s}$, and they have iid $\text{Exp}(\lambda)$ distribution.
>
> Thus $(X_{s+t} - X_s)_{t \geq 0} \sim PP(\lambda) \perp (X_\gamma)_{0 \leq \gamma \leq s}$.

Similarly, we can show the *Strong Markov property* for $PP(\lambda)$. Recall that a random variable $T \in [0, \infty]$ is called a *stopping time* if for all t, the event $\{T \leq t\}$ depends only on $(X_s)_{s \leq t}$.

Theorem 2.2 (Strong Markov Property). *Let $(X_t)_{t \geq 0}$ be $PP(\lambda)$, and T a stopping time. Then conditional on $T < \infty$, the process $(X_{T+t} - X_T)_{t \geq 0}$ is $PP(\lambda)$ and independent of $(X_s)_{s \leq T}$.*

Theorem 2.3. *Let* $(X_t)_{t\geq 0}$ *be an increasing right continuous process taking values in* $\{0,1,2,\ldots\}$ *with* $X_0 = 0$. *Let* $\lambda > 0$. *Then the following are equivalent:*

(a) *The holding times* $S_1, S_2, \ldots$ *are iid* $\mathrm{Exp}(\lambda)$, *and the jump chain is given by* $Y_n = n$.

(b) *(Infinitesimal definition)* X *has independent increments, and as* $h \downarrow 0$ *uniformly in* t, *we have*

$$\mathbb{P}(X_{t+h} - X_t = 1) = \lambda h + o(h),$$
$$\mathbb{P}(X_{t+h} - X_t = 0) = 1 - \lambda h + o(h).$$

This means that $\mathbb{P}(X_{t+h} - X_t \geq 2) = o(h)$.

(c) X *has independent and stationary increments and for all* $t \geq 0$, $X_t \sim$ $\mathrm{Poisson}(\lambda t)$.

Proof: We show (a) implies (b). If (a) holds, then the increments are independent and stationary. Using the fact it is stationary, we have uniformly in t as $h \to 0$,

$$\mathbb{P}(X_{t+h} - X_t = 0) = \mathbb{P}(X_h = 0) = \mathbb{P}(S_1 > h) = e^{-\lambda h} = 1 - \lambda h + o(h),$$

$$\mathbb{P}(X_{t+h} - X_t \geq 1) = \mathbb{P}(X_h \geq 1) = \mathbb{P}(S_1 \leq h) = 1 - e^{-\lambda h} = \lambda h + o(h),$$

$$\mathbb{P}(X_{t+h} - X_t \geq 2) = \mathbb{P}(X_h \geq 2) = \mathbb{P}(S_1 + S_2 \leq h) \leq \mathbb{P}(S_1 \leq h)^2$$
$$= (\lambda h + o(h))^2 = o(h).$$

Now we show (b) implies (c). If X satisfies (b), then $(X_{t+s} - X_s)_{t\geq 0}$ also satisfies (b), so X has independent and stationary increments. Set $p_j = \mathbb{P}(X_t = j)$. Then since the increments are independent and X is increasing, for $j \geq 1$,

$$p_j(t+h) = \mathbb{P}(X_{t+h} = j) = \sum_{i=0}^{j} \mathbb{P}(X_t = j-1)\mathbb{P}(X_{t+h} - X_t = i)$$
$$= p_j(t)(1 - \lambda h + o(h)) + p_{j-1}(t)(\lambda h + o(h)) + o(h).$$

Rearranging this, we get

$$\frac{p_j(t+h) - p_j(t)}{h} = -\lambda p_j(t) + \lambda p_{j-1}(t) + o(1).$$

Setting $s = t + h$, we get

$$\frac{p_j(s+h) - p_j(s)}{h} = -\lambda p_j(s-h) + \lambda p_{j-1}(s-h) + o(1).$$

These equations imply $p_j(t)$ is a continuous function of t, and as $h \to 0$,

$$p'_j(t) = -\lambda p_j(t) + \lambda p_{j-1}(t).$$

Differentiating $e^{\lambda t} p_j(t)$ with respect to t, we get that

$$\left(e^{\lambda t} p_j(t)\right)' = \lambda e^{\lambda t} p_j(t) + e^{\lambda t} p'_j(t) = \lambda e^{\lambda t} p_{j-1}(t).$$

For $j = 0$, we similarly get $p_0(t+h) = p_0(t)(1-\lambda h) + o(h)$. So $p'_0(t) = -\lambda p_0(t)$, implying $p_0(t) = e^{-\lambda t}$.

Since $p'_1(t) = -\lambda p_1(t) + \lambda e^{-\lambda t}$, we get $p_1(t) = \lambda t e^{-\lambda t}$, and by induction,

$$p_k(t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

So $X_t \sim \text{Poisson}(\lambda t)$.

Finally we show (c) implies (a). Since X has independent stationary, we can write

$$\begin{aligned}
\mathbb{P}(X_{t_1} &= n_1, X_{t_2} = n_2, \ldots, X_{t_k} = n_k) \\
&\mathbb{P}(X_{t_1} = n_1)\mathbb{P}(X_{t_2} - X_{t_1} = n_2 - n1) \cdots \mathbb{P}(X_{t_k} - X_{t_{k-1}} = n_k - n_{k-1}) \\
&= \mathbb{P}(X_{t_1} = n_1)\mathbb{P}(X_{t_2-t_1} = n_2 - n_1) \cdots \mathbb{P}(X_{t_k} - X_{t_{k-1}} = n_k - n_{k-1}).
\end{aligned}$$

But these probabilities are known; they are Poisson. So (c) determines the finite dimensional distribution of a right-continuous process X, hence it determines X. As we know the Poisson process produces (c), we get (c) implies (a).

We went through this process to obtain the Poisson distribution for each X_t. Let's see if we can do this directly. Note

$$\begin{aligned}
\mathbb{P}(X_t = n) &= \mathbb{P}(S_1 + \cdots + S_n \leq t < S_1 + \cdots + S_{n+1}) \\
&= \mathbb{P}(S_1 + \cdots + S_n \leq t) - \mathbb{P}(S_1 + \cdots + S_{n+1} \leq t) \\
&= \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} \, \mathrm{d}x - \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^n}{n!} \, \mathrm{d}x \\
&= e^{-\lambda t} \frac{(\lambda t)^n}{n!},
\end{aligned}$$

from integration by parts, and the fact the sums $S_1 + \cdots + S_n \sim \Gamma(n, \lambda)$.

Theorem 2.4 (Superposition)**.** *Let X and Y be two independent Poisson processes*

with parameters λ and μ respectively. Then,

$$Z_t = X_t + Y_t \sim PP(\lambda + \mu).$$

Proof: To prove this, we use (c). Z has stationary with independent processes, and $Z_t \sim \text{Poisson}(\lambda t + \mu_t)$, as it is a sum of independent Poisson distributions.

Theorem 2.5 (Thinning). *Let X be $PP(\lambda)$, and let $(Z_i)_{i \geq 0}$ be iid $\text{Ber}(p)$. Let Y be a process with values in $\{0, 1, 2, \ldots\}$ which jumps at time t if and only if*

- *X jumps at time t, and*

- *$Z_{X_t} = 1$.*

In other words, we keep every point of X with probability p independently over all points. Then $Y \sim PP(\lambda p)$ and $X - Y$ is an independent Poisson process with parameter $\lambda(1 - p)$.

Proof: We will use the infinitesimal definition. The independence of increments for Y is clear. Since $\mathbb{P}(X_{t+h} - X_t \geq 2) = o(h)$, we have

$$\mathbb{P}(Y_{t+h} - Y_t = 1) = p \cdot \mathbb{P}(X_{t+h} - X_t = 1) + o(h) = \lambda p h + o(h).$$

We also get

$$\begin{aligned}\mathbb{P}(Y_{t+h} - Y_t = 0) &= \mathbb{P}(X_{t+h} - X_t = 0) + (1 - p)\mathbb{P}(X_{t+h} - X_t = 1) + o(h) \\ &= 1 - \lambda h + (1 - p)(\lambda h + o(h)) + o(h) = 1 - \lambda p h + o(h).\end{aligned}$$

This proves $Y \sim PP(\lambda p)$. Since $X - Y$ is also a thinning of X with probability $1 - p$, $X - Y \sim PP(\lambda(1 - p))$.

Now we prove that X and $X - Y$ are independent. It is enough to show that the finite dimensional distributions of Y and $X - Y$ are independent, so

$$\begin{aligned}\mathbb{P}(Y_{t_1} = n_1, \ldots, Y_{t_k} &= n_k, X_{t_1} - Y_{t_1} = m_1, \ldots, X_{t_k} - Y_{t_k} = m_k) \\ &= \mathbb{P}(Y_{t_1} = n_1, \ldots, Y_{t_k} = n_k)\mathbb{P}(X_{t_1} - Y_{t_1} = m_1, \ldots, X_{t_k} - Y_{t_k} = m_k).\end{aligned}$$

We will show this for a fixed t. The general case follows from independence

of increments. Indeed,

$$
\begin{aligned}
\mathbb{P}(Y_t = n, X_t - Y_t = m) &= \mathbb{P}(X_t = m + n, Y_t = n) \\
&= \mathbb{P}(X_t = m + n)\mathbb{P}(Y_t = n \mid X_t = m + n) \\
&= e^{-\lambda t}\frac{(\lambda t)^{m+n}}{(m+n)!}\mathbb{P}(\mathrm{Bin}(m+n, p) = n) \\
&= e^{-\lambda t}\frac{(\lambda t)^{m+n}}{(m+n)!}\binom{m+n}{n}p^n(1-p)^n \\
&= e^{-\lambda tp}\frac{(\lambda tp)^n}{n!}e^{-\lambda t(1-p)}\frac{(\lambda t(1-p))^m}{m!} \\
&= \mathbb{P}(Y_t = n)\mathbb{P}(X_t - Y_t = m).
\end{aligned}
$$

Theorem 2.6. *Let X be a Poisson process. Then conditional on the event $\{X_t = n\}$, the jump times $J_1, \ldots, J_n$ are distributed as the order-statistics of n iid $U[0,t]$ random variables, i.e. their joint density is*

$$
f(t_1, \ldots, t_n) = \frac{n!}{t^n}\mathbb{1}(0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq t).
$$

Proof: Since $S_1, S_2, \ldots$ are iid $\mathrm{Exp}(\lambda)$, the joint density of $(S_1, \ldots, S_{n+1})$ is

$$
\lambda^{n+1}e^{-\lambda(S_1 + \cdots + S_{n+1})}\mathbb{1}(S_1, \ldots, S_{n+1} \geq 0).
$$

