Introduction

Schedule:

(I) Basic properties of continuous-time Markov Chains

(II) Qualitative properties of continuous time Markov Chains

(III) Queueing theory

(IV) Renewal theory

(V) Spatial Poisson processes

1 Some basic aspects of continuous-time Markov Chains

Definition. A sequence of random variables is called a *stochastic process* or *process*. The process $X = (X_n)_{n\geq 1}$ is called a discrete-time Markov Chain with state space I if for all $x_0, x_1, \ldots, x_n \in I$

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

If $\mathbb{P}(X_{n+1} = y | X_n = x)$ is independent of n, the chain is called *time-homogeneous*. We then write $P = (P_{x,y})_{x,y\in I}$ for the *transition matrix* where $P_{x,y} = \mathbb{P}(X_1 = y | X_0 = x)$. The data associated to every time-homogeneous Markov Chain is the transition matrix P and the initial distribution μ, i.e $\mathbb{P}(X_0 = x_0) = \mu(x_0)$.

From now on:

- I denotes a countable (or finite) state space.
- $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space on which all the relevant random variables are defined.

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

(a) for all $t \geq 0$, $X(t) = X_t$ is a random variable such that $X(t) : \Omega \to I$;

(b) for all $\omega \in \Omega$, $t \mapsto X_t(\omega)$ is right-continuous (right-continuous sample path). In our case this means for all $\omega \in \Omega$, for all $t \geq 0$, there exists $\varepsilon > 0$ (depending on ω, t) such that

$$X_t(\omega) = X_s(\omega)\ \forall s \in [t, t+\varepsilon].$$

Fact. A right-continuous random process is defined by its finite-dimensional distributions

$$\mathbb{P}(X_{t_0=i}, X_{t_1=i_1}, \ldots, X_{t_n} = i_{t_n}),\ n \geq 0,\ t_k \geq 0,\ i_k \in I.$$

For every $\omega \in \Omega$, the path $t \mapsto X_t(\omega)$ of a right-continuous process stays constant for a while. So there are 3 possibilities:

(i) The path makes infinitely many jumps overall but only finitely many in a given interval $[0, t]$.

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

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

Write $J_0 = 0, J_1, J_2, \ldots$ for the jump times and $S_1, S_2, \ldots$ for the holding times, defined by

$$J_0 = 0,\ J_{n+1} = \inf\{t \geq J_n : X_t \neq X_{J_n}\},$$

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

By right-continuity, $S_n > 0$ for all n. If $J_{n+1} = \infty$ for some n, we define $X_\infty = X_{J_n}$ as the final value, otherwise X_∞ is not defined. The explosion time ζ is defined by

$$\zeta = \sup(J_n) = \sum_{n=1}^{\infty} S_n.$$

We are not going to consider what happens to a chain after explosion. We thus set $X_t = \infty$ for all $t \geq \zeta$ (adjoining a new state '∞'). We call such a chain *minimal*.

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

Definition. A right-continuous random process $X = (X_t)_{t\geq 0}$ has the Markov property (and is called a continuous-time markov chain) if for all $i_1, i_2, \ldots, i_n \in I$ and $0 \leq t_1 < t_2 < \ldots < t_n$,

$$\mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_0} = i_0) = \mathbb{P}(X_{t_n} = i_n | X_{t_0} = i_0).$$

Remark. For all $h > 0$, $Y_n = X(hn)$ defines a discrete-time Markov Chain.

Definition. The transition probabilities are $P_{ij}(s, t) = \mathbb{P}(X_t = j | X_s = i)$, $s \leq t$, $i, j \in I$. It is called *time-homogeneous* if it depends on $t - s$ only, i.e

$$P_{ij}(s, t) = P_{i,j}(0, t - s).$$

In this case we just write $P_{ij}(t - s)$. As in the case of discrete time, a (time-homogeneous) Markov process is characterised by

1. Its initial distribution $\lambda_i = \mathbb{P}(X_0 = i)$, $i \in I$;
2. Its *family of transition matrices* $(P(t))_{t\geq 0} = (P_{ij}(t))_{t\geq 0}$.

The family $(P(t))_{t\geq 0}$ is called the *transition subgroup* of the MC.

A (time-homogeneous) Markov process is characterised by

- its initial distribution;
- its transition subgroup $(P(t))_{t\geq 0}$

$$(P(t))_{t\geq 0} = (P(t))_{\substack{i,j\in I\\ t\geq 0}} = (\mathbb{P}(X_t = j|X_0 = i))_{\substack{i,j\in I\\ t\geq 0}}$$

It is easy to see that

- $P(0)$ is the identity
- $P(t)$ is a stochastic matrix for all t (i.e rows sum to 1)
- $P(t+s) = P(t)P(s)\ \forall s,t$ (Chapman-Kolmogorov equation)

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

Holding times

Let X be a (right-continuous continuous-time time-homogeneous) Markov Chain on a countable state-space I.

Suppose X starts from $x \in I$. Question: how long does X stay in the state x?

Definition. We call S_x the *holding time* at state x ($S_x > 0$ by right-continuity).

Let $s, t \geq 0$. Then

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

Thus S_x has the memoryless property.

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 $\mathbb{P}(S > t + s | S > s) = \mathbb{P}(S > t)$ for all $s, t \geq 0$ if and only if S has the exponential distribution.*

Proof. It is easy to see the exponential distribution is memoryless. So we prove the other direction. 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 such that $F(1/n) = \mathbb{P}(S > 1/n) > 0$. Then $F(1) = F(1/n)^n > 0$. So we can set $F(1) = e^{-\lambda}$ for some $\lambda \geq 0$.

For $k \in \mathbb{N}$, $F(k) = F(1)^k = e^{-\lambda k}$. For p/q rational, $F(p/q) = F(1/q)^p = (F(1/q)^q)^{p/q} = F(1)^{p/q} = e^{-\lambda \frac{p}{q}}$.

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

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

So taking sequences of rationals approaching t, we have $F(t) = e^{-\lambda t}$. □

Poisson Process'

We are now going to look at the simplest (and most important) example of continuous time Markov Chains - the Poisson process.

