UNIVERSITY OF COPENHAGEN
DEPARTMENT OF
MATHEMATICAL SCIENCES

Stochastic processes
Søren Feodor Nielsen
May 11, 2009

Continuous-time homogeneous Markov chains

Contents

1 Poisson processes **2**

1.1 The Poisson process 2
1.2 Some properties of the exponential distribution 4
1.3 Some properties of the Poisson process 5

2 Continuous-time homogeneous Markov chains **6**

2.1 Regular jump continuous-time Markov chains 6
2.2 Infinitesimal generator 10
2.3 Uniform Markov chains 11
2.4 Differential equations 15
2.4.1 Transition probabilities 15
2.4.2 Finite state space: Matrix exponentials 17
2.4.3 Distribution 18
2.5 Explosion 19
2.6 State space decomposition 21
2.6.1 Communication, transience and recurrence 21
2.6.2 Positive and null recurrence 23
2.6.3 Stationarity and reversibility 26
2.7 Long-run behaviour 26
2.7.1 Ergodicity 26
2.7.2 Absorption 27
2.8 Birth-and-death processes 28
2.8.1 Explosion 28

2.8.2 Linear birth-and-death processes 30
2.8.3 Birth-and-death processes with immigration 31
2.8.4 Pure birth and death processes 31

1 Poisson processes

1.1 The Poisson process

In order to define the Poisson process, we need a definition of a continuous-time stochastic process.

Definition 1 (Continuous-time stochastic process) *A **continuous-time stochastic process**, $(X(t))_{t\geq 0}$, with state space E is a collection of random variables $X(t)$ with values in E.*

With an at most countable state space, E, the distribution of the stochastic process, $(X(t))_{t\geq 0}$, is determined by the probabilities

$$P\{X(t_0) = i_0, X(t_1) = i_1, \ldots, X(t_n) = i_n\} \tag{1}$$

for $0 \leq t_0 < t_1 < \ldots < t_n, i_0, i_1, \ldots, i_n \in E$ for all $n \in \mathbb{N}$.

One of the fundamental continuous-time processes, and quite possibly the simplest one, is the Poisson process, which may be defined as follows:

Definition 2 (Homogeneous Poisson process) *Let $S_1, S_2, \ldots$ be a sequence of independent identically exponentially distributed random variables with intensity λ. Put $T_n = \sum_{k=1}^{n} S_k$. Then*

$$N(t) = \sum_{n=1}^{\infty} 1_{\{T_n \leq t\}} \qquad t \geq 0$$

*is a **homogeneous Poisson process** with intensity λ.*

Thus the Poisson process describes the number of "events" happened until time t when the waiting time from one event to the next is exponential and independent of all other waiting times.

Following Brémaud, we will use HPP as an abbreviation of homogeneous Poisson process. We will not discuss inhomogeneous Poisson processes in these notes and will therefore just say "Poisson process" when we mean "homogeneous Poisson process".

The distribution of a Poisson process may in principle be derived from the definition. For instance we have the following partial result:

Theorem 3 *If $(N(t))_{t\geq 0}$ is a Poisson process with intensity $\lambda > 0$ then for each $t > 0$ $N(t) \sim$ **Pois**(λt).*

Proof For each $n \in \mathbb{N}$ we have $T_n \sim \Gamma(n, 1/\lambda)$ and hence

$$P\{N(t) \geq n\} = P\{T_n \leq t\} = 1 - \int_t^\infty \frac{\lambda^n}{\Gamma(n)} u^{n-1} e^{-\lambda u} du = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

where the last equality follows by partial integration. □

The following equivalent definition of a Poisson process is often useful.

Theorem 4 (Equivalent definition of a HPP) *$(N(t))_{t\geq 0}$ is a Poisson process with intensity λ if and only if*

1. $P\{N(0) = 0\} = 1$
2. $\forall n \in \mathbb{N}, 0 < t_0 < t_1 < \ldots < t_n$: *The increments* $N(t_0), N(t_1) - N(t_0), \ldots, N(t_n) - N(t_{n-1})$ *are independent.*
3. $\forall 0 < s < t$: $N(t) - N(s) \sim$ **Pois**$(\lambda(t - s))$

Proof[1] By direct calculation, one sees that a process fulfilling 1–3 has independent identically exponentially distributed waiting times between jumps; see the proof of Brémaud Theorem 8.1.1. Thus a process fulfilling 1–3 is a Poisson process. On the other hand, we have now shown that for any given Poisson process, $(N_t)_{t\geq 0}$, with intensity λ there is a process, $(\tilde{N}_t)_{t\geq 0}$, fulfilling 1–3, which is also a Poisson process with intensity λ. Thus,

$$P\{\tilde{N}(t_0) = i_0, \ldots, \tilde{N}(t_n) = i_n\} = P\{N(t_0) = i_0, \ldots, N(t_n) = i_n\}$$

From this it follows that a Poisson process fulfils 1–3. □

The theorem gives an alternative definition of a Poisson process as a process with independent stationary Poisson distributed increments. Processes with independent stationary increments are known as Levy processes. Another example of a Levy process is the very important Brownian motion, which has independent stationary Gaussian increments.

Note that the distribution of a Poisson process is easier to write down using Theorem 4 than by using Definition 2.

[1]**This remark may be skipped.** In our definition of the Poisson process we have imposed a certain structure on the sample paths of the process: It starts in 0, stays there for an exponential time, then jumps to 1, stays there for another (independent) exponential time, and so on. Due to measure-theoretic subtleties, one may construct a process fulfilling 1–3 of Theorem 4 which does not have this structure. On the other hand one may always choose a version of a process fulfilling 1–3 of Theorem 4 which has the desired structure and it is this version that we prove is a Poisson process in our sense.

1.2 Some properties of the exponential distribution

The exponential distribution is of course essential to the understanding of the Poisson process but also for the Markov chains to be discussed next. Here we collect a few useful results.

Theorem 5 (Memoryless property) *If $X \sim$ **Exp**$(1/\lambda)$ then $X - t | X > t \sim$ **Exp**$(1/\lambda)$. Conversely, if X is a non-negative random variable with a continuous distribution such that the conditional distribution of $X - t$ given $X > t$ is the same as the distribution of X for all $t > 0$, then X is exponentially distributed.*

Proof If $X \sim$ ***Exp***$(1/\lambda)$ then

$$P\{X - t > s | X > t\} = \frac{P\{X > t + s\}}{P\{X > t\}} = e^{-\lambda s} \qquad t, s > 0.$$

Conversely if for all $s, t > 0$ we have

$$P\{X > s\} = P\{X - t > s | X > t\} = \frac{P\{X > t + s\}}{P\{X > t\}}$$

then

$$g(t) = \log P\{X > t\}$$

is a decreasing linear function, $g(s) = -\lambda s$ for some $\lambda > 0$, say, and it follows that $P\{X > s\} = \exp(-s\lambda)$ for some $\lambda > 0$. □

Lemma 6 *Let $X_1, \ldots, X_n$ be independent exponentially distributed random variables with intensity $\lambda_1, \ldots, \lambda_n$. Then $\min_{i=1,\ldots,n} X_i$ is exponentially distributed with intensity $\lambda_1 + \cdots + \lambda_n$.*