Then the jump times $J_1 = S_1, J_2 = S_1 + S_2, \ldots$ have joint density

$$
g(t_1, \ldots, t_{n+1}) = \lambda^{n+1}e^{-\lambda t_{n+1}}\mathbb{1}(0 \leq t_1 \leq \cdots \leq t_{n+1}).
$$

Now taking $A \subset \mathbb{R}^n$,

$$
\mathbb{P}((J_1, \ldots, J_n) \in A \mid X_t = n) = \frac{\mathbb{P}((J_1, \ldots, J_n) \in A, X_t = n)}{\mathbb{P}(X_t = n)}.
$$

We can calculate

$$
\begin{aligned}
\mathbb{P}((J_1, \ldots, J_n) \in A, X_t = n) &= \mathbb{P}((J_1, \ldots, J_n) \in A, J_n \leq t < J_{n+1}) \\
&= \int_{A, t_{n+1}} g(t_1, \ldots, t_{n+1})\mathbb{1}(t_{n+1} > t \geq t_n)\,\mathrm{d}\mathbf{t}\,\mathrm{d}t_{n+1} \\
&= \int_A \int_{t_{n+1}=t}^{\infty} \lambda^{n+1}e^{-\lambda t_{n+1}}\mathbb{1}(0 \leq t_1 \leq \cdots \leq t_n \leq t)\,\mathrm{d}\mathbf{t}\,\mathrm{d}t_{n+1} \\
&= \int_A \lambda^n e^{-\lambda t}\mathbb{1}(0 \leq t_1 \leq \cdots \leq t_n \leq t)\,\mathrm{d}\mathbf{t}.
\end{aligned}
$$

We also know that
$$\mathbb{P}(X_t = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!},$$
so dividing we get
$$\mathbb{P}((J_1, \ldots, J_n) \in A \mid X_t = n) = \int_A \frac{n!}{t^n} \mathbb{1}(0 \le t_1 \le \cdots \le t_n \le t)\, \mathrm{d}\mathbf{t}$$
$$= \int_A f(t_1, \ldots, t_n)\, \mathrm{d}\mathbf{t},$$
proving the joint distribution is as desired.

3   Birth Process

A *birth process* is a generalization of a Poisson process. For a Poisson process, the rate of going from i to $i+1$ is λ, uniform for all i. For a birth process, this is q_i.

For each i, let S_i be $\text{Exp}(q_i)$, such that $S_1, S_2, \ldots$ are independent. Then set $T_i = S_1 + S_2 + \cdots + S_i$, and $X_t = i$ if $J_i \leq t < j_{i+1}$. Then X is called a birth process.

- A *simple birth process* is when $q_i = \lambda i$ for all i.

- A Poisson process is a birth process with $q_i = \lambda$ for all i.

3.1   Motivation for Simple Birth Processes

At time 0, there is only one individual, so $X_0 = 1$. Each individual has en exponential clock of parameter λ independently.

Then, if there are i individuals, then the first clock rings after $\text{Exp}(\lambda i)$ time, and we move from i to $i+1$ individuals. By the memoryless property, the process begins again.

Proposition 3.1. *Let $(T_k)_{k \geq 1}$ be a sequence of independent random variables with $T_k \sim \text{Exp}(q_k)$, and $0 < \sum q_k < \infty$. Let $T = \inf_k T_k$. Then,*

(a) $T \sim \text{Exp}(\sum q_k)$,

(b) *The infimum is attained at a unique point K almost surely, and*

$$\mathbb{P}(K = k) = \frac{q_k}{\sum q_k},$$

(c) *T and K are independent.*

The proof is found on the example sheet.

The main difference between a Poisson process and a birth process is the possibility of explosion in a birth process. Recall that explosion occurs when $\sum S_n < \infty$.

Proposition 3.2. *Let X be a birth process with rates q_i and $X_0 = 1$.*

1. *If $\sum \frac{1}{q_i} < \infty$, then $\mathbb{P}(\zeta < \infty) = 1$, i.e. the birth process is explosive.*

2. *If $\sum \frac{1}{q_i} = \infty$, then $\mathbb{P}(\zeta = \infty) = 1$, i.e. the birth process is non-explosive.*

Note that both the Poisson process and simple birth process are part of case 2, hence both are non-explosive.

Proof:

1. If $\sum \frac{1}{q_n} < \infty$, then

$$\mathbb{E}\left[\sum_{n=1}^{\infty} S_n\right] = \sum_{n=1}^{\infty} \mathbb{E}[S_n] = \sum_{n=1}^{\infty} \frac{1}{q_n} < \infty.$$

Thus $\zeta = \sum_n S_n < \infty$ almost surely.

2. If $\sum \frac{1}{q_n} = \infty$, then

$$\prod_{n=1}^{\infty} \left(1 + \frac{1}{q_n}\right) \geq 1 + \sum_{n=1}^{\infty} \frac{1}{q_n} = \infty.$$

Then,

$$\begin{aligned}
\mathbb{E}[e^{-\zeta}] &= \mathbb{E}\left[\exp\left(-\sum_{n=1}^{\infty} S_n\right)\right] \\
&= \lim_{n \to \infty} \mathbb{E}\left[\exp\left(-\sum_{i=1}^{n} S_i\right)\right] \\
&= \lim_{n \to \infty} \prod_{i=1}^{n} \mathbb{E}[e^{-S_i}] \\
&= \lim_{n \to \infty} \prod_{i=1}^{n} \frac{1}{1 + 1/q_i} = \prod_{i=1}^{n} \frac{1}{1 + 1/q_i} = 0.
\end{aligned}$$

As $e^{-\zeta} \geq 0$ with $\mathbb{E}[e^{-\zeta}] = 0$, we must have $e^{-\zeta} = 0$ almost surely, so $\mathbb{P}(\zeta = \infty) = 1$.

Theorem 3.1 (Markov Property)**.** *Let X be a birth process with rates q_i. Then conditional on $X_s = i$, the process $(X_{s+t})_{t \geq 0}$ is a birth process with rates q_j for $j > i$, starting from i and independent of $(X_\gamma)_{\gamma \leq s}$.*

This proof is the same for Poisson processes. Moreover, we also have the following similarities to Poisson processes:

Theorem 3.2. *Let X be an increasing right-continuous process with values in $\{1, 2, \ldots\} \cup \{\infty\}$. Let $0 \leq q_j < \infty$ for all $j \geq 0$. Then the following are equivalent:*

(a) (Jump chain/holding time definition) *Conditional on $X_s = i$, the holding times $S_1, S_2, \ldots$ are independent exponential with rates $q_i, q_{i+1}, \ldots$ respectively,*

and the jump chain is given by $Y_n = i + n$ for all n.

(b) (Infinitesimal definition) *For all* $t, h \geq 0$, *conditional on* $X_t = i$, *the process* $(X_{t+h})_{h \geq 0}$ *is independent of* $(X_s \mid s \leq t)$ *and as* $h \downarrow 0$ *uniformly in* t,

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

(c) (Transition probability definition) *For all* $n = 0, 1, 2, \ldots$ *and all times* $0 \leq t_0 \leq t_1 \leq \cdots \leq t_{n+1}$, *and all states* $i_0, i_1, \ldots, i_{n+1}$,

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} \mid X_0 = i_0, \ldots, X_{t_n} = i_n) = P_{i_n, i_{n+1}}(t_{n+1} - t_n),$$

where $(P_{i,j}(t))$ *is the unique solution to the equation*

$$P'_{i,j}(t) = q_{j-1} P_{i,j-1}(t) - q_j P_{i,j}(t).$$

As in the Poisson process,

$$P_{i,j}(t + h) = P_{i,j-1}(t) q_{j-1} h + P_{i,j}(t)(1 - q_j h) + o(h).$$

Existence and uniqueness follows since for $i = j$,

$$P'_{i,i}(t) = -q_i P_{i,i}(t),$$

and $P_{i,i}(0) = 1$, so $P_{i,i}(t) = e^{-q_i t}$, so by induction the unique solution for $P_{i,j}(t)$ exists.

Also note that we can write $P'(t) = P(t)Q$ where $P(t) = (P_{ij}(t))$ and

$$Q = \begin{pmatrix} -q_1 & q_1 & 0 & 0 & \cdots \\ 0 & -q_2 & q_2 & 0 & \cdots \\ 0 & 0 & -q_3 & q_3 & \cdots \\ 0 & 0 & 0 & q_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

4   Q-matrix and Construction of Markov Processes

Definition 4.1. $Q = (q_{ij})_{i,j \in I}$ is called a *Q-matrix* if

(a) $0 \leq -q_{ii} < \infty$ for all $i \in I$,

(b) $0 \leq q_{ij} < \infty$ for all $i \neq j \in I$,

(c) $\sum_j q_{ij} = 0$ for all $i \in I$.

We write $q_i = -q_{ii} = \sum_{j \neq i} q_{ij}$ for all $i \in I$.

Given a Q-matrix Q, we define a jump matrix P as follows: for $x \neq y$ with $q_x \neq 0$, set $p_{xy} = \frac{q_{xy}}{q_x}$ and $p_{xx} = 0$. If $q_x = 0$, set $p_{xy} = \mathbb{1}(x = y)$.

Example 4.1.

If
$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 2 & 1 & -1 \end{pmatrix},$$
then
$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

Note that P is a stochastic matrix.

Definition 4.2. Let Q be a Q-matrix and λ be a probability measure on I. Then a (minimal) random process X is a *Markov process* with initial distribution λ and infinitesimal generator Q if:

(a) The jump chain $Y_n = X_{J_n}$ is a discrete-time Markov chain with $Y_0 \sim \lambda$ and transition matrix P.

(b) Conditional on $Y_0, Y_1, \ldots, Y_n$, the holding times $S_1, S_2, \ldots, S_{n+1}$ are independent with $S_i \sim \text{Exp}(q_{Y_{i-1}})$ for $i = 1, 2, \ldots, n + 1$.

We write $X \sim \text{Markov}(\lambda, Q)$.

> **Example 4.2.**
>
> Birth processes are Markov(λ, Q) with $I = \mathbb{N}$ and
>
> $$Q = \begin{pmatrix} -q_1 & q_1 & 0 & 0 & \cdots \\ 0 & -q_2 & q_2 & 0 & \cdots \\ 0 & 0 & -q_3 & q_3 & \cdots \\ 0 & 0 & 0 & q_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$
>
> Note that
>
> $$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$
>
> We know exactly that $Y_n = Y_0 + n$.

4.1   Constructions of Markov continuous processes

1. (Y_n) is a discrete time Markov chain, $Y_0 \sim \lambda$ with transition matrix P.

   (T_i) are iid $\mathrm{Exp}(1)$, independent of Y, and set

   $$S_n = \frac{T_n}{q_{Y_{n-1}}}, \qquad J_n = \sum_{i=1}^{n} S_i.$$

   Then set $X_t = Y_n$ if $J_n \le t < J_{n+1}$, and $X_t = \infty$ otherwise.