Definition. Suppose $S_1, S_2, \ldots$ are iid random variables with $S_1 \sim \text{Exp}(\lambda)$. Define the *jump times* $J_0 = 0, J_1 = S_1$, $J_n = S_1 + \ldots + S_n$ for all n, and set $X_t = i$ if $J_i \leq t < J_{i+1}$. Then $I = \{0, 1, 2, \ldots\}$ and note that X is right-continuous and increasing. X is called a *Poisson process* of parameter/intensity λ. We sometimes refer to the jump times $(J_i)_{i \geq 1}$ as the *points* of the Poisson process, then X =number of points in $[0, t]$.

Theorem 1.2 (Markov property). *Let $(X_t)_{t \geq 0}$ be a Poisson process of intensity λ. Then for all $s \geq 0$, the process $(X_{s+t} - X_s)_{t \geq 0}$ is also a Poisson process of intensity λ, and is independent of $(X_t)_{0 \leq t \leq s}$.*

Proof. Set $Y_t = X_{t+s} - X_s$ for all $t \geq 0$. Let $i \in \{0, 1, 2, \ldots\}$ and condition on $\{X_s = i\}$, Then the jump times for the process Y are $J_{n+1} - s, J_{n+2} - s, \ldots$ and the holding times are

$$\begin{aligned} T_1 &= J_{n+1} - s = S_{i+1} - (s - J_i) \\ T_2 &= S_{i+2} \\ T_3 &= S_{i+3} \\ &\vdots \end{aligned}$$

Since $\{X_s = i\} = \{J_i \leq s\} \cap \{s < J_{i+1}\} = \{J_i \leq s\} \cap \{S_{i+1} > s - J_i\}$, conditional on $\{X_s, i\}$, by the memoryless property of the exponential distribution (and independence of S_{i+1} and J_i) we see that $T_1 \sim \text{Exp}(\lambda)$. Moreover the times J_j, $j \geq 2$ are independent of S_k, $k \leq i$ and hence independent of $(X_r)_{r \leq s}$, and they have iid $\text{Exp}(\lambda)$ distribution. Thus $((X_{s+t} - X_s))_{t \geq 0}$ is a Poisson process of parameter λ and is independent of $(X_t)_{0 \leq t \leq s}$. $\square$

Similar to this, one can show the Strong Markov property for a Poisson process of parameter λ. Recall 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 1.3 (Strong Markov property). *Let $(X_t)_{t \geq 0}$ be a Poisson process of parameter λ and T a stopping time. Then conditional on $T < \infty$, the process $(X_{T+t} - X_T)_{t \geq 0}$ is a Poisson process of parameter λ and independent of $(X_s)_{s \leq T}$.*

Theorem 1.4. *Let $(X_t)_{t\geq 0}$ be an increasing right-continuous process taking values in $\{0,1,2,\dots\}$ with $X_0=0$. Let $\lambda>0$. Then the following are equivalent*

(a) The holding times $S_1, S_2, \dots$ are iid $\mathrm{Exp}(\lambda)$ and the jump chain is given by $Y_n = n$ (i.e X is a poisson process of intensity λ)

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

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

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

Proof. First we show (a)$\Rightarrow$(b). If (a) holds, then by the Markov property, the increments are independent and stationary ($(X_{t+s}-X_s)_{t\geq 0} =^d (X_t - X_0)_{t\geq 0}$). Using stationarity we have (uniformly in t) as $h \to 0$,

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

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

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

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

Thus, $\frac{p_j(t+h)-p_j(t)}{h}=-\lambda p_j(t)+\lambda p_{j-1}(t)+o(1)$. Setting $s=t+h$, we get

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

In particular, $p_j(t)$ is continuous and differentiable with

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

Differentiating

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

For $j = 0$ we have $p_0(t+h) = p_0(t)(1 - \lambda h + o(h))$, i.e $p_0'(t) = -\lambda p_0(t)$ so $p_0(t) = e^{-\lambda t}$. Thus

$$p_1'(t) = -\lambda p_1(t) + \lambda e^{-\lambda t}, \text{ i.e } p_1(t) = \lambda t e^{-\lambda t}.$$

And by induction

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

i.e $X_t \sim \mathrm{Poi}(\lambda t)$.

Finally we show (c)⇒(a). We know X has independent stationary increments, We have for $t_1 \leq \ldots \leq t_k,\ n_1 \leq \ldots \leq n_k$,

$$\begin{aligned}
&\mathbb{P}(X_{t_1} = n_1, \ldots, X_{t_k} = n_k) \\
&= \mathbb{P}(X_{t_1} = n_1)\underbrace{\mathbb{P}(X_{t_2} - X_{t_1} = n_2 - n_1)}_{\sim \mathrm{Poi}(\lambda t_1)} \ldots \underbrace{\mathbb{P}(X_{t_k} - X_{t_{k-1}} = n_k - n_{k-1})}_{\sim \mathrm{Poi}(\lambda(t_2 - t_1))}.
\end{aligned}$$

So (c) determines the finite-dimensional distributions (f.d.d) of a right-continuous process X, hence (c) determines X. So (c)⇒(a).

Question: can we show (a)⇒(c) directly? Indeed note

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

□

Theorem 1.5 (Superposition)**.** *Let X and Y be two independent Poisson processes with parameters λ and μ respectively. Then $(Z_t)_{t\geq 0} = (X_t + Y_t)_{t\geq 0}$ is a Poisson process with parameter $\lambda + \mu$.*

Proof. We use (c) from the previous theorem. So Z has stationary independent increments. Also $Z_t \sim \mathrm{Poi}(\lambda t + \mu t)$. □

Theorem 1.6 (Thinning)**.** *Let X be a Poisson process with parameter λ. Let $(Z_i)_{i\geq 1}$ be a sequence of iid Bernouilli(p) random variables. Let Y be a Posisson prcess with values in $\{0, \ldots, \}$ which jumps at time t if and only if X_t jumps at time t and $Z_{X_t} = 1$.*

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

Proof. We shall 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

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

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

Hence Y is Poisson of paremeter λp. Clearly $X - Y$ is a thinning of X with Bernouilli parameter $1-p$, so $X - Y$ is Poisson of parameter $\lambda(1-p)$.