Proof For any $x > 0$ we have

$$P\{\min_{i=1,\ldots,n} X_i > x\} = \prod_{i=1}^n P\{X_i > x\} = \exp\left(-x \sum_{i=1}^n \lambda_i\right)$$

proving the claim. □

Theorem 7 (Competition theorem) *Let X_i $i \in I$ be independent random variables such that $X_i \sim$ **Exp**$(1/\lambda_i)$, where I is an at most countable set. Suppose that $\lambda = \sum_{i \in I} \lambda_i < \infty$ and put*

$$Z = \inf_{i \in I} X_i \quad \textit{and} \quad K = i \textit{ if } X_i = Z$$

*Then Z and K are independent, $Z \sim$ **Exp**$(1/\lambda)$ and*

$$P\{K = i\} = \frac{\lambda_i}{\lambda} \quad i \in I$$

Proof First suppose that I is finite and pick an arbitrary $i \in I$. Put $U = \inf_{j \in I \setminus \{i\}} X_j$. Then $U \sim \textbf{\textit{Exp}}(1/\sum_{j \in I \setminus \{i\}} \lambda_j)$ by Lemma 6. Then with f denoting the density of U and using the fact that X_i and U are independent we get

$$\begin{aligned} P\{K = i, Z > z\} = P\{z < X_i < U\} &= \int_z^\infty \int_x^\infty f(u)du \cdot \lambda_i e^{-\lambda_i x} dx \\ &= \int_z^\infty e^{-x \sum_{j \in I \setminus \{i\}} \lambda_j} \cdot \lambda_i e^{-\lambda_i x} dx = \frac{\lambda_i}{\lambda} \int_z^\infty \lambda e^{-x\lambda} dx \\ &= \frac{\lambda_i}{\lambda} \cdot e^{-z\lambda} \end{aligned}$$

proving the result when I is finite. If I is infinite but countable we may choose finite subsets $I_n \uparrow I$ and letting $Z_n = \inf_{i \in I_n} X_i$ and $K_n = i$ if $X_i = Z_n$ we see that

$$P\{K = i, Z > z\} = \lim_{n \to \infty} P\{K_n = i, Z_n > z\} = \lim_{n \to \infty} \frac{\lambda_i}{\sum_{i \in I_n} \lambda_i} \cdot e^{-z \sum_{i \in I_n} \lambda_i} = \frac{\lambda_i}{\lambda} e^{-z\lambda}$$

as $\{K_n = i, Z_n > z\} \downarrow \{K = i, Z > z\}$. □

Remark One may worry whether K is well-defined: Could we (in the infinite I case!) have $Z = X_i$ and $Z = X_j$ for some $i \neq j$? If we in the proof let $K = \Delta \notin I$ if more than one X_i equals Z, then the proof may be repeated unchanged and we get

$$P\{K = i\} = \frac{\lambda_i}{\lambda} \quad i \in I$$

From this we see that $P\{K = \Delta\} = 0$ setting our minds at rest.

1.3 Some properties of the Poisson process

Theorem 8 (Sum of independent Poisson processes) *Let $(N_i)_{i \in I}$ be an at most countable family of independent Poisson processes with intensities $(\lambda_i)_{i \in I}$. If $\lambda = \sum_{i \in I} \lambda_i < \infty$ then*

$$N(t) = \sum_{i \in I} N_i(t) \qquad t \geq 0$$

is a Poisson process with parameter λ.

Proof See Brémaud, proof of Theorem 8.1.2. □

At first, summing independent Poisson processes may seem to be a strange idea. However, suppose that we observe different types of events happen over time and that each type of event occur according to its own Poisson process independently of the other types; the different types could be different types of claims in an insurance company or deaths due to different causes. Then ignoring the type of events, i.e. summing the event processes, the resulting event process is still Poisson.

The next theorem proves a kind of converse: If events arrive according to a Poisson process and these events are of different types, then if the type of event is independent of the arrival process, each type specific arrival process is Poisson.

Theorem 9 (Thinning) *Let $(N(t))_{t\geq 0}$ be a Poisson process with parameter λ and let $(X_n)_{n\in\mathbb{N}}$ be a sequence of iid random variables with values in a at most countable set I and distribution given by*

$$P\{X_n = i\} = p_i \qquad i \in I$$

Suppose that $(X_n)_{n\in\mathbb{N}}$ is independent of $(N(t))_{t\geq 0}$ and put

$$N_i(t) = \sum_{n=1}^{\infty} 1_{\{X_n=i\}} \cdot 1_{\{T_n\leq t\}} \qquad t \geq 0, i \in I$$

where $T_1, T_2, \ldots$ are the arrival times for the Poisson process $(N(t))_{t\geq 0}$. Then the processes $(N_i(t))_{t\geq 0}$, $i \in I$ are independent Poisson processes with parameters $p_i\lambda$, $i \in I$.

Proof For simplicity, consider the case where $I = \{1, 2\}$. Here

$$\begin{aligned}
&P\{N_1(t_0) = n_{1,0}, N_1(t_1) - N_1(t_0) = n_{1,1}, N_2(t_0) = n_{2,0}, N_2(t_1) - N_2(t_0) = n_{2,1}\} \\
&\quad = P\{\text{the same}|N(t_0) = n_0, N(t_1) - N(t_0) = n_1\}P\{N(t_0) = n_0, N(t_1) - N(t_0) = n_1\} \\
&\quad = \binom{n_0}{n_{1,0}} p_1^{n_{1,0}} p_2^{n_0-n_{1,0}} \cdot \binom{n_1}{n_{1,1}} p_1^{n_{1,1}} p_2^{n_1-n_{1,1}} \cdot \frac{(\lambda t_0)^{n_0}}{n_0!} e^{-\lambda t_0} \cdot \frac{(\lambda(t_1-t_0))^{n_1}}{n_1!} e^{-\lambda(t_1-t_0)} \\
&\quad = \frac{(p_1\lambda t_0)^{n_{1,0}}}{n_{1,0}!} e^{-p_1\lambda t_0} \cdot \frac{(p_1\lambda(t_1-t_0))^{n_{1,1}}}{n_{1,1}!} e^{-p_1\lambda(t_1-t_0)} \\
&\qquad \cdot \frac{(p_2\lambda t_0)^{n_{2,0}}}{n_{2,0}!} e^{-p_2\lambda t_0} \cdot \frac{(p_2\lambda(t_1-t_0))^{n_{2,1}}}{n_{2,1}!} e^{-p_2\lambda(t_1-t_0)}
\end{aligned}$$

from which it follows that the increments, $N_1(t_0), N_1(t_1) - N_1(t_0), N_2(t_0), N_2(t_1) - N_2(t_0)$, are independent and Poisson distributed. Moreover, we also obtain for instance independence of $N_1(t_1)$ and $N_2(t_0)$ by summing over the possible values of $N_1(t_0)$ and $N_2(t_1) - N_2(t_0)$. Clearly, all of this may be extended to any finite number of increments and to any finite I proving the result for I finite. To obtain the result for an infinite I observe that for any finite subset I_n, N_i, $i \in I_n$ are independent Poisson processes, and this is exactly what is meant by the claim that N_i, $i \in I$ are independent Poisson processes. □

2 Continuous-time homogeneous Markov chains

2.1 Regular jump continuous-time Markov chains

Definition 10 (Continuous-time Markov chains) *A continuous-time stochastic process $(X(t))_{t\geq 0}$ is a homogeneous Markov chain if for all $0 < t_1 < \ldots < t_n, i_0, i_1, \ldots, i_n \in E, n \in \mathbb{N}$*

$$\begin{aligned}
&P\{X(0) = i_0, X(t_1) = i_1, \ldots, X(t_n) = i_n\} \\
&\qquad = \nu(i_0)p_{i_0,i_1}(t_1)p_{i_1,i_2}(t_2 - t_1) \cdots p_{i_{n-1},i_n}(t_n - t_{n-1})
\end{aligned} \tag{2}$$

We see directly from (2) that the distribution of a continuous-time HMC is determined by the initial distribution

$$\nu(i) = P\{X(0) = i\} \qquad i \in E$$

and the transition probabilities

$$P(t) = [p_{i,j}(t)]_{i,j\in E} \qquad t \geq 0$$

where $P(0) = I$, the identity matrix. We note that

$$p_{i,j}(t) = P\{X(t) = j|X(0) = i\} = P\{X(t+s) = j|X(s) = i\} \qquad \text{for all } s > 0$$

The "homogeneity" in the definition of homogeneous Markov chains is due to the fact that the transition probabilities only depends on the difference t between s and $s+t$ and not on the actual times $(s, s+t)$.