2. Let (T_n^y) be iid $\mathrm{Exp}(1)$ random variables, with $Y_0 \sim \lambda$, and inductively define Y_n, S_n as, if $Y_n = x$, then for $y \ne x$, let

   $$S_{n+1}^y = \frac{T_{n+1}^y}{q_{xy}} \sim \mathrm{Exp}(q_{xy}), \qquad S_{n+1} = \inf_{y \ne x} S_{n+1}^y.$$

   If $S_{n+1} = S_{n+1}^z$ for some random z, then take $Y_{n+1} = z$, if $q_x > 0$. If $q_x = 0$, then take $Y_{n+1} = x$.

3. For $x \ne y$, let $(N_t^{x,y})$ be independent Poisson processes with rates q_{xy} respectively. Let $Y_0 \sim \lambda$, $J_0 = 0$ and define inductively

   $$J_{n+1} = \inf\{t > J_n \mid N_t^{Y_n,y} \ne N_{J_n}^{Y_n,y}\},$$

and define

$$Y_{n+1} = \begin{cases} y & J_{n+1} < \infty, N_{J_{n+1}}^{Y_n,Y} \neq N_{J_n}^{Y_n,y}, \\ Y_n & J_{n+1} = \infty. \end{cases}$$

For birth processes, we characterized when non-explosion happens. In general, we only give a sufficient condition.

Theorem 4.1. *Let X be* $\text{Markov}(\lambda, Q)$ *on I. Then $\mathbb{P}(\zeta = \infty) = 1$ if any of the following holds:*

(a) *I is finite.*

(b) *$\sup_{x \in I} q_x < \infty$.*

(c) *$X_0 = x$ and x is recurrent for the jump chain Y.*

> **Proof:**   Note (a) implies (b), so it is enough to prove for (b).
>
> Set $q = \sup_{x \in I} q_x < \infty$. Then the holding times
>
> $$S_n \geq \frac{T_n}{q},$$
>
> where $(T_i) \sim \text{Exp}(1)$. Then
>
> $$\zeta = \sum_{n=1}^{\infty} S_n \geq \frac{1}{q} \sum_{n=1}^{\infty} T_n = \infty$$
>
> with probability one. Hence $\mathbb{P}(\zeta = \infty) = 1$.
>
> For (c), let (N_i) be the times when the jump chain Y_i visits x. By the strong law of large numbers,
>
> $$\zeta \geq \sum_{i=1}^{\infty} S_{N_i+1} = \sum_{i=1}^{\infty} \frac{T_{N_i+1}}{q_{N_i}} = \frac{1}{q_x} \sum_{i=1}^{\infty} T_{N_i+1} = \infty.$$
>
> Hence $\mathbb{P}(\zeta = \infty) = 1$.

Example 4.3.

Let $I = \mathbb{Z}$, and $q_{i,i+1} = q_{i,i-1} = 2^{|i|}$. Then $p_{i,i+1} = p_{i,i-1} = \frac{1}{2}$. So the jump chain Y is the simple random walk on $\mathbb{Z}$, which is recurrent. Hence X is non-explosive.

Now consider $I = \mathbb{Z}$, where $q_{i,i+1} = 2^{|i|+1}$, and $q_{i,i-1} = 2^{|i|}$. Then Y is a simple random walk with $q_{i,i+1} = \frac{2}{3}$, and $q_{i,i-1} = \frac{1}{3}$. This is transient, so we cannot use the theorem. Instead,

$$\mathbb{E}[\zeta] = \mathbb{E}\left[\sum_{n=1}^{\infty} S_n\right] = \sum_{j \in \mathbb{Z}} \mathbb{E}\left[\sum_{k=1}^{V_j} S_{N_k^j + 1}\right]$$

$$= \sum_{j \in \mathbb{Z}} \mathbb{E}[V_j]\mathbb{E}[S_{N_1^j + 1}] = \sum_{j \in \mathbb{Z}} \mathbb{E}[V_j]\frac{1}{q_j}$$

$$= \sum_{j \in \mathbb{Z}} \frac{1}{3 \cdot 2^{|i|}}\mathbb{E}[V_j].$$

Note that $\mathbb{E}[V_j] \leq 1 + \mathbb{E}_j[V_j] = 1 + \mathbb{E}_0[V_0] = c < \infty$ as the jump chain is transient. Hence

$$\mathbb{E}[\zeta] = \sum_{j \in \mathbb{Z}} \frac{1}{3 \cdot 2^{|j|}}\mathbb{E}[V_j] \leq \sum_{j \in \mathbb{Z}} \frac{c}{3 \cdot 2^{|j|}} < \infty.$$

Therefore, $\mathbb{P}(\zeta < \infty) = 1$, so the process is explosive.

Theorem 4.2. *Let X be* $\mathrm{Markov}(\lambda, Q)$ *and T a stopping time. Then conditional on $T < \zeta$, and $X_T = x$, the process $(X_{T+t})_{t \geq 0}$ is* $\mathrm{Markov}(\delta_x, Q)$ *and independent of $(X_s)_{s \leq T}$.*

This is not proved: it uses measure theory. A reference can be found in James Norris notes.

4.2   Kolmogorov's Forward and Backward Equations

We work with a countable state space I.

Theorem 4.3. *Let X be a minimal right continuous process with values in a countable set I. Let Q be a Q-matrix with jump matrix P. Then the following are equivalent:*

(a) *X is a continuous time Markov chain with generator Q.*

(b) *For all $n \geq 0$, $0 \leq t_0 \leq \cdots \leq t_n \leq t_{n+1}$ and all states $x_0, \ldots, x_{n+1} \in I$,*

$$\mathbb{P}(X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n, \ldots, X_{t_0} = x_0) = P_{x_n, x_{n+1}}(t_{n+1} - t_n),$$

where $(P(t)) = (P_{xy}(t))$ is the minimal non-negative solution to the backward

equation

$$P'(t) = QP(t), \qquad P(0) = I.$$

Minimality means that if $\tilde{P}$ is another non-negative solution to the backward equation, then

$$P_{xy}(t) \leq \tilde{P}_{xy}(t).$$

In fact, if the chain is non-explosive, the solution is unique.

(c) *$P(t)$ is also the minimal non-negative solution to the forwards equation*

$$P'(t) = P(t)Q, \qquad P(0) = I.$$

We will show (a) is equivalent to (b), and skip the proof of the equivalence of (c). The proof of this is found in Norris' notes.

Proof: To show (a) implies (b), let (J_n) denote the jump times. Then conditioning on the first jump time,

$$\mathbb{P}_x(X_t = y, J_1 > t) = \mathbb{1}(x = y)e^{-q_x t}.$$

Integrating on the values of $J_1 < t$ and using the independence of the jump chain to the values of the holding times, for $z \neq x$,

$$\mathbb{P}_x(X_t = y, J_1 \leq t, X_{J_1} = z) = \int_0^t q_x e^{-q_x s} \frac{q_{xz}}{q_x} P_{zy}(t - s)\, ds$$

$$= \int_0^t e^{-q_x s} q_{xz} P_{zy}(t - s)\, ds$$

Summing over all $z \neq x$,

$$\mathbb{P}_x(X_t = y, J_1 \leq t) = \int_0^t \sum_{z \neq x} e^{-q_x s} q_{xz} P_{zy}(t - s)\, ds.$$

Hence

$$P_{xy}(t) = \mathbb{P}_x(X_t = y) = e^{-q_x t} \mathbb{1}(x = y) + \int_0^t \sum_{z \neq x} e^{-q_x s} q_{xz} P_{zy}(t - s)\, ds.$$

Making a change of variable $u = t - s$, we get

$$e^{q_x t} P_{xy}(t) = \mathbb{1}(x = y) + \int_0^t \sum_{z \neq x} e^{q_x u} q_{xz} P_{zy}(u)\, du.$$

Thus $P_{xy}(t)$ is a continuous function in t, and hence

$$\sum_{z \neq x} e^{q_x u} q_{xz} P_{zy}(u)$$

is a series of continuous functions, and is also uniformly convergent, and hence continuous. Therefore the integral is differentiable, and

$$e^{q_x t}(q_x P_{xy}(t) + P'_{xy}(t)) = \sum_{z \neq x} e^{q_x t} q_{xz} P_{zy}(t).$$

Thus

$$P'_{xy}(t) = \sum_z q_{xz} P_{zy}(t),$$

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

To show minimality, let $\tilde{P}$ be another non-negative solution of the backward equation. We will show

$$P_{xy}(t) \leq \tilde{P}_{xy}(t),$$

for all x, y, t. As before,

$$\mathbb{P}_x(X_t = y, t < J_{n+1}) = \mathbb{P}_x(X_t = y, J_1 > t) + \mathbb{P}_x(X_t = y, J_1 \leq t < J_{n+1})$$

$$= e^{-q_x t} \mathbb{1}(x = y) + \sum_{z \neq x} \int_0^t q_x e^{-q_x s} \frac{q_{xz}}{q_x} \mathbb{P}_z(X_{t-s} = y, t - s < J_n) \, \mathrm{d}s.$$

Now as $\tilde{P}$ satisfies the backwards equation, we get as before

$$\tilde{P}_{xy}(t) = e^{-q_x t} \mathbb{1}(x = y) + \sum_{z \neq x} \int_0^t e^{-q_x s} q_{xz} \tilde{P}_{zy}(t - s) \, \mathrm{d}s.$$

Now we prove by induction that

$$\mathbb{P}_x(X_t = y, t < J_n) \leq \tilde{P}_{xy}(t),$$

for all n. For $n = 1$,

$$e^{-q_x t} \mathbb{1}(x = y) \leq \tilde{P}_{xy}(t),$$

by our equation for $\tilde{P}_{xy}(t)$. Assuming for n, then for $n + 1$,

$$\mathbb{P}_x(X_t = y, t < J_{n+1}) \leq e^{-q_x t} \mathbb{1}(x = y) + \sum_{z \neq x} \int_0^t q_x e^{-q_x s} P_{zy}(t - s) \, \mathrm{d}s = \tilde{P}_{xy}(t).$$

Hence,

$$\mathbb{P}_x(X_t = y, t < \zeta) = \lim_{n \to \infty} \mathbb{P}_x(X_t = y, t < J_n) \leq \tilde{P}_{xy}(t).$$

Using minimality,

$$P_{xy}(t) = \mathbb{P}_x(X_t = y) = \mathbb{P}_x(X_t = y, t < \zeta) \leq \tilde{P}_{xy}(t).$$

4.3 Finite State Spaces

Definition 4.3. If A is a finite-dimension square matrix, its *matrix exponential* is given by

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + A + \frac{A^2}{2!} + \cdots .$$