Now we show Y and $X - Y$ are independent. It is enough to show that the f.d.d of Y and $X - Y$ are independent, i.e if $0 \leq t_1 \leq t_2 \leq \ldots \leq t_k$, $n_1 \leq \ldots \leq n_k$ and $m_1 \leq \ldots \leq m_k$, then we want to prove

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

We will only show this for fixed t ($k = 1$) the general case follows similarly using independence of increments. We have

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

as required. ∎

Theorem 1.7. *Let X be a Poisson Process. Conditional on the event $(X_t = n)$, the jump times $J_1, J_2, \dots, J_n$ are distributed as the order statistics of n iid $U[0,t]$ random variables. That is, they have joint density*

$$f(t_1, \dots, t_n) = \frac{n!}{t^n} \mathbb{1}(0 \le t_1 \le \dots \le t_n \le t).$$

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

$$\lambda^{n+1} e^{-\lambda(S_1 + \dots + S_{n+1})} \mathbb{1}(S_i \ge 0 \ foralli).$$

Then the jump times $J_1 = S_1, J_2 = S_1 + S_2, \dots, J_{n+1} = S_1 + \dots + S_{n+1}$ have joint density

$$g(t_1, \dots, t_{n+1}) = \lambda^{n+1} e^{-\lambda t_{n+1}} \mathbb{1}(0 \le t_1 \le t_2 \le \dots t_{n+1}).$$

(Noting the Jacobian of the transformation is 1.) Now take $A \subseteq \mathbb{R}^n$ so

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

Note

$$\begin{aligned}
&\mathbb{P}((J_1, \dots, J_n) \in A, X_t = n) \\
&= \mathbb{P}((J_1, \dots, J_n) \in A, J_n \le t < J_{n+1}) \\
&= \int_{(t_1, \dots, t_{n+1}) \in A \times \mathbb{R}} g(t_1, \dots, t_n) \mathbb{1}(t_{n+1} \ge t \ge t_n) \mathrm{d}t_1 \dots \mathrm{d}t_{n+1} \\
&= \int_A \int_t^\infty \lambda^{n+1} e^{-\lambda t_{n+1}} \mathbb{1}(0 \le t_1 \le \dots \le t_n \le t) \mathrm{d}t_{n+1} \mathrm{d}t_1 \dots \mathrm{d}t_n \\
&= \int_A \lambda^n e^{-\lambda t} \mathbb{1}(0 \le t_1 \le \dots \le t_n \le t) \mathrm{d}t_1 \dots \mathrm{d}t_n.
\end{aligned}$$

Then we get

$$\mathbb{P}((J_1, \dots, J_n) \in A | X_t = n) = \int_A \frac{n!}{t^n} \mathbb{1}(0 \le t_1 \le \dots \le t_n \le t) \mathrm{d}t_1 \dots \mathrm{d}t_n.$$

As required. □

Now we look at a generalisation of a Poisson Process: called a Birth Process. For a Poisson Process, the rate of going from i to $i+1$ is λ. For a Birth Process, this is q_i (can depend on i). More precisely:

Definition (Birth Process)**.** For each i, let $S_i = \text{Exp}(q_i)$ with $S_1, S_2, \dots$ independent. Set $J_i = S_1 + \dots + S_i$ and $X_t = i$ if $J_i \le t < J_{i+1}$. Then X is called a *Birth Process*.

We have some special cases:

1. Simple birth process: when $q_i = \lambda i$ for $i = 1, 2, \ldots$;
2. Poisson Proces $q_i = \lambda$ for all i.

Motivation for Simple Birth Process (SBP): at time 0 there is only one 'individual' i.e $X_0 = 1$. Each individual has an exponential clock of parameter λ independently. Then if there are i individuals, the first clock rings after $\mathrm{Exp}(\lambda i)$ time, and we jump from i to $i+1$ individuals. Indeed, by the memoryless property, the process begins afresh after each jump.

Proposition 1.8. *Let $(T_k)_{k\geq 1}$ be a sequence of independent random variables with $T_K \sim \mathrm{Exp}(q_k)$ and $\sum_k q_k < \infty$. Let $T = \inf_k T_k$. Then*

(a) $T \sim \mathrm{Exp}\left(\sum_k q_k\right)$

(b) The infimum is attained at a point T_K almost surely, and

$$\mathbb{P}(K = n) = \frac{q_n}{\sum_k q_k}.$$

(c) T and K are independent.

Proof. See example sheet. $\square$

The main difference between a Poisson Process and a Birth Process is that there is the possibility of explosion in the Birth Process. Recall explosion occurs when $\zeta := \sum_n S_n < \infty$.

Proposition 1.9. *Let X be a Birth Process with rates q_i and $X_0 = 1$. Then*

1. *If $\sum_{i=1}^{\infty} \frac{1}{q_i} < \infty$, then X is explosive, i.e $\mathbb{P}(\zeta < \infty) = 1$;*
2. *If $\sum_{i=1}^{\infty} \frac{1}{q_i} = \infty$, then X is non-explosive, i.e $\mathbb{P}(\zeta = \infty) = 1$.*

Remark. This shows the SBP (as well as the PP) is non-explosive.

Proof.

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

$$\mathbb{E}[\zeta] = \mathbb{E}\left[\sum_n S_n\right] = \sum_n \mathbb{E}S_n = \sum_n \frac{1}{q_n} < \infty.$$

Where we have swapped summation and expectation by the MCT (monotone convergence theorem). Thus $\zeta = \sum_n S_n < \infty$ almost surely.

2. If $\sum_n \frac{1}{q_n} = \infty$, then $\prod_n \left(1 + \frac{1}{q_n}\right) \geq 1 + \sum_n \frac{1}{q_n} = \infty$. Then

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

Since $e^{-\zeta} \geq 0$, since $\mathbb{E}(e^{-\zeta}) = 0$ we have $e^{-\zeta} = 0$ almsot surely, i.e $\mathbb{P}(\zeta = \infty) = 1$.