Moreover, we directly see that (2) is equivalent to

$$\begin{aligned} &P\{X(t_{n+1}) = j|X(0) = i_0, X(t_1) = i_1, \dots, X(t_n) = i\} \\ &\qquad = P\{X(t_{n+1}) = j|X(t_n) = i\} = p_{i,j}(t_{n+1} - t_n) \end{aligned} \tag{3}$$

so that continuous-time Markov chains have a Markov property similar to the discrete-time Markov chains. Note that if $X_0 \sim \mu = (\mu_i)_{i\in E}$ then

$$\mu_j(t) = P\{X(t) = j\} = \sum_{i\in E} P\{X(t) = j|X(0) = i\}P\{X(0) = i\} = \sum_{i\in E} \mu(i)p_{ij}(t) \quad j \in E$$

i.e. $X \sim \mu(t) = (\mu^\top P(t))^\top$.

Theorem 11 (Chapman-Kolmogorov equations) *For any $t, s > 0$ we have*

$$P(t+s) = P(t)P(s)$$

Proof For any $i, j \in E$ we have

$$\begin{aligned} p_{i,j}(t+s) = P\{X(t+s) = j|X(0) = i\} &= \sum_{k\in E} P\{X(t+s) = j, X(t) = k|X(0) = i\} \\ &= \sum_{k\in E} P\{X(t+s) = j|X(t) = k, X(0) = i\}P\{X(t) = k|X(0) = i\} \\ &= \sum_{k\in E} p_{i,k}(t)p_{k,j}(s) \end{aligned}$$

□

The collection of transition matrices $(P(t))_{t\geq 0}$ is called the *transition semigroup* of the Markov chain.

Definition 10 does not ensure that the Markov chain is in any way "well-behaved". For instance, without additional assumptions the process may change state all the time so that $X(t) \neq X(s)$ for all s and t regardless of how small $t - s$ is. So we will impose the condition on all Markov chains to be considered from now on that they are so-called "regular jump processes":

Definition 12 (Jump processes/Regular jump processes) *A stochastic process $(X(t))_{t\geq 0}$ on an at most countable state space is a* ***jump process*** *if whenever $X(t)$ jumps from one state to another it will remain in the new state at least for a short (random) while. It is a* ***regular jump process*** *if it is a jump process and it only has finitely many jumps in $[0;t]$ for any $t > 0$.*

The critical reader will argue that this definition is not sufficiently precise (and such a reader is referred to Brémaud's Definition 8.2.5). However, for the present course it is sufficient. A HMC which is not a regular jump process will (at least with positive probability) jump from one state to another and then directly following this transition jump again. Processes behaving in this manner are rarely useful in applications so our restriction is not one that should cause us any concern.

Theorem 13 (HPP is a regular jump HMC) *The Poisson process with intensity λ is a regular jump homogeneous Markov chain with transition probabilities given by*

$$p_{i,j}(t) = \begin{cases} \dfrac{(\lambda t)^{j-i}}{(j-i)!}e^{-\lambda t} & \textit{for } j \geq i \\ 0 & \textit{for } j < i \end{cases} \qquad t > 0$$

Proof Note that by construction the Poisson process given by Definition 2 is a jump process. It is regular as

$$P\{N(t) = \infty\} = 0 \quad \text{for all } t > 0.$$

That it is also a Markov chain with the transition probabilities given above follows easily from the alternative definition (Theorem 4; see Brémaud, Example 8.2.1 for details). □

For a Markov chain $(X(t))_{t\geq 0}$ with state space E the sequence of *transition times*, $(\tau_n)_{n\in\mathbb{N}}$, are the times when $X(t)$ jumps i.e.

$$\tau_n = \inf\{t \geq \tau_{n-1} : X(t) \neq X(\tau_{n-1})\} \quad (\text{with } \tau_0 = 0 \text{ and } \inf\emptyset = \infty)$$

Observe that $\tau_n = \infty$ implies that $\tau_{n+k} = \infty$ for all $k \in \mathbb{N}$. The times between transition times $\tau_{n+1} - \tau_n$ are called *holding times* or *inter-arrival times*. Note that for a regular jump process $\tau_n \to \infty$ as $n \to \infty$. The *embedded process* is given by $X_0 = X(0)$ and

$$X_n = \begin{cases} X(\tau_n) & \text{if } \tau_n < \infty \\ \Delta & \text{if } \tau_n = \infty \end{cases} \qquad n \in \mathbb{N}$$

where Δ is an arbitrary element not in E. Clearly the distribution of the Markov chain is given by the joint distribution of the transition times (or equivalently the sequence of holding times) and the embedded process. The following important result gives such a characterisation:

Theorem 14 (Regenerative structure) *Let $(X(t))_{t\geq 0}$ be a continuous-time homogeneous Markov chain with state space E. Then there exists $Q = [q_{ij}]_{i,j\in E}$, $(q_i)_{i\in E}$ such that*

1. *Given the embedded process $(X_n)_{n\in\mathbb{N}_0}$ the sequence of holding times, $(\tau_{n+1} - \tau_n)_{n\in\mathbb{N}_0}$, are independent with distribution given by*

$$P\{\tau_{n+1} - \tau_n > x | (X_k)_{k\in\mathbb{N}_0}\} = \begin{cases} \exp(-q_{X_n} x) & \text{if } X_n \neq \Delta \\ 1 & \text{if } X_n = \Delta \end{cases} \qquad x > 0$$

2. *The embedded process $(X_n)_{n\in\mathbb{N}_0}$ is a discrete-time homogeneous Markov chain with state space $E \cup \{\Delta\}$ and transition probabilities given by*

$$p_{ij} = \begin{cases} \dfrac{q_{ij}}{q_i} & \text{if } i,j \in E,\ i \neq j, \text{ and } q_i \neq 0 \\ 0 & \text{if } i,j \in E \text{ and } i = j \text{ or } q_i = 0 \text{ or if } i = \Delta \neq j \\ 1 & \text{if } i \in E,\ j = \Delta \text{ and } q_i = 0 \text{ or if } i = j = \Delta \end{cases}$$

Note that if $X_n = \Delta$ or $q_{X_n} = 0$ then the n'th holding time $\tau_{n+1} - \tau_n$ is infinitely long and the embedded process is absorbed in Δ. Note also that for any $i \in E$ with $q_i \neq 0$ we have

$$1 = \sum_{j\in E} p_{ij} = \sum_{j\in E\setminus\{i\}} \frac{q_{ij}}{q_i}$$

so that $q_i = \sum_{j\in E\setminus\{i\}} q_{ij}$ holds for all $i \in E$.

The Poisson process with intensity $\lambda > 0$ has iid exponential inter-arrival times with intensity λ and embedded process with transition probabilities given by

$$p_{i,i+1} = 1 \qquad i \in \mathbb{N}_0$$

We shall not prove Theorem 14. Nor will we prove that a process with a distribution given as in the theorem is actually a Markov chain. But a few comment may be in order:

1. That the embedded process is a homogeneous Markov chain follows from a suitable strong Markov property.
2. That the holding times must be exponentially distributed with a parameter only depending on the current state of the Markov chain follows from the Markov property and the fact that the exponential distribution is only memoryless continuous distribution: If the time points $t_1, \ldots, t_n = t, \ldots, t_{n+k} = t + s$ are chosen sufficiently close to each other then

$$\begin{aligned} &P\{\tau_1 > t + s | \tau_1 > t, X_0 = i\} \\ &\qquad \approx P\{X(t_{n+1}) = \ldots = X(t_{n+k}) = i | X(0) = \ldots = X(t_n) = i\} \\ &\qquad = P\{X(t_{n+1}) = \ldots = X(t_{n+k}) = i | X(t_n) = i\} \end{aligned}$$

   which implies first that the distribution of τ_1 may only depend on the past through the value i of $X(t_n)$. Moreover as

$$\begin{aligned} &P\{X(t_{n+1}) = \ldots = X(t_{n+k}) = i | X(t_n) = i\} \\ &\quad = P\{X(t_{n+1} - t_n) = \ldots = X(t_{n+k} - t_n) = i | X(0) = i\} \approx P\{\tau_1 > s | X(0) = i\} \end{aligned}$$

   it follows that the conditional distribution of τ_1 given $X(0) = i$ must be exponential.