Proposition 4.1. *For any $r \times r$ matrix A, the exponential e^A is an $r \times r$ matrix, and if A_1, A_2 commute, then*

$$e^{A_1 + A_2} = e^{A_1} e^{A_2}.$$

Proposition 4.2. *Let Q be a Q-matrix on a finite set I, and let*

$$P(t) = e^{tQ}.$$

Then,

(i) $P(t + s) = P(t)P(s)$ *for all s, t.*

(ii) $(P(t))$ *is the unique solution to the forward equation $P'(t) = P(t)Q$, $P(0) = I$.*

(iii) $(P(t))$ *is the unique solution to the backward equation $P'(t) = QP(t)$, $P(0) = I$.*

(iv) *For $k = 0, 1, 2, \ldots$*

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} \right)^k P(t) \bigg|_{t=0} = Q^k.$$

Proof:

(i) Since tQ and sQ commute,

$$P(t + s) = e^{(t+s)Q} = e^{tQ} e^{sQ} = P(t)P(s).$$

(ii) The sum in e^{tQ} has infinite radius of convergence, hence we can differentiate term by term to get

$$\frac{\mathrm{d}}{\mathrm{d}t}P(t) = \frac{\mathrm{d}}{\mathrm{d}t}e^{tQ} = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=0}^{\infty}\frac{(tQ)^k}{k!} = \sum_{k=0}^{\infty}\frac{\mathrm{d}}{\mathrm{d}t}\frac{t^k Q^k}{k!}$$

$$= \sum_{k=1}^{\infty}\frac{kt^{k-1}}{k!}Q^k = \sum_{k=1}^{\infty}\frac{t^{k-1}}{(k-1)!}Q^{k-1}Q = Qe^{tQ} = P(t)Q.$$

(iii) Similarly $P'(t) = QP(t)$.

We get that

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{tQ} = Qe^{tQ} = e^{tQ}Q.$$

For uniqueness, if $\tilde{P}$ is another solution to the forward equation, then $\tilde{P}'(t) = Q\tilde{P}(t)$ and $\tilde{P}(0) = I$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\tilde{P}(t)e^{-tQ}\right) = \tilde{P}'(t)e^{-tQ} + \tilde{P}(t)\frac{\mathrm{d}}{\mathrm{d}t}e^{-tQ}$$

$$= \tilde{P}(t)Qe^{-tQ} + \tilde{P}(t)(-Qe^{-tQ}) = 0,$$

so $\tilde{P}(t)e^{-tQ}$ is a constant matrix. Since $\tilde{P}(0)e^0 = I$, we get $\tilde{P}(t) = e^{tQ}$.

Similarly, if $\tilde{P}$ is a solution to the backward equation, we can find $e^{-tQ}\tilde{P}(t) = I$.

(iv) is obtained by differentiating term-by-term k times.

Example 4.4.

Let

$$Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}.$$

To find $P_{11}(t)$, we can diagonalise $Q = PDP^{-1}$, where

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

to get

$$e^{tQ} = \sum_{k=01}^{\infty} \frac{t^k Q^k}{k!} = \sum_{k=0}^{\infty} t^k \frac{PD^k P^{-1}}{k!}$$

$$= P \begin{pmatrix} e^{t\lambda_1} & 0 & 0 \\ 0 & e^{t\lambda_2} & 0 \\ 0 & 0 & e^{t\lambda_3} \end{pmatrix} P^{-1},$$

which gives $P_{11}(t) = ae^{t\lambda_1} + be^{t\lambda_2} + ce^{t\lambda_3}$. To find the constants a, b, c, note $P_{11}(0) = 1$, $P'_{11}(0) = q_{11}$ and $P''_{11}(0) = Q^2_{11}$.

Theorem 4.4. *Let I be a finite state space and Q be a matrix. Then, it is a Q-matrix if and only if $P(t) = e^{tQ}$ is a stochastic matrix for all $t \geq 0$.*

Proof: For $t \downarrow 0$, $e^{tQ} = I + tQ + \mathcal{O}(t^2)$. Hence, for $x \neq y$, $q_{xy} \geq 0$ if and only if $P_{xy}(t) \geq 0$ for all $t \geq 0$ sufficiently small.

Since $P(t) = P(t/n)^n$ for all n, we get $q_{xy} \geq 0$ if and only if, $p_{xy}(t) \geq 0$ for all $t \geq 0$.

Assume now that Q is a Q-matrix, so $\sum_y q_{xy} = 0$ for all x. Then,

$$\sum_y Q^n_{xy} = \sum_y \sum_z Q^{n-1}_{xz} Q_{zy} = \sum_z Q^{n-1}_{xz} \sum_y Q_{zy} = 0.$$

Hence, since

$$P_{xy}(t) = \delta_{xy} + \sum_{k=1}^{\infty} \frac{t^k}{k!} Q^k_{xy},$$

we must have

$$\sum_y P_{xy}(t) = 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_y Q^k_{xy} = 1,$$

proving $P(t)$ is a stochastic matrix.

Assuming that $P(t)$ is a stochastic matrix, then as

$$Q = \frac{d}{dt}\bigg|_{t=0} P(t),$$

then

$$\sum_y q_{xy} = \frac{d}{dt}\bigg|_{t=0} \sum_y P_{xy}(t) = 0,$$

hence Q is a Q-matrix.

Theorem 4.5. *Let X be a right-continuous process with values in a finite set I, and let Q be a Q-matrix on I. Then the following are equivalent:*

(a) *The process is Markov with generator Q (written* $\mathrm{Markov}(Q)$).

(b) *Infinitesimal definition. Conditional on $X_s = x$, the process $(X_{s+t})_{t \geq 0}$ is independent of $(X_r)_{r \leq s}$, and uniformly in t as $h \downarrow 0$,*

$$\mathbb{P}(X_{t+h} = y \mid X_t = x) = \mathbb{1}(x = y) + q_{xy}h + o(h).$$

(c) *For all $n \geq 0$, $0 \leq t_0 \leq \cdots \leq t_n$ and all states $x_0, \ldots, x_n$,*

$$\mathbb{P}(X_{t_n} = x_n \mid X_{t_0} = x_0, \ldots, X_{t_{n-1}} = x_{n-1}) = P_{x_{n-1}x_n}(t_n - t_{n-1}),$$

where $(P_{xy}(t))$ is the solution to the forward equation

$$P'(t) = P(t)Q, \qquad P(0) = I.$$

Proof: We have shown the equivalence of (a) and (c), so it is enough to show the equivalence of (b) and (c).

To show (c) implies (b), we know $P(t) = e^{tQ}$ is the solution to the forward equation. As $t \downarrow 0$, $P(t) = I + tQ + \mathcal{O}(t^2)$, thus for all $t > 0$ and as $h \downarrow 0$, for all x and y,

$$\mathbb{P}(X_{t+h} = y \mid X_t = x) = \mathbb{P}(X_h = y \mid X_0 = x) = P_{xy}(h) = \delta_{xy} + hq_{xy} + o(h).$$

To show (b) implies (c), note

$$P_{xy}(t + h) = \sum_z P_{xz}(t) \cdot (\mathbb{1}(z = y) + q_{zy}h + o(h)).$$

So rearranging,

$$\frac{P_{xy}(t + h) - P_{xy}(t)}{h} = \sum_z P_{xz}(t)q_{zy} + o(1).$$

As $h \downarrow 0$,

$$P'_{xy}(t) = \sum_z P_{xz}(t)q_{zy} = P(t)Q.$$

5   Qualitative Properties of Continuous Time Markov Chains

We consider minimal chains, and countable state spaces.

5.1   Class Structure

Definition 5.1. For states $x, y \in I$, we write $x \to y$ if

$$\mathbb{P}_x(X_t = y \text{ for some } t > 0) > 0.$$

We write $x \leftrightarrow y$ if $x \to y$ and $y \to x$.

We can define *communicating classes*, *irreducibility*, *closed class* and *absorbing states* exactly as in the discrete-time case.

Proposition 5.1. *Let X be* $\mathrm{Markov}(Q)$ *with transition semigroup* $(P(t))_{t \geq 0}$*. For any two states $x, y \in I$, the following are equivalent:*

   (a) $x \to y$.

   (b) $x \to y$ *for the jump chain.*

   (c) $q_{x_0 x_1} \cdots q_{x_{n-1} x_n} > 0$ *for some $x = x_0$, $x_1, \ldots, x_{n-1}$, $x_n = y$.*

   (d) $P_{xy}(t) > 0$ *for all $t > 0$.*

   (e) $P_{xy}(t) > 0$ *for some $t > 0$.*

Proof: Clearly (d) implies (e) implies (a) implies (b).

To show (b) implies (c), since $x \to y$ for the jump chain, there exists $x_0 = x, x_1, \ldots, x_n = y \in I$ such that

$$P_{x_0 x_1} P_{x_1 x_2} \cdots P_{x_{n-1} x_n} > 0.$$

By the definition of the jump chain, we immediately get

$$q_{x_0 x_1} q_{x_1 x_2} \cdots q_{x_{n-1} x_n} > 0.$$

Now to show (c) implies (d), for any two states w, z with $q_{wz} > 0$, and for any $t > 0$,

$$P_{wz}(t) \geq \mathbb{P}_\omega(J_1 \leq t, Y_1 = z, S_2 > t) = (1 - e^{-q_w t})\frac{q_{wz}}{q_w}e^{-q_z t} > 0.$$

Hence if $q_{wz} > 0$, $P_{wz}(t) > 0$. Therefore if (c) holds then $P_{x_i x_{i+1}}(t) > 0$ for all t, hence

$$P_{xy}(t) \geq P_{x_0 x_1}\left(\frac{t}{n}\right) P_{x_1 x_2}\left(\frac{t}{n}\right) \cdots P_{x_{n-1} x_n}\left(\frac{t}{n}\right) > 0.$$

5.2   Hitting Times

Let Y be the jump chain associated with X, and $A \subset I$. Se $T_A = \inf\{t > 0 \mid X_t \in A\}$ be the *hitting time*, and $H_A = \inf\{n \geq 0 \mid Y_n \in A\}$.

We also define the *hitting probability* $h_a(x) = \mathbb{P}_x(T_A < \infty)$, and *mean hitting probability* $k_A(x) = \mathbb{E}[T_A]$.

Note the hitting probability of X is the same as that for Y, but the mean hitting times will differ.

Theorem 5.1. $(h_A(x))_{x \in I}$ and $(k_A(x))_{x \in I}$ are the minimal non-negative solutions to

$$\begin{cases} h_A(x) = 1 & x \in A, \\ Qh_A(x) = \sum_y q_{xy} h_A(y) = 0 & x \notin A, \end{cases}$$

$$\begin{cases} k_A(x) = 0 & x \in A, \\ Qk_A(x) = \sum_y q_{xy} k_A(y) = -1 & x \notin A, \end{cases}$$

assuming $q_x > 0$ *for all* $x \notin A$.