$\square$

Theorem 1.10 (Markov Property)**.** *Let X be a BP with parameters (q_i). Conditional on $X_s = i$, the process $(X_{s+t})_{t\geq 0}$ is a birth process with rates $(q_j)_{j\geq i}$ starting from i, and independent of $(X_r)_{r\leq s}$.*

Proof. As in the Poisson Process case. □

Theorem 1.11. *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:*

1. *(jump chain/holding time definition) conditional on $X_s = i$, the holding times $S_1, S_2, \ldots$ are independent exponentials with rates $q_i, q_{i+1}, \ldots$ respectively and the jump chain is given $Y_n = i + n$ for all n.*
2. *(infinitesimal definition) for all $t, h \geq 0$, conditional on $X_t = i$, the process $(X_{t+h})_{h\geq 0}$ is independent of $(X_s)_{s\leq t}$ and as $h \to 0$, uniformly in t we have*
$$\mathbb{P}(X_{t+h} = i|X_t = i) = 1 - q_i h + o(h),$$
$$\mathbb{P}(X_{t+h} = i + 1|X_t = i) = q_i h + o(h).$$
3. *(transition probability definition) for all $n = 0, 1, 2, \ldots$ and all times $0 \leq t_0 \leq t_1 \leq \ldots \leq t_{n+1}$, and all states $i_0, i_1, \ldots, i_{n+1}$,*
$$\mathbb{P}(X_{t_{n+1}} = i_{n+1}|X_0 = i_0, \ldots, X_{t_n} = i_n) = p_{i_n,i_{n+1}}(t_{n+1} - t_n),$$
where $(p_{ij}(t) : i, j = 0, 1, 2, \ldots)$ is the unique solution to the equation (called Kolmogorov's forward equation)
$$p'_{ij}(t) = q_{j-1}p_{i,j-1}(t) - q_j p_{i,j}(t). \tag{$*$}$$
(as in the Poisson Process, $p_{ij}(t+h) = p_{i,j-1}(t)q_j h+p_{i,j}(t)(1-q_j h)+o(h)$.)

Existence and uniqueness of a solution in (3) gollow since for $i = j$ $p'_{i,i}(t) = -q_i p_{i,i}(t)$ and $p_{i,i}(0) = 1$, so $p_{i,i}(t) = e^{-q_i t}$. Then by induction, if the unique solution for $p_{i,j}(t)$ exists, then plug into $(*)$ to see there exists a unique solution for $p_{i,j+1}(t)$.

Also note that we can write the equation in matrix form:

$$P'(t) = P(t)Q, \text{ where } Q = \begin{pmatrix} -q_1 & q_1 & 0 & \ldots \\ 0 & -q_2 & q_2 & \ldots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Q-matrix and construction of Markov Processes

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

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

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

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

Write $q_i = -q_{ii} = \sum_{j \notin I} q_{ij}$ 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.

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

Definition. Let Q be a Q-matrix and λ a probability measure on the state space 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 starting from $Y_0 \sim \lambda$ with transition matrix P.

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

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

Example. Birth-Processes are $\mathrm{Markov}(\lambda, Q)$ with $I = \mathbb{N}$ and

$$Q = \begin{pmatrix} -q_1 & q_1 & 0 & \cdots \\ 0 & -q_2 & q_2 & \cdots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \text{ and } P = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

And jump chain $Y_n = Y_0 + n$.

We have multiple constructions of a Markov(λ, Q) process:
Construction 1:

- $(Y_n)_{n\geq 1}$ is a discrete-time Markov chain, $Y_0 \sim \lambda$ & transition matrix P.
- $(T_i)_{i\geq 1}$ iid Exp(1) random variables, independent of Y and set $S_n = \frac{T_n}{q_{X_{n-1}}}$ and $J_n = \sum_{i=1}^n S_i$ (this implies $S_n \sim \text{Exp}(q_{X_{n-1}})$) and set $X_t = Y_n$ if $J_n \leq t < J_{n+1}$ and $X_t = \infty$ otherwise.

Construction 2:

- Let $(T_n^y)_{\substack{n\geq 1 \\ y\in I}}$ be iid Exp(1) random variables
- $Y_0 \sim \lambda$ and inductively define Y_n, S_n: if $Y_n = x$ then for $y \neq x$ define $S_{n+1}^y = \frac{T_{n+1}^y}{q_{xy}} \sim \text{Exp}(q_{xy})$ and $S_{n+1} = \inf_{y\neq x} S_{n+1}^y \sim \text{Exp}\left(\sum_{y\neq x} q_{xy}\right)$, and if $S_{n+1} = S_{n+1}^Z$ for some random Z (since the infimum is attained), take $Y_{n+1} = Z$ (if $q_x > 0$). If $q_x = 0$ take $Y_{n+1} = x$.

(Proof of equivalence: see Example Sheet)

Construction 3:

- For $x \neq y$, let $(N_t^{x,y})$ be independent Posisson Processes with rates q_{xy} respectively. Let $Y_0 \sim \lambda$, $J_0 = 0$ and define inductively:

$$J_{n+1} = \inf\{t > J_n : N_t^{Y_n,y} \neq N_{J_n}^{Y_n,y} \text{ for some } y \neq Y_n\},$$

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

For a birth process, we characterised when explosion happens. In general, the next theorem gives a sufficient condition:

Theorem 1.12. *Let X be* Markov(λ, Q) *on I. Then $\mathbb{P}(\zeta = \infty) = 1$ (non-explosive) if any of the following hold:*

(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 that (a)$\Rightarrow$(b) so it is enough to show in the cases we have (b) or (c). If (b) holds, set $q = \sup_{x\in I} q_x < \infty$. Since $S_n = \frac{T_n}{q_{X_{n-1}}}$, $S_n \geq \frac{T_n}{q}$. Hence

$$\zeta = \sum_{n=1}^{\infty} S_n > \frac{1}{q}\sum_{n=1}^{\infty} T_n = \infty \text{ almost surely (SLLN)},$$

i.e $\mathbb{P}(\zeta = \infty) = 1$.