The theorem states that a continuous-time homogeneous Markov chain is just a discrete time homogeneous Markov chain (the embedded process) with exponentially distributed time spans between jumps (the holding times) only depending on the current state. Equivalently, we may think of a continuous-time Markov chain as a process with transitions given by the Competition theorem (Theorem 7) as follows: As soon as we jump to state i we initiate independent exponential waiting times, one for each state in the state space, with intensities q_{ij}, $j \neq i$. When the first waiting time "runs out", we jump to the state corresponding to this waiting time. Thus we jump after an exponential time with intensity $q_i = \sum_{j \in E \setminus \{i\}} q_{ij}$ and to state j with probability q_{ij}/q_i.

2.2 Infinitesimal generator

The diagonal elements of the matrix Q are not determined by Theorem 14. It is customary to let $q_{ii} = -q_i$. Thus the matrix $Q = [q_{ij}]_{i,j \in E}$ has

$$q_{ij} \geq 0 \quad i \neq j \qquad \text{and} \qquad \sum_{j \in E} q_{ij} = 0 \tag{4}$$

This matrix is called the *infinitesimal generator* of the Markov chain. Clearly, this matrix determines the transition probabilities of the Markov chain. Alternatively (and equivalently) one may specify the distribution of the Markov chain by specifying the *jump matrix*, containing the transition probabilities for the embedded chain, and the q_i's. These are usually denoted by

$$\Pi = [\pi_{ij}]_{i,j \in E} \quad \text{and} \quad \lambda = (\lambda_i)_{i \in E} \tag{5}$$

where

$$\pi_{ij} = p_{ij} \quad \text{and} \quad \lambda_i = q_i$$

are given in Theorem 14.

If $(P(t))_{t \geq 0}$ is the transition semigroup for a regular jump HMC, then we must have:

$$p_{ii}(t) = P\{X(t) = i | X(0) = i\} \geq P\{\tau_1 > t | X_0 = i\} = e^{-tq_i} \to 1 \quad \text{as } t \to 0 \tag{6}$$

and for $i \neq j$

$$p_{ij}(t) = P\{X(t) = j | X(0) = i\} \leq P\{\tau_1 \leq t | X_0 = i\} = 1 - e^{-tq_i} \to 0 \quad \text{as } t \to 0 \tag{7}$$

Thus the semigroup is *continuous* in the sense that

$$P(t) \to P(0) \qquad \text{component-wise as } t \to 0$$

From this it follows that for all $t \geq 0$

$$P(t+s) \to P(t) \qquad \text{component-wise as } s \to 0$$

(see Brémaud, Problem 8.2.1).

Returning to the infinitesimal generator, we see that

$$\begin{aligned} p_{ii}(t) &= P\{\tau_1 > t | X_0 = i\} + P\{\tau_2 \le t, X(t) = i | X(0) = i\} \\ &= 1 - tq_i + o(t) \end{aligned} \tag{8}$$

where $o(t)$ denotes any term which divided by t goes to 0 as $t \to 0$, since

$$\begin{aligned} P\{\tau_2 \le t, & X(t) = i | X(0) = i\} \le P\{\tau_2 \le t | X(0) = i\} \le P\{\tau_1 \le t, \tau_2 - \tau_1 \le t | X_0 = i\} \\ &= \sum_{j \in E} P\{\tau_1 \le t, \tau_2 - \tau_1 \le t, X_1 = j | X_0 = i\} \\ &= \sum_{j \in E} \frac{q_{ij}}{q_i}(1 - e^{-q_i t})(1 - e^{-q_j t}) = (1 - e^{-q_i t}) \sum_{j \in E} \frac{q_{ij}}{q_i}(1 - e^{-q_j t}) = o(t) \end{aligned}$$

as

$$\sum_{j \in E} \frac{q_{ij}}{q_i}(1 - e^{-q_j t}) \to 0 \qquad \text{as } t \to 0.$$

by dominated convergence. Equation (8) may be written in a perhaps more familiar manner

$$\frac{p_{ii}(t) - 1}{t} \to -q_i = q_{ii} \quad \text{as } t \to 0.$$

Similarly we get

$$\begin{aligned} p_{ij}(t) &= P\{X(t) = j, \tau_2 > t | X(0) = i\} + P\{X(t) = j, \tau_2 \le t | X(0) = i\} \\ &= P\{\tau_1 \le t, \tau_2 > t, X_1 = j | X_0 = i\} + P\{\tau_2 \le t, X(t) = j | X(0) = i\} \\ &= P\{\tau_1 \le t, X_1 = j | X_0 = i\} - P\{\tau_1 \le t, \tau_2 \le t, X_1 = j | X_0 = i\} \\ &\qquad + P\{\tau_2 \le t, X(t) = j | X(0) = i\} \\ &= (1 - e^{-tq_i})\frac{q_{ij}}{q_i} + o(t) = q_{ij}t + o(t) \end{aligned} \tag{9}$$

We see that the "parameters" Q not only determine the distribution of the Markov chain but also have interpretations of "instantaneous risks" or *intensities* of jumping from i to j. Hence, a common way of specifying a Markov chain model is to give the intensities of various jumps.

A useful tool for obtaining insight into the structure of a continuous-time Markov chain is the *intensity diagram*, which is just the transition graph of the underlying embedded Markov chain (usually ignoring any transitions to Δ). If one puts numbers next to the edges, then one would put the intensities, rather than the jump probabilities, next to the edges in the intensity diagram.

2.3 Uniform Markov chains

Proving Theorem 14 is too involved for this course. However, there is a nice subset of the regular jump Markov chains, which are easily seen to be Markov chains.

Definition 15 (Uniform Markov chain) *Let $(N_t)_{t\geq 0}$ be a Poisson process with intensity $\lambda > 0$ and let $(\hat{X}_n)_{n\in\mathbb{N}_0}$ be a discrete-time homogeneous Markov chain, independent of $(N_t)_{t\geq 0}$, with state space E and transition probabilities $K = [k_{ij}]_{i,j\in E}$. The **uniform Markov chain** $(X_t)_{t\geq 0}$ with **clock** $(N_t)_{t\geq 0}$ and **subordinated chain** $(\hat{X}_n)_{n\in\mathbb{N}_0}$ is given by*

$$X_t = \hat{X}_{N(t)} \quad t \geq 0$$

Observe that the subordinated chain of a uniform Markov chain is not generally the embedded chain of the uniform Markov chain: The subordinated chain may have $k_{ii} > 0$ for some i whereas the embedded chain will have $p_{ii} = 0$ for all $i \in E$. Thus it is not clear that a uniform Markov chain is a Markov chain at all. Luckily we may prove the following result.