Proof: For h, the hitting probability is the same as those for the jump chain, so $h_A(x) = 1$ for all $x \in A$, and

$$h_A(x) = \sum_{y \neq x} h_A(y) p_{xy} = \sum_{y \neq x} h_A(y) \frac{q_{xy}}{q_x}.$$

Hence we can rewrite this as

$$\sum_y h_A(y) q_{xy} = 0.$$

Now we consider the mean hitting times, k. If $x \in A$, then $T_A = 0$ so

$k_A(x) = 0$. Let $x \notin A$. Then $J_1 < T_A$, and hence

$$k_A(x) = \mathbb{E}_x[T_A] = \mathbb{E}_x[J_1] + \mathbb{E}_x[T_A - J_1] = \mathbb{E}_x[J_1] + \sum_{y \neq x} \mathbb{E}_x[T_a - J_1 \mid Y_1 = y] p_{xy}$$

$$= \frac{1}{q_x} + \sum_{y \neq x} \mathbb{E}_y[T_A] \frac{q_{xy}}{q_x}$$

$$= \frac{1}{q_x} + \sum_{y \neq x} k_A(y) \frac{q_{xy}}{q_x}.$$

Rearranging, we get
$$\sum_y q_{xy} k_A(x) = -1.$$

The minimality of the solutions is exactly as in the discrete time chain.

5.3   Recurrence and Transience

Definition 5.2. The state x is called *recurrent* if
$$\mathbb{P}(\sup\{t \mid X_t = x\} = \infty) = 1.$$

The state x is called *transient* if
$$\mathbb{P}(\sup\{t \mid X_t = x\} = \infty) = 0.$$

Remark. If X explodes with positive probability starting from x, i.e. if $\mathbb{P}(\zeta < \infty) > 0$, then
$$\mathbb{P}(\sup\{t \mid X_t = x\} < \infty) \geq \mathbb{P}(\zeta < \infty) > 0,$$

so x cannot be recurrent.

Theorem 5.2. *Let X be* Markov(Q) *with jump chain Y. Then,*

(a) *If x is recurrent for Y, then x is recurrent for X.*

(b) *If x is transient for Y, then x is transient for X.*

(c) *Every state is either recurrent or transient.*

(d) *Recurrence and transience are class properties.*

Proof: Note that (a) and (b) and results for discrete chains imply (c) and (d).

For (a), suppose x is recurrent for Y and $X_0 = x$. Then X is not explosive, i.e. $\mathbb{P}(\zeta = \infty) = 1$, and $J_n \to \infty$ with probability one, starting from x. Since $X_{J_n} = Y_n$ for all n, and Y visits x infinitely many times,

$$\mathbb{P}(\sup\{t \mid X_t = x\} = \infty) = 1.$$

For (b), if x is transient for Y, then $q_x > 0$. Also, there is almost surely a last visit to x for Y, i.e. $N = \sup\{n \mid Y_n = x\} < \infty$.

Also, $J_{N+1} < \infty$ almost surely, as $q_x > 0$, and if $t \in \{S \mid X_s = x\}$, then $t < J_{N+1}$. Hence

$$\sup\{s \mid X_s = x\} \le J_{N+1} < \infty.$$

We have an analogue for the discrete time result on recurrence and transience, namely

$$\sum_n p_{xx}(n) = \infty \iff x \text{ is recurrent.}$$

Theorem 5.3. *x is recurrent for X if and only if*

$$\int_0^\infty p_{xx}(t)\,\mathrm{d}t = \infty.$$

Proof: If $q_x = 0$, then x is absorbing, hence recurrent, and $p_{xx}(t) = 1$, so the integral is infinite. Assume $q_x > 0$. Then,

$$\int_0^\infty p_{xx}(t)\,\mathrm{d}t = \int_0^\infty \mathbb{E}_x\big[\mathbb{1}(X_t = x)\big]\,\mathrm{d}t$$

$$= \mathbb{E}_x\left[\int_0^\infty \mathbb{1}(X_t = x)\,\mathrm{d}t\right]$$

$$= \mathbb{E}_x\left[\sum_{n=0}^\infty \mathbb{1}(Y_n = x)S_{n+1}\right]$$

$$= \sum_{n=0}^\infty \mathbb{E}_x\big[\mathbb{1}(Y_n = x)S_{n+1}\big]$$

$$= \sum_{n=0}^\infty \mathbb{P}_x(Y_n = x)\mathbb{E}_x[S_{n+1} \mid Y_n = x]$$

$$= \sum_{n=0}^\infty p_{xx}(n)\frac{1}{q_x} = \frac{1}{q_x}\sum_{n=0}^\infty p_{xx}(n).$$

So the result follows from the discrete time case and the previous theorem.

5.4   Invariant Distributions

For a discrete Markov chain Y, π is an *invariant measure* for Y if and only if $\pi P = \pi$. Moreover if $\sum_i \pi_i = 1$, then π is called an *invariant distribution*. If $Y_0 \sim \pi$, then $Y_n \sim \pi$ for $n \geq 1$.

In fact recall:

Theorem 5.4. *If Y is a discrete time Markov chain, irreducible and recurrent. Then,*

$$\nu^x(y) = \mathbb{E}_x\left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y)\right],$$

where $H_x = \inf\{n \geq 1 \mid Y_n = x\}$. THen ν^x is an invariant measure, and $0 < \nu^x(y) < \infty$ for all y, with $\nu^x(x) = 1$.

Theorem 5.5. *Let Y be irreducible. If λ is any invariant measure with $\lambda(x) = 1$, then*

$$\lambda(y) \geq \nu^x(y).$$

If Y is recurrent then $\lambda(y) = \nu^x(y)$.

Definition 5.3. Let $X \sim \text{Markov}(Q)$ and let λ be a measure. It is calledn *invariant* or *infinitesimally invariant* if $\lambda Q = 0$.

Lemma 5.1. *If $|I|$ is finite, then $\lambda Q = 0$ if and only if $\lambda P(s) = \lambda$ for all $s \geq 0$.*

Proof: If $|I|$ is finite, then $P(s) = e^{sQ}$, so if $\lambda Q = 0$, then

$$\lambda P(s) = \lambda e^{sQ} = \lambda \sum_{k=0}^{\infty} \frac{(sQ)^k}{k!}$$

$$= \lambda I + \lambda(sQ) + \lambda\frac{(sQ)^2}{2!} + \lambda\frac{(sQ)^3}{3!} + \cdots$$

$$= \lambda + s\lambda Q + s^2\frac{\lambda Q^2}{2!} + s^3\frac{\lambda Q^3}{3!} + \cdots$$

$$= \lambda.$$

Now if $\lambda P(s) = \lambda$ for all s, then since

$$Q = \frac{\mathrm{d}}{\mathrm{d}s}P(s)\bigg|_{s=0},$$

then

$$\lambda Q = \frac{\mathrm{d}}{\mathrm{d}s}\lambda P(s)\bigg|_{s=0} = 0.$$

Lemma 5.2. *Let X be* Markov(Q) *and Y its jump chain. Then π is invariant for X if and only if μ defined by $\mu_x = q_x \pi_x$ is invariant for Y.*

Proof: We want to show $\pi Q = 0$ if and only if $\mu P = \mu$. Since $q_x(p_{xy} - \delta_x) = q_{xy}$, we have

$$
\begin{aligned}
(\pi Q)_y = \sum_{x \in I} \pi_x q_{xy} &= \sum_{x \in I} \pi_x q_x (p_{xy} - \delta_{xy}) \\
&= \mu_x (p_{xy} - \delta_{xy}) = \sum_x \mu_x p_{xy} - \mu_y \\
&= (\mu P)_y - \mu_y.
\end{aligned}
$$

Theorem 5.6. *Let X be irreducible and recurrent with generator Q. Then X has an invariant measure, which is unique up to scalar multiplication.*

Proof: Assume $|I| > 1$. Then by irreducibility, $q_x > 0$ for all x. For Y, $\nu^x(y)$ is an invariant measure which is unique up to scalar multiplication. Hence, $\frac{\nu^x(y)}{q_y}$ is an invariant measure for X, which is also unique.

Definition 5.4. Let $T_x = \inf\{t \geq J_1 \mid X_t = x\}$ be the *first return time* to x.

Lemma 5.3. *Assume $q_y > 0$. Define*

$$\mu^x(y) = \mathbb{E}_x\left[\int_0^{T_x} \mathbb{1}(X_t = y)\,\mathrm{d}t\right].$$

Then we have

$$\mu^x(y) = \frac{\nu^x(y)}{q_y}.$$

Proof:

$$\mu^x(y) = \mathbb{E}_x\left[\int_0^{T_x} \mathbb{1}(X_t = y)\,\mathrm{d}t\right]$$

$$= \mathbb{E}_x\left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y)S_{n+1}\right]$$

$$= \mathbb{E}_x\left[\sum_{n=0}^{\infty} S_{n+1}\mathbb{1}(Y_n = y, n \le H_x - 1)\right]$$

$$= \sum_{n=0}^{\infty} \mathbb{E}_x[S_{n+1} \mid Y_n = y, n \le H_x - 1]\mathbb{P}_x(Y_n = y, n \le H_x - 1).$$

By the Markov property, this equals

$$\sum_{n=0}^{\infty} \mathbb{E}_x[S_{n+1} \mid Y_n = y]\mathbb{P}_x(Y_n = y, n \le H_x - 1)$$

$$= \sum_{n=0}^{\infty} \frac{1}{q_y}\mathbb{P}_x(Y_n = y, n \le H_x - 1)$$

$$= \frac{1}{q_y}\sum_{n=0}^{\infty} \mathbb{E}_x\big[\mathbb{1}(Y_n = y, n \le H_x - 1)\big]$$

$$= \frac{1}{q_y}\mathbb{E}_x\left[\sum_{n=0}^{\infty} \mathbb{1}(Y_n = y, n \le H_x - 1)\right]$$

$$= \frac{1}{q_y}\mathbb{E}_x\left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y)\right] = \frac{\nu^x(y)}{q_y}.$$

Definition 5.5. A recurrent state x is called *positive recurrent* if

$$m_x = \mathbb{E}_x[T_x] < \infty.$$

Else, the state is *null recurrent*.

Theorem 5.7. *Let $X \sim \mathrm{Markov}(Q)$ be irreducible. Then the following are equivalent:*

(a) *Every state is positive recurrent.*

(b) *Some state is positive recurrent.*

(c) *X is non-explosive and has an invariant distribution λ.*

Moreover, when (c) *holds,*

$$\lambda(x) = \frac{1}{q_x m_x}.$$

Proof: Clearly (a) implies (b). We show (b) implies (c). Assume that $q_x > 0$ for all x. Let x be a positive recurrent state. Then all states are recurrent, so Y is recurrent and hence X is non-explosive.