Now suppose (c) holds. Let $(N_i)_{i \in I}$ be the times when the jump chain Y visits x. By the SLLN,

$$\zeta \geq \sum_{i=1}^{\infty} S_{N_i+1} = \sum_{i=1}^{\infty} \frac{T_{N_i+1}}{q_{N_i}} = \frac{1}{q_x} \sum_{i=1}^{\infty} T_{N_i+1} = \infty \text{ almost surely,}$$

i.e $\mathbb{P}(\zeta = \infty) = 1$. □

Example. Suppose $I = \mathbb{Z}$, $q_{i,i+1} = q_{i,i-1} = 2^|i|$ for all i. Then $p_{i,i+1} = p_{i,i-1} = 1/2$ and the jump chain is the symmetric simple random walk on $\mathbb{Z}$, which is recurrent. Hence X is non-explosive.

Example. Suppose $I = \mathbb{Z}$, $q_{i,i+1} = 2^{|i|+1}$, $q_{i,i-1} = 2^{|i|}$ so $q_i = 2^{|i|} + 2^{|i|+1}$. Then the jump chain Y is a simple random walk with 1/3 probabilty of moving towards 0 and 2/3 probability of moving away from 0, hence is transient. We have

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

where V_j is the total number of visits to j and N_k^j is the time of the kth visit to j. Hence

$$\sum_{j \in \mathbb{Z}} \mathbb{E}\left[\sum_{k=1}^{V_j} S_{N_k^j+1}\right] = \sum_{j \in \mathbb{Z}} \mathbb{E}[V_j]\mathbb{E}[S_{N_1^j+1}] = \sum_{j \in \mathbb{Z}} \mathbb{E}[V_j]\frac{1}{q_j} = \sum_{j \in \mathbb{Z}} \frac{1}{3 \cdot 2^{|j|}} \mathbb{E}V_j.$$

Since $\mathbb{E}V_j \leq 1 + \mathbb{E}_j V_j = 1 + \mathbb{E}_0 V_0 := C < \infty$ (transience) we have

$$\sum_{j \in \mathbb{Z}} \frac{1}{3 \cdot 2^{|j|}} \mathbb{E}V_j \leq \sum_{j \in \mathbb{Z}} \frac{C}{2 \cdot 2^{|j|}} < \infty.$$

So $\mathbb{E}[\zeta] < \infty$ and $\mathbb{P}(\zeta < \infty) = 1$, i.e explosive.

Theorem 1.13 (Strong Markov Property)**.** *Let X be* Markov(λ, Q) *and let T be a stopping time. Then conditional on $T < \zeta$ and $X_T = x$, the process $(X_{T+t})_{t \geq 0}$ is* Markov(δ_x, Q) *and independent of $(X_s)_{s \leq T}$.*

Proof. Omitted (uses measure theory, see Norris (6.5)). □

Kolmogorov's forward & backward equations

We work on a countable state space I.

Theorem 1.14. *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 $n \geq 0$, $0 \leq t_0 \leq \ldots \leq t_{n+1}$, and all states $x_0, \ldots, x_{n+1} \in I$,*

$$\mathbb{P}(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_{t_n}, \ldots, X_{t_0} = x_1) = p_{x_n x_{n+1}}(t_{n+1} - t_n).$$

Where $(P(t)) = (p_{xy}(t))$ is the minimal non-negative solution to the backward equation

$$P'(t) = QP(t), \text{ with } P(0) = I.$$

(Minimality means that if $\tilde{P}$ is another non-negative solution, we have $p_{xy}(t) \leq \tilde{p}_{xy}(t)$ for all t and all $x, y \in I$.) In fact, if the chain is non-explosive, the solution is unique.

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

$$P'(t) = P(t)Q, \text{ with } P(0) = I.$$

Note. We shall skip the proof of the equivalence of (c) (see Norris (2.8)).

Proof. First we show (a)$\Rightarrow$(b). If $(J_n)_{n\geq 1}$ denote the jump times, then

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

Integrating over the values of $J_1 \leq t$ and using independence of the jump chain, for $z \neq x$,

$$\begin{aligned}\mathbb{P}_x(X_t = y, J_1 \leq t, X_{J_1} = z) &= \int_0^t q_x e^{-q_x s}\frac{q_{xz}}{q_x} p_{zy}(t-s)\mathrm{d}s \\ &= \int_0^t e^{-q_x s} q_{xz} p_{zy}(t-s)\mathrm{d}x\end{aligned}$$

Summing over all $z \neq x$ (and by the MCT),

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

So

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

And by a substitution

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

Hence $p_{xy}(t)$ 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 convergence (Weierstrass-M test), so continuous. Hence $e^{q_x t}p_{xy}(t)$ is differentiable with derivative

$$e^{q_x t}\left(q_x p_{xy}(t) + p'_{xy}(t)\right) = \sum_{z\neq x} e^{q_x t} q_{xz} p_{zy}(t).$$

Thus

$$p'_{xy}(t) = \sum_z q_{xz} p_{zy}(t) \implies P'(t) = QP(t).$$

Now we 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,

$$\begin{aligned}&\mathbb{P}_x(X_t = y, t < J_{n+1}) = \mathbb{P}_x(X_t = y, J_1 > t) + \mathbb{P}_x(X_t = y, J_1 \leq t < J_{n+1}) \\ &= e^{-q_x t}\mathbb{1}(x = y) + \sum_{z\neq x}\int_0^t q_x e^{-q_x s}\frac{q_{xz}}{q_x}\mathbb{P}_z\left(X_{t-s} = y, t-s < J_n\right)\mathrm{d}s.\end{aligned}$$

Now, as $\tilde{P}$ satisfies the backward equation, we get as before (retracing previous steps)

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

Now we prove by induction that

$$\mathbb{P}_x(X_t = y, t < J_n) \leq \tilde{p}_{xy}(t) \text{ for all } n.$$

For $n = 1$,

$$e^{-q_x t}\mathbb{1}(x=y) \leq \tilde{p}_{xy}(t) \text{ by } (*).$$

Assume true for some $n \in \mathbb{N}$. Then for $n+1$,

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

So it holds for all n. Hence

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

(Since $J_n \uparrow \zeta$.) Now by minimality,

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

□

Finite state space:

Definition. If A is a finite-dimensional square matrix, its matrix exponential is given by

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

Claim. *For any $r \times r$ matrix A, the exponential e^A is an $r \times r$ matrix. If A_1 and A_2 commute, then $e^{A_1+A_2} = e^{A_1}e^{A_2}$.*

Proof. Example Sheet. □

Proposition 1.15. *Let Q be a Q-matrix on a finite set I and $P(t) = e^{tQ}$. Then*

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

(ii) $(P(t))_{t\geq 0}$ *is the unique solution to the forward equation* $P'(t) = P(t)Q$, $P(0) = I$*;*

(iii) $(P(t))_{t\geq 0}$ *is the unique solution to the backward equation* $P'(t) = QP(t)$, $P(0) = I$*;*

(iv) For $k = 0, 1, 2, \dots$, $\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^k P(t)\Big|_{t=0} = Q^k$.

Proof.

(i) Since tQ and sQ commute, $\exp((t+s)Q) = \exp(tQ)\exp(sQ)$.

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

(iii) Same as (ii).

(iv) Same again.

Now we'll show uniqueness in (ii) and (iii). If $\tilde{P}$ is another solution to the forward equation, $\tilde{P}'(t) = \tilde{P}(t)Q$, $\tilde{P}(0) = I$, then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\tilde{P}(t)e^{-tQ}\right) &= \tilde{P}'(t)e^{-tQ} + \tilde{P}(t)\left(-Qe^{-tQ}\right) \\ &= \tilde{P}(t)Qe^{-tQ} - \tilde{P}(t)Qe^{-tQ} = 0\end{aligned}$$

So $\tilde{P}(t)e^{-tQ}$ is a constant matrix. Since $\tilde{P}(0) = I$, this implies $\tilde{P}(t) = e^{tQ}$. The same thing works for the backward equation. $\square$

Example. Let $Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}$. To find $p_{11}(t)$, we can diagonalise $Q = PDP^{-1}$ for a diagonal matrix

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

so

$$e^{tQ} = Pe^{tD}P^{-1} = P\begin{pmatrix} e^{t\lambda_1} & 0 & 0 \\ 0 & e^{t\lambda_2} & 0 \\ 0 & 0 & e^{t\lambda_3} \end{pmatrix}P^{-1}.$$

i.e $p_{11}(t) = ae^{t\lambda_1} + be^{t\lambda_2} + ce^{t\lambda_3}$, which we can solve by considering $p_{11}(0), p'_{11}(0), p''_{11}(0)$.

Theorem 1.16. *Let I be a finite state space and Q be a matrix. Then it is a Q-matrix iff $P(t) = e^{tQ}$ is a stochastic matrix for all t.*

Proof. For t sufficiently small, $p(t) = e^{tQ} = I + tQ + \mathcal{O}(t^2)$, so for all $x \neq y$, $q_{xy} \geq 0$ iff $p_{xy}(t) \geq 0$ for all t suffiently small.

Since $P(t) = (P(t/n))^n$ for all n, we get $q_{xy} \geq 0$ for all $x \neq y$ iff $p_{xy}(t) \geq 0$ for all $t \geq 0$.

Assume now that Q is a Q-matrix, i.e $\sum_y q_{xy} = 0$ 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 $Q^n \mathbf{1} = Q^{n-1} Q \mathbf{1} = 0$ ($\mathbf{1}$ is vector will all entries 1). Hence, since

$$p_{xy}(t) = \delta_{xy} + \sum_{k=1}^{\infty} \frac{t^k}{k!} (Q^k)_{xy}$$

we have $\sum_y p_{xy}(t) = 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_y (Q^k)_{xy} = 1$. i.e $P(t)$ is a stochastic matrix.

Assume now that $P(t)$ is a stochastic matrix. Then as $Q = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} P(t)$, we have

$$\sum_y q_{xy} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \sum_y p_{xy}(t) = 0.$$

i.e Q is a Q-matrix. ☐

Theorem 1.17. *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 X is Markov with generator Q (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 $h \downarrow 0$, for all x, y

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

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

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

where $(p_{xy}(t))$ is the solution to the forward equation $P'(t) = P(t)Q$, $P(0) = I$.

Proof. We have already shown (a) $\iff$ (b) (from countable setting), so it is enough to show (b) $\iff$ (c).

First we show (c)$\Rightarrow$(b). $P(t) = e^{tQ}$ is the solution (as I is finite). As $t \downarrow 0$, $P(t) = I + tQ + \mathcal{O}(t^2)$. Thus for all $t > 0$ and as $h \downarrow 0$, $\forall x, y$,

$$\mathbb{P}(X_{t+h} = y | X_t = x) = \mathbb{P}(X_h = y | X_0 = x) = p_{xy}(h) = \delta_{xy} + h q_{xy} + o(h).$$

Now we show (b)⇒(c). We have

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

So

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

As $h \downarrow 0$,

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

□

Remark. To get the backward equation we could write

$$p_{xy}(t+h) = \sum_z p_{xz}(h) p_{zy}(t)$$

and continue similarly.

Qualitative Properties of Continuous Time Markov Chains

We have minimal chains, and countable state space.

Class Structure

Definition. For states $x, y \in I$, write $x \to y$ ("*x leads to y*") if $\mathbb{P}_x(X_t = y \text{ for some } t \geq 0) > 0$. We write $x \leftrightarrow y$ ("*x communicates with y*") if $x \to y$ and $y \to x$. Clearly this is an equivalence relation and we call the equivalence classes *communicating classes*. We define *irreducibility, closed class* and *absorbing states* exactly as in discrete Markov Chains.