Theorem 16 (A uniform Markov chain is a HMC) *The uniform Markov chain $(X_t)_{t\geq 0}$ with clock $(N_t)_{t\geq 0}$ with intensity λ and subordinated chain $(\hat{X}_n)_{n\in\mathbb{N}_0}$ with transition matrix K is a regular jump homogeneous Markov chain with transition probabilities given by*

$$P(t) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} K^n \tag{10}$$

and infinitesimal generator given by

$$Q = \lambda(K - I). \tag{11}$$

Proof Clearly, the process is a regular jump process by construction: The regularity follows by observing that the number of jumps made by X_t in finite time is less than the number of jumps made by the underlying Poisson process, which is a regular jump process. To show that it is a homogeneous Markov chain with transition probabilities given by (10), it suffices to show that

$$\begin{aligned} &P\{X(t_0) = i_0, \dots, X(t_{n-1}) = i_{n-1}, X(t_n) = i, X(t_{n+1}) = j\} \\ &\qquad = P\{X(t_0) = i_0, \dots, X(t_{n-1}) = i_{n-1}, X(t_n) = i\} p_{ij}(t_{n+1} - t_n) \end{aligned}$$

We may split the right hand side according to the number of jumps made by the clock until time t_n and the number of jumps made between time t_n and t_{n+1}:

$$\begin{aligned} &P\{X(t_0) = i_0, \dots, X(t_{n-1}) = i_{n-1}, X(t_n) = i, X(t_{n+1}) = j\} \\ &\quad = \sum_{k,l=0}^{\infty} P\{X(t_0) = i_0, \dots, X(t_n) = i, X(t_{n+1}) = j, N(t_n) = k, N(t_{n+1}) - N(t_n) = l\} \end{aligned}$$

Each of these terms may be rewritten in terms of the clock and the subordinated chain:

$$\begin{aligned} &P\{X(t_0) = i_0, \dots, X(t_n) = i, X(t_{n+1}) = j, N(t_n) = k, N(t_{n+1}) - N(t_n) = l\} \\ &\qquad = P\Big(A \cap \{\hat{X}_k = i, \hat{X}_{k+l} = j, N(t_{n+1}) - N(t_n) = l\}\Big) \end{aligned}$$

where

$$\begin{aligned} A \cap \{\hat{X}_k = i\} &= \{X(t_0) = i_0, \dots, X(t_n) = i, N(t_n) = k\} \\ &= \Big\{ \Big((\hat{X}_m)_{m=0,\dots,k}, N(t_0), N(t_1) - N(t_0), \dots, N(t_n) - N(t_{n-1}) \Big) \in B \Big\} \end{aligned}$$

The set B may be written out explicitly but it is complicated and all we need is the fact that it can be done. Using the Markov property of the subordinated chain, the independent increments of the clock, and the fact that the subordinated chain and the clock are independent we may proceed as follows:

$$\begin{aligned} P\Big(A \cap & \{\hat{X}_k = i, \hat{X}_{k+l} = j, N(t_{n+1}) - N(t_n) = l\} \Big) \\ &= P\{\hat{X}_{k+l} = j, N(t_{n+1}) - N(t_n) = l | \hat{X}_k = i\} \cdot P\Big(A \cap \{\hat{X}_k = i\} \Big) \\ &= P\{\hat{X}_{k+l} = j | \hat{X}_k = i\} \cdot P\{N(t_{n+1}) - N(t_n) = l\} \cdot P\Big(A \cap \{\hat{X}_k = i\} \Big) \\ &= k_{ij}(l) \cdot \frac{(\lambda(t_{n+1} - t_n))^l}{l!} e^{-\lambda(t_{n+1}-t_n)} \cdot P\Big(A \cap \{\hat{X}_k = i\} \Big) \end{aligned}$$

It follows that

$$\begin{aligned} P\{X(t_0) = i_0, \dots, & X(t_{n-1}) = i_{n-1}, X(t_n) = i, X(t_{n+1}) = j\} \\ &= \sum_{k,l=0}^{\infty} k_{ij}(l) \frac{(\lambda(t_{n+1} - t_n))^l}{l!} e^{-\lambda(t_{n+1}-t_n)} P\Big(A \cap \{\hat{X}_k = i\} \Big) \\ &= \sum_{k=0}^{\infty} P\Big(A \cap \{\hat{X}_k = i\} \Big) \sum_{l=0}^{\infty} k_{ij}(l) \frac{(\lambda(t_{n+1} - t_n))^l}{l!} e^{-\lambda(t_{n+1}-t_n)} \\ &= P\{X(t_0) = i_0, \dots, X(t_{n-1}) = i_{n-1}, X(t_n) = i\} p_{ij}(t_{n+1} - t_n) \end{aligned}$$

proving that a uniform Markov chain is a regular jump Markov chain and that the transition probabilities are given by (10). To find the infinitesimal generator, note that the embedded chain of the uniform Markov chain is the Markov chain

$$X_n = \begin{cases} \hat{X}_{\hat{\tau}_n} & \text{if } \hat{\tau}_n < \infty \\ \Delta \, (\notin E) & \text{if } \hat{\tau}_n = \infty \end{cases} \qquad n \in \mathbb{N}_0$$

where $\hat{\tau}_0 = 0$ and

$$\hat{\tau}_n = \inf\{k > \hat{\tau}_{n-1} : \hat{X}_k \neq \hat{X}_{\hat{\tau}_{n-1}}\} \quad n \in \mathbb{N} \quad (\text{with } \inf \emptyset = \infty)$$

It is easily shown (see Brémaud, Problem 2.7.1) that the transition probabilities of $(X_n)_{n \in \mathbb{N}_0}$ are given by

$$p_{ij} = P\{X_1 = j | X_0 = i\} = \sum_{l=1}^{\infty} P\{\hat{X}_l = j, \tau_1 = l | \hat{X}_0 = i\} = \sum_{l=1}^{\infty} k_{ij} k_{ii}^{l-1} = \frac{k_{ij}}{1 - k_{ii}}$$

when $k_{ii} \neq 1$. Moreover, the transition times of the uniform Markov chain are given by

$$\tau_n = \begin{cases} T_{\hat{\tau}_n} & \text{if } \hat{\tau}_n < \infty \\ \infty & \text{if } \hat{\tau}_n = \infty \end{cases}$$

where $0 = T_0 < T_1 < T_2 < \ldots$ are the arrival times of the clock. Now

$$\begin{aligned}
P\{\tau_1 > t | X_0 = i\} &= \sum_{l=1}^{\infty} P\{T_l > t, \hat{\tau}_1 = l | X_0 = i\} \\
&= \sum_{l=1}^{\infty} P\{T_l > t | \hat{\tau}_1 = l, X_0 = i\} P\{\hat{\tau}_1 = l | X_0 = i\} \\
&= \sum_{l=1}^{\infty} \sum_{j=0}^{l-1} \frac{(\lambda t)^j}{j!} e^{-\lambda t} k_{ii}^{l-1}(1 - k_{ii}) = e^{-\lambda t} \sum_{j=0}^{\infty} \sum_{l=j+1}^{\infty} k_{ii}^{l-1}(1 - k_{ii}) \cdot \frac{(\lambda t)^j}{j!} \\
&= e^{-\lambda t} \sum_{j=0}^{\infty} k_{ii}^j \frac{(\lambda t)^j}{j!} = e^{-t \cdot \lambda(1 - k_{ii})}
\end{aligned}$$

from which is follows that $q_i = \lambda(1 - k_{ii})$. Thus

$$q_{ii} = -q_i = \lambda(k_{ii} - 1)$$

and

$$q_{ij} = p_{ij} q_i = \frac{k_{ij}}{1 - k_{ii}} \lambda(1 - k_{ii}) = \lambda k_{ij}$$

proving (11). □

We may think of a uniform Markov chain as a Markov chain where the time between transitions are iid exponentially distributed –there is a uniform rate of transitions– but where some of these transitions –when the subordinated chain "moves" from one state to the same state– are unobservable. The uniform rate of transitions makes the uniform Markov chain easy to handle mathematically; the price to pay is that we need to include unobservable transitions.

A natural question is which regular jump Markov chains are *uniformisable*, i.e may be constructed as a uniform Markov chain. The answer is given by

Theorem 17 (Uniformisation) *A regular jump Markov chain with state space E and infinitesimal generator Q such that $\sup_{i \in E} q_i < \infty$ is uniformisable.*

Proof Put $\lambda \geq \sup_{i \in E} q_i$ and construct a uniform Markov chain with clock with intensity λ and subordinated chain with transition probabilities given by

$$k_{ij} = \begin{cases} \dfrac{q_{ij}}{\lambda} & \text{if } i \neq j \\ 1 - \dfrac{q_i}{\lambda} & \text{if } i = j \end{cases}$$

Then this uniform Markov chain has the same distribution as the regular Markov chain with infinitesimal generator Q. □

In particular, any regular jump Markov chain on a finite state space is uniformisable.