As Y is recurrent, ν^x is an invariant measure for Y. So $\mu^x(y) = \frac{\nu^x(y)}{q_y}$ is invariant for X. Also,

$$\mu^x(y) = \mathbb{E}_x\left[\int_0^{T_x} \mathbb{1}(X_t = y)\,\mathrm{d}t\right].$$

Hence,

$$\sum_{y\in I}\mu^x(y) = \mathbb{E}_x\left[\int_0^{T_x}\sum_{y\in I}\mathbb{1}(X_t = y)\,\mathrm{d}t\right] = \mathbb{E}_x[T_x] < \infty.$$

Hence we have an invariant distribution of the form

$$\frac{\mu^x(y)}{\mathbb{E}_x[T_x]}.$$

Finally, we show (c) implies (a). From our lemma, the measure

$$\beta(y) = \lambda(y)q_y$$

is an invariant measure for Y. Since $\sum \lambda(y) = 1$, $\lambda(x) > 0$ for some x. Since Y is irreducible, for any $y \in I$, x leads to y, in say n steps. Since $\beta P^n = \beta$, we have $\lambda(y) > 0$.

Fix some $x \in I$. Then $\lambda(x) > 0$. Define

$$a^x(y) = \frac{\beta(y)}{\lambda(x)q_x},$$

for all $y \in I$. This is invariant for Y, and $a^x(y) = 1$. By the theorem for discrete time chains, $a^x(y) \geq \nu^x(y)$ for all $y \in I$. Also if we define

$$\mu^x(y) = \mathbb{E}_x\left[\int_0^{T_x}\mathbb{1}(X_t = y)\,\mathrm{d}t\right],$$

then we know
$$\mu^x(y) = \frac{\nu^x(y)}{q_y},$$
and $\sum \mu^x(y) = \mathbb{E}_x[T_x] = m_x$. Then,

$$m_x = \sum_y \mu^x(y) = \sum_y \frac{\nu^x(y)}{q_y}$$

$$\leq \sum_y \frac{a^x(y)}{q_y} = \sum_y \frac{\beta(y)}{\lambda(x)q_x q_y}$$

$$= \sum_y \frac{\lambda(y)q_y}{\lambda(x)q_x q_y} = \frac{1}{\lambda(x)q_x} \sum_y \lambda(y)$$

$$= \frac{1}{\lambda(x)q_x} < \infty.$$

Hence x is positive recurrent, so all states are positive recurrent.

Also if (c) holds, then X is recurrent, so Y is recurrent, so $a^x(y) = \nu^x(y)$. Hence
$$m_x = \frac{1}{\lambda(x)q_x}.$$

Example 5.1.

On $\mathbb{Z}^+$, consider the birth and death process with $q_{i,i+1} = \lambda q_i$, $q_{i,i-1} = \mu q_i$ and $q_{ii} = -(\lambda + \mu)q_i$.

Then $(\lambda/\mu)^L$ is invariant for Y, so

$$\pi_i = \frac{1}{q_i}\left(\frac{\lambda}{\mu}\right)^i$$

is invariant for X. Take $q_i = 2^i$ and $\lambda = \frac{3\mu}{2}$. Then,

$$\pi_i = \left(\frac{3}{4}\right)^i$$

is invariant for X. This is summable, so X has an invariant distribution. However, X is not positive recurrent as Y is transient. Therefore X is transient.

In order to not contradict our previous theorem, this means X is explosive.

Lemma 5.4. *Fix $t > 0$ and set $Z_n = X_{nt}$, where (Z_n) is a discrete-time Markov chain.*

Then x is recurrent for X if and only if x is recurrent for Z.

This is an exercise in the example sheets.

Theorem 5.8. *Let $X \sim \mathrm{Markov}(Q)$ be recurrent, irreducible, and λ be a measure. Then*
$$\lambda Q = 0 \iff \lambda P(s) = \lambda$$
for all $s > 0$.

> **Proof:**   Any measure λ such that $\lambda Q = 0$ is unique (up to scalar multiplication).
>
> Similarly any measure λ such that $\lambda P(s) = \lambda$ is unique (up to scalar multiplication). Then $(X_n)_{n=0}^{\infty}$ is a discrete time chain with transition matrix $P(1)$, and moreover it is irreducible as X is irreducible, and recurrent from our previous lemma, hence the invariant measure λ is unique.
>
> Then it is enough to show that $\mu^x Q = 0$ and $\mu^x P(s) = \mu^x$ for all s, where
>
> $$\mu^x(y) = \mathbb{E}_x \left[\int_0^{T_x} \mathbb{1}(X_t = y)\, \mathrm{d}t \right].$$
>
> Moreover $\mu^x(y) = \frac{\nu^x(y)}{q_y}$, and since Y is recurrent $\nu^x(y)$ is an invariant measure for Y. So $\mu^x(y)$ is an invariant measure for X, meaning $\mu^x Q = 0$.
>
> Also, by the Strong Markov property,
>
> $$\mathbb{E}_x \left[\int_0^s \mathbb{1}(X_t = y)\, \mathrm{d}t \right] = \mathbb{E}_x \left[\int_{T_x}^{T_x + s} \mathbb{1}(X_t = y)\, \mathrm{d}t \right].$$

Thus,

$$\mu^x(y) = \mathbb{E}_x\left[\int_0^{T_x} \mathbb{1}(X_t = y)\,\mathrm{d}t\right]$$

$$= \mathbb{E}_x\left[\int_0^s \mathbb{1}(X_t = y)\,\mathrm{d}t\right] + \mathbb{E}_x\left[\int_s^{T_x} \mathbb{1}(X_t = y)\,\mathrm{d}t\right]$$

$$= \mathbb{E}_x\left[\int_{T_x}^{T_x+s} \mathbb{1}(X_t = y)\,\mathrm{d}t\right] + \mathbb{E}_x\left[\int_s^{T_x} \mathbb{1}(X_t = y)\,\mathrm{d}t\right]$$

$$= \mathbb{E}_x\left[\int_s^{T_x+s} \mathbb{1}(X_t = y)\,\mathrm{d}t\right]$$

$$= \mathbb{E}_x\left[\int_0^\infty \mathbb{1}(X_{u+s} = y, u < T_x)\,\mathrm{d}u\right]$$

$$= \int_0^\infty \mathbb{P}_x(X_{u+s} = y, u < T_x)\,\mathrm{d}u$$

$$= \int_0^\infty \sum_{z \in I} \mathbb{P}_x(X_u = z, X_{u+s} = y, u < T_x)\,\mathrm{d}u$$

$$= \sum_{z \in I} P_{zy}(s)\mathbb{E}_x\left[\int_0^{T_x} \mathbb{1}(X_u = z)\,\mathrm{d}u\right]$$

$$= \sum_{z \in I} \mu^x(z)P_{zy}(s).$$

Hence $\mu^x = \mu^x P(s)$ as required.

5.5   Convergence to Equilibrium

Theorem 5.9. *Let $X \sim \mathrm{Markov}(Q)$ be irreducible and non-explosive, and let λ be an invariant distribution. Then for all $x, y \in S$,*

$$P_{xy}(t) \to \lambda(y),$$

as $t \to \infty$.

Lemma 5.5. *For the semigroup $P(t)$ and for all $t \geq 0$, $h \geq 0$,*

$$|P_{xy}(t + h) - P_{xy}(t)| \leq 1 - e^{-q_x h} \leq q_x h.$$

Proof: The latter inequality is easy to show, from $e^x \geq 1 + x$. For the former, note

$$
\begin{aligned}
|P_{xy}(t+h) - P_{xy}(t)| &= \left| \sum_{z \in I} P_{xz}(h) P_{zy}(t) - P_{xy}(t) \right| \\
&= \left| \sum_{z \neq x} P_{xz}(h) P_{zy}(t) - P_{xy}(t)(1 - p_{xx}(h)) \right| \\
&\leq 1 - p_{xx}(h) = \mathbb{P}_x(X(h) \neq x) \\
&\leq \mathbb{P}_x(J_1 \leq h) = 1 - e^{-q_x h}.
\end{aligned}
$$

Now we return to proving our theorem.

Proof: Fix $\varepsilon > 0$ and $h > 0$ such that $q_x h < \varepsilon/2$. Consider the Markov chain $(Z_n) = (X_{nh})_{n=0}^{\infty}$. Then Z_n is irreducible and aperiodic. As X is positive recurrent, $\lambda P(h) = \lambda$, meaning λ is an invariant distribution for (Z_n).

By discrete-time Markov chain results, for all x, y,

$$
P_{xy}(nh) \overset{n \to \infty}{\to} \lambda(y)'.
$$

Hence there exists some n_0 such that for all $n \geq n_0$, $|p_{xy}(nh) - \lambda(y)| \leq \varepsilon/2$. Let $t \geq n_0 h$. Then there exists $n \geq n_0$ such that $nh \leq t < (n+1)h$, so from our previous lemma,

$$
\begin{aligned}
|p_{xy}(t) - \lambda(y)| &\leq |p_{xy}(t) - p_{xy}(nh)| - |p_{xy}(nh) - \lambda(y)| \\
&\leq q_x(t - nh) + \frac{\varepsilon}{2} \leq q_x h + \frac{\varepsilon}{2} \\
&\leq \varepsilon.
\end{aligned}
$$

5.6   Ergodic Theory

Theorem 5.10. *Let $X \sim \mathrm{Markov}(\lambda, Q)$ be irreducible. Then as $t \to \infty$,*

$$
\frac{1}{t} \int_0^t \mathbb{1}(X_s = x) \, \mathrm{d}s \to \frac{1}{q_x m_x},
$$

with probability 1.

If X is positive recurrent and π is the unique invariant distribution, if $f : I \to \mathbb{R}$

is bounded, then

$$\frac{1}{t}\int_0^t f(X_s)\,\mathrm{d}s \to \sum_{x\in I} f(x)\pi(x).$$

The proof of this is skipped.

5.7   Reversibility

Theorem 5.11. *Let $X \sim \mathrm{Markov}(Q)$ be irreducible, non-explosive with invariant distribution π. Let $X_0 \sim \pi$. Fix $T > 0$ and set*

$$\hat{X}_t = X_{T-t},$$

for $0 \le t \le T$. Then $\hat{X} \sim \mathrm{Markov}(\hat{Q})$ and invariant distribution π, where

$$\hat{q}_{xy} = \pi(y)\frac{q_{yx}}{\pi(x)}.$$

Moreover, $\hat{Q}$ is irreducible and non-explosive.