Proposition 1.18. *Let X be* Markov(Q) *with transition semigroup $(P(t))_{t\geq 0}$. For any 2 states $x, y \in I$, the following are equivalent*

(a) $x \to y$;

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

(c) $q_{x_0x_1} \dots q_{x_{n-1}x_n} > 0$ for some $x = x_0, x_1, \dots, x_{n-1}, x_n = y$;

(d) $p_{xy}(t) > 0$ for all $t > 0$;

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

Proof. Clearly (d)⇒(e)⇒(a)⇒(b). Now we show (b)⇒(c). Since $x \to y$ for the jump chain, there exist $x_0 = x, x_1, \dots, x_{n-1}, x_n = y \in I$ such that

$$p_{x_0x_1} p_{x_1x_2} \dots p_{x_{n-1}x_n} > 0.$$

Hence $q_{x_0x_1}q_{x_1x_2}\dots q_{x_{n-1}x_n}$ since $q_{xy}/q_x = p_{xy}$.

Now we show (c)⇒(d). For any 2 states w, z with $q_{wz} > 0$, and for any $t > 0$,

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

i.e $q_{wz} > 0$ implies $q_{wz}(t) > 0$ for all t. Hence if (c) holds, $p_{x_i x_{i+1}}(t) > 0$ for all t and all $0 \leq i \leq n-1$. Then $p_{xy}(t) = p_{x_0x_1}(t/n)p_{x_1x_2}(t/n)\dots p_{x_{n-1}x_n}(t/n) > 0$. □

Hitting times

Definition. Let Y be the jump chain associated with X, and $A \subseteq I$. Set $T_A = \inf\{t > 0 : X_t \in A\}$, $H_A = \inf\{n \geq 0 : Y_n \in A\}$, $h_A(x) = \mathbb{P}_x(T_A < \infty)$ (hitting probability), $k_A(x) = \mathbb{E}_x T_A$ (mean hitting time).

Note. The hitting probability for X is the same as that for Y but the mean hitting times will differ in general.

Theorem 1.19. *$(h_A(x))_{x\in I}$ and $(k_A(x))_{x\in I}$ are the minimal non-negative solutions to*

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

and

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

respectively (assume $q_x > 0$ for all $x \notin A$).

Proof. The hitting probabilities are the same as those for the jump chain. Hence $h_A(x) = 1$ for all $x \in A$ and $h_A(x) = \sum_{y\neq x} p_{xy} h_A(y)$ for all $x \notin A$. Hence for all $x \notin A$

$$q_x h_A(x) = \sum_{y\neq x} h_A(y) q_{xy} \implies \sum_y h_A(y) q_{xy} = 0.$$

Clearly if $x \in A$, $T_A = 0$, so $k_A(x) = 0$. Let $x \notin A$. Then $J_1 \leq T_A$, and hence

$$\begin{aligned} k_A(x) &= \mathbb{E}_x T_A \\ &= \mathbb{E}_x J_1 + \mathbb{E}_x(T_A - J_1) \\ &= \mathbb{E}_x J_1 + \sum_{y\neq x} \mathbb{E}_x(T_A - J_1 | Y_1 = y) p_{xy} \\ &= \frac{1}{q_x} + \sum_{y\neq x} k_A(y) \frac{q_{xy}}{q_x}. \end{aligned}$$

Therefore

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

The minimality of solutions is as in the discrete chain. □

Recurrence and Transience

Definition. The state x is called *recurrent* for X if

$$\mathbb{P}(\{t : X_t = x\} \text{ is unbounded}) = 1.$$

The state x is called *transient* if

$$\mathbb{P}(\{t : X_t = x\} \text{ is unbounded}) = 0.$$

Remark. If X explodes with positive probability starting from x, i.e $\mathbb{P}(\zeta < \infty) > 0$, then $\sup_t\{t : X_t = x\} \leq \zeta < \infty$ with positive probability so x cannot be recurrent.

Theorem 1.20. *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. (a) & (b) will imply (c) & (d) through the results for the discrete chain. So we prove (a) and (b).

First we prove (a). Suppose x is recurrent for Y and $X_0 = x$. Then X is not explosive, i.e $\mathbb{P}(\zeta = \infty) = 1$, so $J_n \to \infty$ with probability 1 (starting from x). Since $X_{J_n} = Y_n$ for all n, and Y visits x infinitely often with probability 1, $\{t : X_t = x\}$ is unbounded with probability 1.

Now we prove (b). If x is transient for Y, $q_x > 0$ (otherwise x is an absorbing state). Also, almost surely there is a last visit to x for Y, i.e

$$N := \sup\{n : Y_n = x\} < \infty \text{ almost surely.}$$

Also, $J_{N+1} < \infty$ almost surely (as $q_x > 0$) and if $t \in \{s : X_s = x\}$, then $t \leq J_{N+1}$, i.e $\sup\{s : X_s = x\} \leq J_{n+1} < \infty$ almost surely. □

Like in the discrete-time chain, $\sum_{n\geq 1} p_{xx}(n) = \infty$ implies x is recurrent; and $\sum_{n\geq 1} p_{xx}(n) < \infty$ implies x is transient.

Theorem 1.21. *x is recurrent for X if and only if $\int_0^\infty p_{xx}(t)\mathrm{d}t = \infty$, and x is transient for X if and only if $\int_0^\infty p_{xx}(t)\mathrm{d}t < \infty$.*

Proof. If $q_{xx} = 0$, then x is absorbing, i.e $p_{xx}(t) = 1$ for all t and $\int_0^\infty p_{xx}(t)\mathrm{d}t = \infty$. Assume $q_x > 0$. Then

$$\begin{aligned}
\int_0^\infty p_{xx}(t)\mathrm{d}t &= \int_0^\infty \mathbb{E}[\mathbb{1}(X_t = x)]\mathrm{d}t \\
&= \mathbb{E}_x\left[\int_0^\infty \mathbb{1}(X_t = x)\mathrm{d}t\right] && \text{(Fubini)} \\
&= \mathbb{E}_x\left[\sum_{n=0}^\infty \mathbb{1}(Y_n = x)S_{n+1}\right] \\
&= \sum_{n=0}^\infty \mathbb{E}_x\left[\mathbb{1}(Y_n = x)S_{n+1}\right] && \text{(Fubini)} \\
&= \sum_{n=0}^\infty \mathbb{P}_x(Y_n = x)\mathbb{E}_x\left[S_{n+1} | Y_n = x\right] \\
&= \sum_{n=0}^\infty p_{xx}(n)\frac{1}{q_x}.
\end{aligned}$$

□

Invariant Distributions

Definition. For a discrete Markov Chain Y, π is an *invariant measure* for Y if $\pi P = \pi$. If in addition $\sum \pi_i = 1$, π is called a *invariant distribution.* Then if $Y_0 \sim \pi$, $Y_n \sim \pi$ for all $n \geq 1$.