2.4 Differential equations

2.4.1 Transition probabilities

Equations (8) and (9) may be interpreted as the derivatives of the transition probabilities at 0. We may write them more compactly as

$$P'(0) = Q$$

Using the Chapman-Kolmogorov equations (Theorem 11), it is tempting to write

$$P'(t) \approx \frac{1}{h}(P(t+h) - P(t)) = \begin{cases} \frac{1}{h}(P(h) - P(0)) \cdot P(t) \approx QP(t) \\ P(t) \cdot \frac{1}{h}(P(h) - P(0)) \approx P(t)Q \end{cases}$$

Here the first "$\approx$" should be interpreted as if the left hand side has a (componentwise) limit as $h \to 0$, then this limit is the matrix of derivatives of the transition probabilities at t. The second "$\approx$" represents something which we would hope to be true as $h \to 0$. If the state space E is infinite then each element of the various matrix products is (potentially) an infinite sum and our "$\approx$" is a question of whether we may interchange summation and taking the limit as $h \to 0$. If E is finite, the sums are finite and we obtain two systems of differential equations. Well, almost: So far we have only obtained right-hand derivatives as $h > 0$ has been implicitly assumed.

It turns out that for regular jump Markov chains, both system of differential equations are true, though we need an additional assumption to prove the second one.

Theorem 18 (Kolmogorov's differential systems) *If $(X(t))_{t\geq 0}$ is a regular jump Markov chain on a at most countable state space E with infinitesimal generator Q, then* ***Kolmogorov's backward differential system***

$$P'(t) = QP(t) \qquad t \geq 0 \tag{12}$$

i.e.

$$p'_{ij}(t) = \sum_{k\in E} q_{ik}p_{kj}(t) = -q_i p_{ij}(t) + \sum_{k\neq i} q_{ik}p_{kj}(t) \quad i,j \in E, t \geq 0 \tag{13}$$

is satisfied. If moreover,

$$\sum_{j\in E} p_{ij}(t)q_j < \infty \tag{14}$$

then also ***Kolmogorov's forward differential system***

$$P'(t) = P(t)Q \qquad t \geq 0 \tag{15}$$

i.e.

$$p'_{ij}(t) = \sum_{k\in E} p_{ik}(t)q_{kj} = -p_{ij}(t)q_j + \sum_{k\neq j} p_{ik}(t)q_{kj} \quad i,j \in E, t \geq 0 \tag{16}$$

is satisfied.

Note that assumption (14) is satisfied for Markov chains on finite state spaces as well as uniformisable Markov chains. Note also that the assumption ensures that $\sum_{k\in E} p_{ik}(t)q_{kj}$ is well-defined.

Before giving the proof we state the following lemma without proof:

Lemma 19 *Let $f:\mathbb{R}\to\mathbb{R}$ be a continuous function. Suppose that the limit*

$$f'(t)=\lim_{h\downarrow 0}\frac{1}{h}(f(t+h)-f(t))$$

exists for all $t\in\mathbb{R}$ and is continuous. Then f is differentiable and has derivative f'.

Thus it suffices to find right-hand derivatives and show that these are continuous.

Proof of Theorem 18 For the backward equations we wish to show that for an arbitrary $i\in E$

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h}=\frac{p_{ii}(h)-1}{h}p_{ij}(t)+\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\to\sum_{k\in E}q_{ik}p_{ik}(t)\quad\text{as } h\to 0$$

Let $E_N\uparrow E\setminus\{i\}$ with each E_N finite. Then

$$\begin{aligned}\sum_{k\in E_N}\frac{p_{ik}(h)}{h}p_{kj}(t)\le\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)&\le\sum_{k\in E_N}\frac{p_{ik}(h)}{h}p_{kj}(t)+\sum_{k\notin E_N\cup\{i\}}\frac{p_{ik}(h)}{h}\\&\le\sum_{k\in E_N}\frac{p_{ik}(h)}{h}p_{kj}(t)+\frac{1-p_{ii}(h)}{h}-\sum_{k\in E_N}\frac{p_{ik}(h)}{h}\end{aligned}$$

Letting $h\to 0$ we obtain

$$\begin{aligned}\sum_{k\in E_N}q_{ik}p_{kj}(t)&\le\liminf_{h\to 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\\&\le\limsup_{h\to 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\le\sum_{k\in E_N}q_{ik}p_{kj}(t)-q_{ii}-\sum_{k\in E_N}q_{ik}\end{aligned}$$

Letting $N\to\infty$ gives the desired result. As $\sum_{k\neq i}q_{ik}p_{ik}(t)\le\sum_{k\neq i}q_{ik}=q_i<\infty$, $t\to\sum_{k\in E}q_{ik}p_{ik}(t)$ is continuous, and the right-hand derivative is also the left-hand derivative.

For the forward equations we wish to show that

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h}=\sum_{k\in E}p_{ik}(t)\frac{p_{kj}(h)-\delta_{kj}}{h}\to\sum_{k\in E}p_{ik}(t)q_{kj}\quad\text{as } h\to 0$$

However, since by (6) and (7), we have

$$\left|\frac{p_{kj}(t)-\delta_{kj}}{h}\right|\le\frac{1-e^{-hq_k}}{h}\le q_k$$

the result follows by dominated convergence. As $q_{kj}\le q_k$, $t\to\sum_{k\in E}p_{ik}(t)q_{kj}$ is continuous, and the right-hand derivative is also the left-hand derivative. □

It may be useful to write the differential equations as integral equations. The backward equations may be written as

$$p_{ij}(t) = \delta_{ij} e^{-q_i t} + \int_0^\top \sum_{k \neq i} q_{ik} e^{-q_i(t-s)} p_{kj}(s) ds \quad i, j \in E, t \geq 0, \tag{17}$$

where δ_{ij} is Kronecker's delta, and the forward equations as

$$p_{ij}(t) = \delta_{ij} e^{-q_j t} + \int_0^\top \sum_{k \neq j} p_{ik}(s) q_{kj} e^{-q_j(t-s)} ds \quad i, j \in E, t \geq 0. \tag{18}$$

(Differentiate (17) and (18) to obtain (13) and (16)). Thus an interpretation of the backward equations is that to move from i to j in time t we first stay in i until time $t - s$ ($\lambda_i e^{-\lambda_i(t-s)}$), then jump to k (q_{ik}/q_i) and finally move from k to j in the remaining time ($p_{kj}(s)$); if $i = j$ then an extra term ($e^{-q_i t}$) is added to account for the fact that we can "move" from i to i by not changing state at all. In the forward equations we instead move from i to k in time s, jump to j and stay there. Thus in the backward equation we look back to the first state change; in the forward we look ahead to the last change of state.

2.4.2 Finite state space: Matrix exponentials

In practice, the differential equation will often have to be solved numerically. If the Markov chain is uniformisable, using (10) may be numerically better than solving the differential equations. It is tempting to rewrite (10) formally as

$$P(t) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(t\lambda K)^n}{n!} = e^{-\lambda t} e^{t\lambda K} = e^{tQ}$$

Indeed when Q is an infinitesimal generator on a finite state space E, defining

$$e^{tQ} = \sum_{n=0}^{\infty} \frac{(tQ)^n}{n!} \qquad t \geq 0 \tag{19}$$

turns out to be unproblematic, and it is easily shown that

$$P(t) = e^{tQ} \qquad t \geq 0$$

is the unique solution to the forward and backward differential equations (see Brémaud, appendix 2.2 for details).