Proof: Note that $\hat{Q}$ is a Q-matrix, as we can check $\hat{q}_{xy} \ge 0$ for all $x \ne y$, and

$$\sum_y \hat{q}_{xy} = \frac{1}{\pi(x)}\sum_y \pi(y)q_{yx} = \frac{1}{\pi(x)}(\pi Q)_x = 0.$$

Also $\hat{Q}$ is irreducible, and

$$(\pi\hat{Q})_x = \sum_x \pi(x)\hat{q}_{xy} = \sum_x \pi(y)q_{yx} = \pi(y)\sum_x q_{yx} = 0.$$

So π is invariant for $\hat{Q}$. Now let $0 = t_0 \le t_1 \le \cdots \le t_n = T$, $x_1, \ldots, x_n \in I$ and set $s_i = t_i - t_{i-1}$. Then, if we define

$$\hat{P}_{xy}(t) = \frac{\pi(y)}{\pi(x)}P_{yx}(t),$$

then we get

$$
\begin{aligned}
\mathbb{P}(\hat{X}_{t_0} &= x_0, \hat{X}_{t_1} = x_1, \ldots, \hat{X}_{t_n} = x_n)\\
&= \mathbb{P}(X_0 = x_n, \ldots, X_{T-t_1} = x_1, X_T = x_0)\\
&= \pi(x_n)P_{x_n x_{n-1}}(s_n)\cdots P_{x_1 x_0}(s_1)\\
&= \pi(x_n)\hat{P}_{x_{n-1}x_n}\frac{\pi(x_{n-1})}{\pi(x_n)}\cdots\hat{P}_{x_0 x_1}(s_1)\frac{\pi(x_0)}{\pi(x_1)}\\
&= \pi(x_0)\hat{P}_{x_0 x_1}(s_1)\cdots\hat{P}_{x_{n-1}x_n}(s_n).
\end{aligned}
$$

Hence $\hat{X}$ is Markov with transition semigroup $(\hat{P}(t))$, so now we show that $\hat{Q}$ generates $(\hat{P}(t))$. Hence we show $\hat{P}(t)$ is the minimal non-negative solution to the Kolmogorov backwards equation with $\hat{Q}$, or

$$(\hat{P}(t))' = \hat{Q}\hat{P}(t).$$

Indeed, we have

$$\hat{P}'_{xy}(t) = \frac{\pi(y)}{\pi(x)}P'_{yx}(t) = \frac{\pi(y)}{\pi(x)}\sum_z P_{yz}(t)q_{zx}$$

$$= \frac{1}{\pi(x)}\sum_z \pi(z)\hat{P}_{zy}(t)q_{zx}$$

$$= \sum_z \hat{q}_{xz}\hat{p}_{zy}(t) = (\hat{Q}\hat{P}(t))_{xy}.$$

To show minimality, suppose R is another solution to $R'(t) = \hat{Q}R(t)$. Then defining

$$\bar{R}_{xy}(t) = \frac{\pi(y)}{\pi(x)}R_{yx}(t),$$

as before we would get $\bar{R}'(t) = \bar{R}(t)Q$. Then as $P'(t) = P(t)Q$ and is minimal, $\bar{R} \geq P$. Hence $R \geq \hat{P}$.

Also, as X is irreducible and non-explosive with invariant distribution π, X is (positive) recurrent. Hence $\pi P(t) = \pi$ for all t. Thus,

$$\sum_y \hat{P}_{xy}(t) = \frac{1}{\pi(x)}\sum_y \pi(y)P_{yx}(t) = \frac{1}{\pi(x)}(\pi P(t))_x = \frac{1}{\pi(x)}\pi(x) = 1.$$

Hence,

$$1 = \sum_y \hat{P}_{xy}(t) = \sum_y \mathbb{P}_x(Z_t = y) = \sum_y \mathbb{P}_x(Z_t = y, t < \zeta) = \mathbb{P}_x(t < \zeta).$$

Now since $\mathbb{P}_x(\zeta > t) = 1$ for all t, we get $\mathbb{P}_x(\zeta = \infty) = 1$, meaning the process is non-explosive.

Definition 5.6. Let $X \sim \text{Markov}(Q)$. It is called *reversible* if for all $T > 0$, $(X_t)_{0 \leq t \leq T}$ and $(X_{T-t})_{0 \leq t \leq T}$ have the same distribution.

Definition 5.7. A measure λ and a Q-matrix Q are said to be in *detailed balance*

if for all x, y,

$$\lambda(x)q_{xy} = \lambda(y)q_{yx}.$$

Lemma 5.6. *If Q and λ are in detailed balance, then λ is invariant for Q.*

Proof:

$$(\lambda Q)_y = \sum_x \lambda(x)q_{xy} = \sum_x \lambda(y)q_{yx} = \lambda(y)\sum_x q_{yx} = 0.$$

Remark. To find an invariant measure, first check the detailed balance equations.

Lemma 5.7. *Let $X \sim \text{Markov}(Q)$ be irreducible and non-explosive, π a distribution, and $X_0 \sim \pi$. Then π and Q are in detailed balance if and only if $(X_t)_{t \geq 0}$ is reversible.*

Proof: X is reversible if and only if $Q = \hat{Q}$ and π is an invariant distribution. But this holds if and only if

$$\pi(x)q_{xy} = \pi(y)q_{yx},$$

i.e. π and Q are in detailed balance.

Definition 5.8. A *birth and death chain* X is a continuous time Markov chain on $\mathbb{N} = \{0, 1, 2, \ldots\}$ with, for $x \geq 1$,

$$q_{x,x-1} = \mu_x, \qquad q_{x,x+1} = \lambda_x, \qquad q_{x,x} = -\mu_x - \lambda_x,$$

and for $x = 0$, $q_{0,1} = \lambda_0$ and $q_{0,0} = -\lambda_0$.

Lemma 5.8. *A measure π is invariant for a birth and death chain if and only if it solves the detailed balance equation.*

Proof: We have seen that if it solves the detailed balance equation, then it is invariant.

For the converse, let π be an invariant measure. Then $\pi Q = 0$, so for all $j \geq 1$,

$$\begin{aligned}
(\pi Q)_j &= 0 \\
\Longleftrightarrow \pi_{j-1}q_{j-1,j} &+ \pi_j q_{j,j} + \pi_{j+1}q_{j+1,j} = 0 \\
\Longrightarrow \pi_{j-1}\lambda_{j-1} &+ \pi_{j+1}\mu_{j+1} - \pi_j(\lambda_j + \mu_j) = 0 \\
\Longrightarrow \pi_{j+1}\mu_{j+1} &- \pi_j\lambda_j = \pi_j\mu_j - \pi_{j-1}\lambda_{j-1}.
\end{aligned}$$

For $j = 0$, we get $\pi_1\mu_1 - \pi_0\mu_0 = 0$. Using induction, we get $\pi_{j+1}\mu_{j_1} = \pi_j\lambda_j$.

6   Queueing Theory

Queueing theory is the study of a continuous Markov process, where

- Customers arrive, to get served.

- Serving happens at some rate.

- After getting served, the customers leave.

Applications of queueing theory include telecommunication, traffic control, industrial engineering and project management.

Shout out Kendall, Kingman, Frank Kelly (big ups Cambridge), Burke, Jackson, Khinchin? and some others.

Examples of queueing theory questions are:

Q. What is the equilibrium queue length?

Q. What is the busy period?

Q. What is the time spent by a customer in the queue? What is the waiting time (including the service time)?

We use $M/G/k$ notation. Here,

- M is the Markovian arrival: meaning that the customers arrive according to a Poisson process of rate λ.

- G is the general distribution, which is the service time distribution (which is iid). If we swap the G for a M, that indicates that we have iid $\text{Exp}(\mu)$ service times.

- k is the number of servers (we will consider $k = 1$ or $k = \infty$).

Let X_t be the queue length (including customers being served) at time t. Then $I = \{0, 1, 2, \ldots\}$.

For $M/M/1$ and $M/M/\infty$, X is a Markov chain. Moreover, since the number of customers increases or decreases by one, X is a birth and death chain.

For $M/M/1$, we have $q_{i,i+1} = \lambda$, $q_{i,i-1} = \mu$.

For $M/M/\infty$, we have $q_{i,i+1} = \lambda$, $q_{i,i-1} = i\mu$.

Theorem 6.1. *Consider a $M/M/1$ process with rates λ, μ. Let $\rho = \lambda/\mu$. Then the queue length X is:*

- *Transient $\iff \rho > 1$.*

- *Recurrent $\iff \rho \leq 1$.*

- *Positive recurrent $\iff \rho < 1$.*

In the positive recurrent case, the invariant distribution is

$$\pi(n) = (1 - \rho)\rho^n.$$

If $\rho < 1$ and $X_0 \sim \pi$, then the wait time for a customer that arrives at time t is $\mathrm{Exp}(\mu - \lambda)$.

Proof: The jump chain Y is a biased SRW on $\mathbb{N}$ (with reflection at the origin). Thus Y (and hence X) is transient if $\lambda > \mu$, and recurrent if $\lambda \leq \mu$.

There is no explosion, as $q_i = (\lambda + \mu)$ is bounded, hence positive recurrence is equivalent to having an invariant distribution. As X is a birth and death chain, an invariant measure solves the detailed balance equation.

Therefore $\pi(n)\lambda = \pi(n + 1)\mu$, or

$$\pi(n + 1) = \pi(n)\frac{\lambda}{\mu} = \cdots = \pi(0)\left(\frac{\lambda}{\mu}\right)^{n+1}.$$

Hence π is normalizable if and only if $\frac{\lambda}{\mu} = \rho < 1$. We can normalise it to $\pi(n) = (1 - \rho)\rho^n$.

Hence ρ is the distribution of a (shifted) geometric random variable with parameter $1 - \rho$, i.e. $\pi \sim Z - 1$, where $Z \sim \mathrm{Geo}(1 - \rho)$.

If $\rho < 1$ and $X_0 \sim \pi$, then $X_t \sim \pi$, as X is recurrent. The wait time W of a customer arriving at time t is

$$W = \sum_{i=1}^{X_t + 1} T_i,$$

where $T_i \sim \mathrm{Exp}(\mu)$, independent of X_t. As $X_t + 1 \sim \mathrm{Geo}(1 - \rho)$ independent of (T_i), we get

$$W \sim \mathrm{Exp}(\mu(1 - \rho)) = \mathrm{Exp}(\mu - \lambda).$$

The expected queue length at equilibrium is

$$\mathbb{E}_\pi[X_t] = \mathbb{E}_\pi[Z - 1] = \mathbb{E}_\pi[Z] - 1 = \frac{1}{1 - \rho} - 1$$

$$= \frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda}.$$

Now we look at $M/M/\infty$ processes.