Recall:

Theorem 1.22. *If Y is a discrete time Markov Chain which is irreducible, recurrent and $x \in I$. Then*

$$\nu^x(y) = \mathbb{E}_x\left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y)\right] \text{ where } H_x = \inf\{n \geq 1 : Y_n = x\}.$$

Then $\nu^x(\cdot)$ is an invariant measure and $0 < \nu^x(y) \leq 1$ for all y, $\nu^x(x) = 1$.

Theorem 1.23. *If Y is irreducible, λ is any invariant measure with $\lambda(x) = 1$, then*

$$\lambda(y) \geq \nu^x(y) \text{ for all } y.$$

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

Definition. Let $X \sim \text{Markov}(Q)$ and let λ be a measure. Then λ is called invariant/infinitesimally invariant if $\lambda Q = 0$.

Lemma 1.24. *If $|I|$ is finite, then $\lambda Q = 0$ if and only if $\lambda P(s) = \lambda$ for all $s \geq 0$.*

Proof. $P(s) = e^{sQ}$ since I is finite. If $\lambda Q = 0$, then

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

If $\lambda P(s) = \lambda$ for all s, then

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

$\square$

Lemma 1.25. *Let X be* Markov(Q) *and Y its jump chain. π is invariant for X if and only if μ defined by $\mu_x = q_x \pi_x$ is invariant for Y (i.e $\pi Q = 0$ if and only if $\mu P = \mu$).*

Proof. Since $q_x\left(p_{xy} - \delta_{xy}\right) = q_{xy}$,

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

$\square$

Theorem 1.26. *Let X be irreducible & 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, $\nu^x(y) = \mathbb{E}_x\left[\sum_{n=0}^{H_x - 1} \mathbb{1}(Y_n = y)\right]$ where $H_x = \inf\{n \geq 1 : Y_n = x\}$ is an invariant measure as Y is irreducible & recurrent (since X is), hence ν^x is an invariant measure for Y which is unique up to scalar multiplication. By the previous lemma, $\frac{\nu^x(y)}{q_y}$ is an invariant measure for X, and also unique up to scalar multiplication. $\square$

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

Lemma 1.27. *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 $\mu^x(y) = \frac{\nu^x(y)}{q_y}$.

Proof.

$$\begin{aligned}
\mu^x(y) &= \mathbb{E}_x\left[\int_0^{T_x} \mathbb{1}(X_t = y)\mathrm{d}t\right] \\
&= \mathbb{E}_x\left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y)S_{n+1}\right] \\
&= \mathbb{E}_x\left[\sum_{n=0}^{\infty} S_{n+1}\mathbb{1}(Y_n = y, n \le H_x - 1)\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_x\left[S_{n+1} | Y_n = y, n \le H_x - 1\right] \mathbb{P}_x(Y_n = y, n \le H_x - 1)
\end{aligned}$$

Since $\{n < H_x\}^c = \{H_x \le n\} \in \sigma\{Y_1, \dots, Y_n\}$ (i.e depends on $Y_1, \dots, Y_n$ only) its a stopping time so the Strong Markov Property says

$$\begin{aligned}
\mu^x(y) &= \sum_{n=0}^{\infty} \mathbb{E}_x\left[S_{n+1} | Y_n = y\right] \mathbb{P}_x(Y_n = y, n \le H_x - 1) \\
&= \sum_{n=0}^{\infty} \frac{1}{q_y}\mathbb{P}_x(Y_n = y, n \le H_x - 1) \\
&= \frac{1}{q_y}\sum_{n=0}^{\infty} \mathbb{E}_x\left[\mathbb{1}(Y_n = y, n \le H_x - 1)\right] \\
&= \frac{1}{q_y}\mathbb{E}\left[\sum_{n=0}^{\infty} \mathbb{1}(Y_n = y, n \le H_x - 1)\right] \\
&= \frac{1}{q_y}\mathbb{E}_x\left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y)\right] \\
&= \frac{\nu^x(y)}{q_y}.
\end{aligned}$$

□

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

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

Otherwise, we call x *null recurrent*.

Theorem 1.28. *Let $X \sim$ 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.

Also, when (c) holds, the invariant distribution λ is given by $\lambda(x) = \frac{1}{q_x m_x}$ for all x.

Proof. Clearly (a)⇒(b). Now we show (b)⇒(c). Assume without loss of generality that $q_x > 0$. Let x be a positive recurrent state. Then all states are recurrent, so Y is recurrent and the chain is non-explosive starting from any y. As Y is recurrent, ν^x is an invariant measure for Y. So $\mu^x = \frac{\nu^x}{q_y}$ (as defined previously) is an invariant measure for X. Also

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

so

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

So μ_x is normalisable, and $\frac{\mu_x}{\mathbb{E}_x T_x}$ is an invariant distribution for X.

Now we show (c)⇒(a). By a previous lemma, the measure $\beta(y) = \lambda(y)q_y$ is an invariant measure for Y. Since $\sum_{y\in I}\lambda(y) = 1$, $\lambda(x) > 0$ for some x. Since Y is irreducible, for any $y \in I$, $x \to y$, i.e $p_{xy}(n) > 0$ for some n. As β is invariant for Y, $\beta P^n = \beta$. So

$$\lambda(y)q_y = \beta y = \sum_{z\in I}\beta_z p_{zy}(n) \geq \beta_x p_{xy}(n) = \lambda(x)q_x p_{xy}(n) > 0$$

so $\lambda(y) > 0$ for all y. □