Note that if u is a left eigenvector for Q corresponding to the eigenvalue of λ , i.e.

$$u^\top Q = \lambda u^\top \qquad \text{(or equivalently } Q^\top u = \lambda u\text{)}$$

then

$$u^\top P(t) = u^\top \exp(tQ) = \sum_{n=0}^{\infty} \frac{t^n}{n!} u^\top Q^n = \sum_{n=0}^{\infty} \frac{t^n}{n!} \lambda^n u^\top = e^{t\lambda} u^\top$$

Hence if we can find eigenvectors $u_1, \ldots, u_k$ with eigenvalues $\lambda_1, \ldots, \lambda_k$ such that the i'th unit vector e_i can be written as $\sum_{l=1}^k c_l u_l$ then

$$(p_{ij}(t))_{j\in E} = e_i^\top P(t) = \sum_{l=1}^k c_l u_l^\top P(t) = \sum_{l=1}^k c_l \exp(t\lambda_l) u_l^\top$$

More generally, if Q is diagonalisable in the sense that there exist matrices U and V such that $VU^\top = I$ and $\Lambda = U^\top QV$ is a diagonal matrix with elements $\lambda_1, \ldots, \lambda_k$, then (e.g.)

$$P(t) = V\boldsymbol{diag}\left(e^{\lambda_1 t}, \ldots, e^{\lambda_k t}\right) U^\top$$

If Q has k distinct eigenvalues, then the columns of U are the left eigenvectors, the columns of V the right eigenvectors and $\lambda_1, \ldots, \lambda_k$ the eigenvalues (of which one is 0); see Brémaud, Example 8.3.1 and appendix 2.

2.4.3 Distribution

We can also give a system of differential equations for the distribution of the Markov chain at time t:

Theorem 20 (Kolmogorov's global differential system) *Let $\mu(t)$ denote the distribution of X_t when $(X(t))_{t\geq 0}$ is a continuous-time Markov chain with infinitesimal generator Q and initial distribution $\mu = \mu(0)$. If*

$$\sum_{i\in E} q_i \mu_i(t) < \infty$$

*then $\mu(t)$ satisfy **Kolmogorov's global differential system***

$$\mu'(t)^\top = \mu(t)^\top Q$$

i.e.

$$\mu_i'(t) = -\mu_i(t)q_i + \sum_{j\neq i} \mu_j(t) q_{ji} \quad i \in E, t \geq 0$$

Proof Since

$$\mu(t+h)^\top - \mu(t)^\top = \mu(t)^\top (P(h) - I)$$

we obtain

$$\frac{\mu_i(t+h) - \mu_i(t)}{h} = \sum_{j\in E} \frac{p_{ij}(h) - \delta_{ij}}{h} \mu_j(t) \to \sum_{j\in E} q_{ij}\mu_j(t)$$

by dominated convergence as in the proof of the forward differential equations. □

2.5 Explosion

Given a matrix Q satisfying (4) we can construct a process using the construction in Theorem 14. Clearly this process will be a jump process; it will even be Markov, though that is less obvious and difficult to show. The question is whether it will be a *regular* jump process? Stated differently, the question is if $\lim_{n\to\infty} \tau_n$ is finite or not. If it is finite with positive probability, then there is positive probability of infinitely many jumps in finite time, and the process is not a regular jump process.

Example Let $q_{ij} = 0$, $j \notin \{i, i+1\}$, and $q_{i,i+1} > 0$. Thus the embedded chain has transition probabilities given by $p_{i,i+1} = 1$. By monotone convergence,

$$E_i[\lim_{n\to\infty} \tau_n] = \lim_{n\to\infty} E_i[\tau_n] = \lim_{n\to\infty} E_i[\sum_{k=1}^{n} S_k] = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{q_{i+k-1}} = \sum_{k=i}^{\infty} \frac{1}{q_k} \tag{20}$$

where S_k is the kth holding time. We see that $E_i[\lim_{n\to\infty} \tau_n]$ is finite if $q_i \to \infty$ sufficiently fast as $i \to \infty$ (for instance if $q_i = i^2$). Hence if the waiting times between transitions become shorter and shorter sufficiently fast, then there will (with probability 1) be infinitely many transitions in finite time.

We say that Q is *non-explosive* if for any initial distribution

$$P\left\{\lim_{n\to\infty} \tau_n = \infty\right\} = 1 \tag{21}$$

or equivalently

$$P_i\left\{\lim_{n\to\infty} \tau_n = \infty\right\} = 1 \qquad \text{for all } i \in E$$

The example above shows that explosion may occur if the intensities are chosen in a certain way. Note that even if $E_i[\lim_{n\to\infty} \tau_n] = \infty$ we may still have $P\{\lim_{n\to\infty} \tau_n = \infty\} < 1$. Thus, calculating the expectation of $\lim_n \tau_n$ may prove that the Markov chain explodes but it cannot (on its own) ensure non-explosion.

Writing

$$\lim_{n\to\infty} \tau_n = \sum_{k=1}^{\infty} S_k$$

where $S_k = \tau_k - \tau_{k-1}$, then Q is non-explosive if and only if $\sum_{k=1}^{\infty} S_k = \infty$ with probability 1. As

$$\exp\left(-\lambda \sum_{k=1}^{\infty} S_k\right) \in [0;1]$$

for all $\lambda > 0$ and only equal to 0 when $\sum_{k=1}^{\infty} S_k = \infty$ we see that if for all $i \in E$ and some $\lambda > 0$

$$E_i\left[\exp\left(-\lambda \sum_{k=1}^{\infty} S_k\right)\right] = 0 \tag{22}$$

then the chain is non-explosive. And if the chain is non-explosive, then (22) holds for all $i \in E$ and all $\lambda > 0$. The following result translates this criterion for non-explosion into a condition on the infinitesimal generator.

Theorem 21 (Reuter's criterion) *Let Q be an infinitesimal generator satisfying* (4). *Then the following are equivalent:*

1. *Q is non-explosive*
2. *There is a $\lambda > 0$ such that the only non-negative bounded solution to*

$$\lambda x = Qx \tag{23}$$

   is the trivial solution $x = 0$

3. *For all $\lambda > 0$, the only non-negative bounded solution to* (23) *is the trivial solution $x = 0$*

Proof We prove that $2 \Rightarrow 1$ and $1 \Rightarrow 3$; that $3 \Rightarrow 2$ is of course trivial.

$2 \Rightarrow 1$: We prove that "not 1" implies "not 2". Suppose that $P_i\{\lim_{n\to\infty} \tau_n = \infty\} < 1$ for some $i \in E$. Then for any $\lambda > 0$

$$x_i = E_i\left[\exp\left(-\lambda \sum_{k=1}^{\infty} S_k\right)\right] \leq 1 \quad i \in E.$$

Note that $x_i > 0$ for some $i \in E$, since there is a positive probability of explosion. Moreover,

$$\begin{aligned} x_i = E_i\left[e^{-\lambda S_1} \exp\left(-\lambda \sum_{k=2}^{\infty} S_k\right)\right] &= \sum_{j\neq i} E_i\left[e^{-\lambda S_1} 1_{\{X_1=j\}} \exp\left(-\lambda \sum_{k=2}^{\infty} S_k\right)\right] \\ = \frac{q_i}{q_i+\lambda} \cdot \sum_{j\neq i} \frac{q_{ij}}{q_i} E\left[\exp\left(-\lambda \sum_{k=2}^{\infty} S_k\right) \middle| X_1 = j\right] &= \frac{1}{q_i+\lambda} \sum_{j\neq i} q_{ij} x_j \end{aligned}$$

showing that $(x_i)_{i\in E}$ is a solution to (23).