Theorem 6.2. *The queue length X is positive recurrent, with invariant distribution* Poisson(ρ), *with $\rho = \frac{\lambda}{\mu}$.*

Proof: As X is a birth-and-death process, we just solve the detailed balance equation:

$$\lambda \pi_{n-1} = n\mu \pi_n \implies \pi_n = \frac{1}{n} \cdot \frac{\lambda}{\mu} \pi_{n-1} = \cdots = \frac{1}{n!} \left(\frac{\lambda}{\mu} \right)^n \pi_0.$$

This is always normalizable with

$$\pi_n = e^{-\lambda/\mu} \left(\frac{\lambda}{\mu} \right)^n \frac{1}{n!},$$

hence $\pi \sim$ Poisson(ρ). We also need to show X is non-explosive to conclude X is positive recurrent.

One such way is to show the jump chain Y is recurrent. In fact, we will show it is positive recurrent, hence that is has an invariant measure. Define $\mu_i = \pi_i q_i$. Then μ is an invariant measure for Y, so it is enough to check that μ is normalizable:

$$\mu_i = (i\mu + \lambda)e^{-\rho}\frac{\rho^i}{i!} = \rho\mu \left(e^{-\rho} \frac{\rho^{i-1}}{i!}(i + \rho) \right),$$

which is summable as

$$\sum \frac{\rho^{i-1}}{i!}(i + \rho) = \sum \frac{\rho^{i-1}}{(i-1)!} + \sum \frac{\rho^i}{i!} < \infty.$$

6.1   Burke's theorem

Let A and D denote the arrival and departure processes associated with a queue. Then,

- A increases by 1 when X increases by 1.

- D increases by 1 when X decreases by 1.

So $X_t = X_0 + A_t - D_t$. Then we are given that A is a Poisson process of rate λ.

Remark. A Poisson process does not have an invariant distribution, but still has

the following time-reversal property: if N is a Poisson process of rate λ, then for any $T > 0$, $\hat{N}_t = N_T - N_{T-t}$ is again a Poisson process of rate λ on $[0, T]$.

Proof: Conditioned on $N_T = n$, the distribution of jump times is

$$\frac{n!}{T^n} \mathbb{1}(0 \leq t_1 \leq \cdots \leq t_n \leq T),$$

which gives the result, and also we can use the fact the independent increments are Poisson.

Theorem 6.3 (Burke). *Consider an $M/M/1$ queue with $\mu > \lambda > 0$, or an $M/M/\infty$ queue with $\mu, \lambda > 0$. At equilibrium (i.e. $X_0 \sim \pi$), D is a Poisson process of rate λ and X_t is independent of $(D_s \mid s \leq t)$.*

Remark.

1. $X_0 \sim \pi$ is essential. If $X_0 = 5$ in an $M/M/1$ queue, then the first departure happens at $\text{Exp}(\mu)$ time, and not $\text{Exp}(\lambda)$.

2. The process $(X_s \mid s \leq t)$ and $(D_s \mid s \leq t)$ are not independent. This is because D jumps exactly when X decreases by one.

Proof: As X is a birth-and-death process, π satisfies the detailed balance equation. That is, if $X_0 \sim \pi$, then X is reversible. Then for a fixed $T > 0$ with $\hat{X}_t = X_{T-t}$, we have

$$(\hat{X}_t \mid 0 \leq t \leq T) \stackrel{d}{=} (X_t \mid 0 \leq t \leq T).$$

The arrival process $\hat{A}$ for $\hat{X}$ (till time T) is Poisson process λ. But $\hat{A}_t = D_T - D_{T-t}$ is a Poisson process of rate λ, and since the time reversal of a Poisson process on $[0, T]$ is again a Poisson process on $[0, T]$, this implies $(D_t \mid 0 \leq t \leq T)$ is a Poisson process of rate λ on $[0, T]$.

Since $T > 0$ is arbitrary, this determines the finite dimensional distribution of D, and hence determines the distribution of D, meaning $D \sim PP(\lambda)$.

As X_0 is independent of $(A_s \mid 0 \leq s \leq T)$, $\hat{X}_0$ is independent of $(\hat{A}_s \mid 0 \leq s \leq T)$, meaning X_T is independent of $(D_s \mid 0 \leq s \leq T)$.

6.2   Queues in Tandem

Suppose that there is an $M/M/1$ queue with parameters λ and μ_1. After a customer is served, they immediately join a second single-server queue with service rate μ_2.

Let X and Y denote the queue lengths of the two queues respectively. For (X, Y), $I = \{0, 1, 2, \ldots\}^2$, and the rates are

$$
(m, n) \to \begin{cases} (m+1, n) & \text{rate } \lambda, \\ (m-1, n+1) & \text{rate } \mu_1, \\ (m, n-1) & \text{rate } \mu_2. \end{cases}
$$

Theorem 6.4. (X, Y) *is positive recurrent if and only if* $\lambda < \mu_1$ *and* $\lambda < \mu_2$, *and then the invariant distribution is given by*

$$
\pi(m, n) = (1 - \rho_1)\rho_1^m (1 - \rho_2)\rho_2^n
$$

with $\rho_1 = \frac{\lambda}{\mu_1}$, $\rho_2 = \frac{\lambda}{\mu_2}$. *Hence, at equilibrium,* X_t *and* Y_t *are independent.*

> **Proof:**   We can directly check that $\pi Q = 0$. As the rates are bounded, (X, Y) is non-explosive.
>
> However, we can find a more elegant proof. Note that the marginal X is an M/M/1 queue. Thus X is positive recurrent if and only if $\lambda < \mu_1$ with invariant distribution
> $$
> \pi^1(m) = (1 - \rho_1)\rho_1^m.
> $$
>
> By Burke's theorem, if $X_0 \sim \pi^1$, then the departure process of the first queue is $PP(\lambda)$, which is the arrival process for the second queue. So the second queue is again M/M/1, with rates (λ, μ_2), with invariant distribution
> $$
> \pi^2(n) = (1 - \rho_2)\rho_2^n,
> $$
> if $\lambda < \mu_2$.
>
> So if $X_0 \sim \pi^1$ and $Y_0 \sim \pi^2$ are independent, then $X_t \sim \pi^1$ (as X is recurrent), and also by Burke's theorem, X_t is independent of $(D_s \mid s \leq t)$, and also independent of Y_0.
>
> Hence X_t is independent of Y_t. Also $Y_t \sim \pi^2$ (as Y is recurrent), so $(X_t, Y_t) \sim (\pi^1, \pi^2) = \pi$. Then π is invariant for (X, Y), from the next proposition.

Proposition 6.1. *Let* Z *be irreducible,* π *a distribution, and* $\pi P(t) = \pi$ *for all* t. *Then* π *is invariant for* Z.

6.3   Jackson's Network

Consider a network of N single-server queues, with arrival rates λ_k and service rates μ_k, for $k = 1, 2, \ldots, N$.

After service, each customer in queue i moves to queue j with probability p_{ij}, or exits the system with probability $p_{i0} = 1 - \sum p_{ij}$. We assume that $p_{ii} = 0$, i.e. they don't return to the same queue, and $p_{i0} > 0$, and also that the system is irreducible, i.e. a customer in queue i has a positive probability of visiting queue j at some later time.

Thus $I = \{0, 1, 2, \ldots\}^N$, where if $x = (x_1, x_2, \ldots, x_N)$, then x_i is the number of customers in queue i.

If $n \in I$, i.e. $n = (n_1, n_2, \ldots, n_N)$ and $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, then

$$
\begin{cases}
q_{n,n+e_i} = \lambda_i, \\
q_{n,n-e_i+e_j} = \mu_i p_{ij}, \\
q_{n,n-e_i} = \mu_i p_{i0}.
\end{cases}
$$

Definition 6.1. We say a vector $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_N)$ satisfies the *traffic equation* if for all $1 \le i \le N$,

$$
\bar{\lambda}_i = \lambda_i + \sum_{j=1}^{N} \bar{\lambda}_j p_{ji} - \bar{\lambda}_i p_{ii}.
$$

Remark. $\bar{\lambda}_i$ is the effective arrival rate at queue i.

Lemma 6.1. *There exists a unique solution to the traffic equation.*

> **Proof:**   The uniqueness is an exercise in example sheet 3.
>
> For existence, let $p_{00} = 1$. Then $P + (p_{ij})^N$ is a stochastic matrix. The corresponding (discrete time) Markov chain (Z_n) is absorbing at 0. Thus the communicating class $\{1, 2, \ldots, N\}$ is *transient*.
>
> Thus is V_i is the number of visits to state i by Z, then starting fro Z_0, $\mathbb{E}[V_i] < \infty$ for all $i = 12, \ldots, N$.
>
> Let $\mathbb{P}(Z_0 = i) = \frac{\lambda_i}{\lambda}$ for $i = 1, 2, \ldots, N$, where λ is the sum of λ_i. Then, for

all $1, 2, \ldots, N$,

$$\mathbb{E}[V_i] = \mathbb{E}\left[\sum_{i=1}^{\infty} \mathbb{1}(Z_n = i)\right]$$

$$= \mathbb{P}(Z_0 = i) + \sum_{n=0}^{\infty} \mathbb{P}(Z_{n+1} = i)$$

$$= \frac{\lambda_i}{\lambda} + \sum_{j=1}^{N} p_{ji} \sum_{n=0}^{\infty} \mathbb{P}(Z_n = j)$$

$$= \frac{\lambda_i}{\lambda} + \sum_{j=1}^{N} p_{ji} \mathbb{E}[V_j].$$

Multiplying through by λ and setting $\bar{\lambda}_i = \lambda \mathbb{E}[V_i]$, we get that

$$\bar{\lambda}_i = \lambda_i + \sum_{j=1}^{N} \bar{\lambda}_j p_{ji}.$$

Theorem 6.5. *Assume that the traffic equations have the solution $\bar{\lambda}_i$ such that $\bar{\lambda}_i < \mu_i$ for all $i = 1, \ldots, N$. Then the Jackson Network is positive recurrent with invariant distribution*

$$\pi(n) = \sum_{i=1}^{N} (1 - \bar{\rho}_i) \bar{\rho}_i^{n_i},$$

where $\bar{\rho}_i = \frac{\bar{\lambda}_i}{\mu_i}$.

At equilibrium, the departure processes from each queue form independent Poisson processes with rates $\bar{\lambda}_i p_{i0}$.

Remark.

1. At equilibrium, the queue lengths X_t^i are independent for a fixed time t.

2. The equilibrium for Jackson's network is non-reversible, but there is partial reversibility.

Lemma 6.2 (Partial detailed balance). *Let X be a Markov process on I and π be a measure. Assume that for each $\in I$, there is a partition of $I \setminus \{x\}$ as*

$$I \setminus \{x\} = I_1^x \cup I_2^x \cup \cdots,$$

such that for all $i \geq i$,

$$\sum_{y \in I_i^x} \pi(x) q_{xy} = \sum_{y \in I_i^x} \pi(y) q_{yx}.$$

Index