$1 \Rightarrow 3$: We prove that "not 3" implies "not 1". Assume that there is a non-trivial, non-negative bounded solution to (23) for some $\lambda > 0$; call this $(x_i)_{i\in E}$. Then

$$E_i\left[\exp(-\lambda S_1) x_{X_1}\right] = E_i\left[\exp(-\lambda S_1)\right] \sum_{j\neq i} \frac{q_{ij}}{q_i} x_j = \frac{q_i}{q_i+\lambda} \sum_{j\neq i} \frac{q_{ij}}{q_i} x_j = x_i$$

as $(x_i)_{i\in E}$ solves (23). Thus

$$x_i = E_i\left[\exp(-\lambda S_1) x_{X_1}\right] \quad i \in E$$

By induction and the Markov property, we see that for any $i \in E$

$$\begin{aligned} E_i\left[\exp\left(-\lambda \sum_{k=1}^{n} S_k\right) x_{X_n}\right] &= \sum_{j\neq i} E_i\left[\exp\left(-\lambda S_1\right) 1_{\{X_1=j\}} \exp\left(-\lambda \sum_{k=2}^{n} S_k\right) x_{X_n}\right] \\ &= \frac{q_i}{q_i+\lambda} \sum_{j\neq i} \frac{q_{ij}}{q_i} E_i\left[\exp\left(-\lambda \sum_{k=2}^{n} S_k\right) x_{X_n} \middle| X_1 = j\right] \\ &= \frac{1}{q_i+\lambda} \sum_{j\neq i} q_{ij} E_i\left[\exp\left(-\lambda \sum_{k=1}^{n-1} S_k\right) x_{X_{n-1}}\right] \\ &= \frac{1}{q_i+\lambda} \sum_{j\neq i} q_{ij} x_j = x_i \end{aligned}$$

Since x_i is bounded (in $i \in E$), it follows that

$$x_i \leq E_i\left[\exp\left(-\lambda\sum_{k=1}^{n} S_k\right)\right]\cdot\sup_{j\in E} x_j \to E_i\left[\exp\left(-\lambda\sum_{k=1}^{\infty} S_k\right)\right]\cdot\sup_{j\in E} x_j \text{ as } n\to\infty.$$

As $x_i > 0$ for some i we must have $P_i\{\lim_{n\to\infty}\tau_n = \infty\} < 1$. $\square$

Reuter's criterion has a number of useful corollaries:

Corollary 22 *If E is finite, then any generator is non-explosive.*

Proof It follows from Reuter's criterion since a $k \times k$ matrix has at most k eigenvalues. Thus for any $\lambda > 0$, which is not an eigenvalue, the only solution to $Qx = \lambda x$ is $x = 0$. $\square$

Corollary 23 *Suppose that $\sup_{i\in E} q_i < \infty$. Then the generator is non-explosive.*

Of course, this corollary contains the former.

Proof The chain is uniformisable and a uniform Markov chain does not explode, since its clock does not allow more that finitely many jumps in finite time. $\square$

Corollary 24 *Suppose that all $q_i > 0$. Then the generator is non-explosive if the corresponding jump matrix is irreducible and recurrent.*

Proof Suppose that $x = (x_i)_{i\in E}$ is a bounded, non-negative solution to (23) for some $\lambda > 0$. Then

$$\Pi x = x + \lambda z \qquad \text{where } z_i = x_i/q_i$$

and Π is the jump matrix. By induction we obtain

$$\Pi^{n+1}x = x + \lambda\sum_{k=0}^{n}\Pi^k\cdot z$$

As Π is irreducible and recurrent, each element in $\sum_{k=0}^{n}\Pi^k$ tends to infinity as $n\to\infty$. On the other hand, each element in Π^{n+1} is bounded by 1. It follows that z must be 0, proving that Q is non-explosive. $\square$

2.6 State space decomposition

2.6.1 Communication, transience and recurrence

Similarly to the discrete-time Markov chains we say that a state j is *accessible* from the state i if $p_{ij}(t) > 0$ for some $t > 0$. For $p_{ij}(t) > 0$ for some $t > 0$, there must be an

$n \in \mathbb{N}$ such that $\pi_{ij}(n) > 0$, i.e. j must be accessible from i for the embedded process too. Moreover, we have for this choice of n that

$$\begin{aligned} p_{ij}(t) &\geq P\{X_n = j, \tau_n \leq t, \tau_{n+1} > t | X_0 = i\} \\ &= \pi_{ij}(n) P\{\tau_n \leq t, \tau_{n+1} > t | X_0 = i, X_n = j\} > 0 \end{aligned}$$

Since the joint distribution of $(\tau_n, \tau_{n+1} - \tau_n)$ is continuous with support equal to $\mathbb{R}_+^2$, the right hand side is positive for all $t > 0$ if it is for one value of $t > 0$. Thus we see that j is accessible from i if and only if $p_{ij}(t) > 0$ for all $t > 0$ and if and only if $i \to j$ for the embedded chain. The latter condition is of course the more useful one for practical application. As in discrete-time we write $i \to j$ if j is accessible from i and if also $j \to i$ we write $i \leftrightarrow j$ and say that i and j *communicate*. If all states communicate, then the Markov chain is *irreducible*.

The fact that $p_{ij}(t) > 0$ for some $t > 0$ implies that $p_{ij}(t) > 0$ shows that periodicity is not an issue for continuous-time Markov chains (even though the embedded process may be periodic).

Since the communication structure is the same for the continuous-time Markov chain and its embedded process, the state space is decomposed into the union of the set of transient states (for the embedded process) and a number of closed, recurrent (for the embedded process) communication classes:

$$E = T \bigcup \cup_k R_k$$

Clearly, if a state is transient for the embedded process, it is visited only finitely often and this is inherited by the continuous-time Markov chain. Similarly, if a state is recurrent for the embedded process, it is visited infinitely often by the embedded process and therefore also by the continuous-time process. Thus it makes sense to define:

Definition 25 (Recurrence and transience) *For a continuous-time Markov chain a state is* ***recurrent*** *if it is recurrent for the embedded process; if not recurrent, it is* ***transient****.*

Note that for continuous-time Markov chains recurrence and transience are class properties since they are class properties for the embedded process. Note also that an absorbing state for a continuous-time Markov chain is transient, not recurrent. This is due to the fact that a continuous-time Markov chain never leaves an absorbing state and therefore cannot return.

Theorem 26 *If i is recurrent for a continuous-time Markov chain, then the* ***return time***

$$R_i = \inf\{t \geq E_i : X(t) = i\}, \tag{24}$$

where $E_i = \inf\{t \geq 0 : X(t) \neq i\}$ is the ***escape time****, is finite almost surely given $X(0) = i$, and the time spent in state i*

$$\int_0^\infty 1_{\{X(s)=i\}} ds$$

is almost surely infinite. If i is transient and not absorbing, then

$$P_i\{R_i < \infty\} < 1 \quad \textit{and} \quad P_i\left\{\int_0^\infty 1_{\{X(s)=i\}} ds = \infty\right\} = 0.$$

Proof We see that

$$R_i = \begin{cases} \tau_{T_i} & \text{if } T_i < \infty \\ \infty & \text{if } T_i = \infty \end{cases}$$

where T_i is the return time to i for the embedded process. Clearly $R_i = \infty$ if and only if $T_i = \infty$. Moreover

$$\int_0^\infty 1_{\{X(s)=i\}} ds = \sum_{k=1}^{N_i} Y_k$$

where $Y_1, Y_2, \ldots$ are independent exponentially distributed random variables with intensity λ_i, independent of N_i the number of visits to i made by the embedded process. If i is transient and consequently N_i is finite almost surely, then so is the integral. If i is recurrent and hence $N_i = \infty$ then the sum (and hence the integral) is infinite. If the sum is finite, the Y_ks would have to be smaller than 1 eventually and this happens with probability 0. □

2.6.2 Positive and null recurrence

We define positive and null recurrence similarly to the discrete-time case:

Definition 27 (Positive recurrence) *A recurrent state i is **positive recurrent** for a continuous-time Markov chain if $E_i[R_i] < \infty$. If not positive recurrent, it is **null recurrent**.*

When it comes to positive and null recurrence, matters are a bit more complicated. It is quite possible to have $E_i[T_i] < \infty$ and $E_i[R_i] = \infty$ or vice versa as we shall see.

Definition 28 (Invariant measure) *A non-trivial vector $\nu = (\nu(i))_{i\in E}$ is an **invariant measure** if*

$$\nu^\top P(t) = \nu^\top \quad \textit{for all } t \geq 0$$

*If an invariant measure is a probability measure it is called an **invariant distribution** or a **stationary distribution**.*

It follows immediately from Theorem 20 that an invariant distribution, π, for a Markov chain with infinitesimal generator Q satifies the equations

$$\pi^\top Q = 0$$

provided that $\sum_{i\in E} q_i\pi_i < \infty$.

Theorem 29 (Regenerative form of invariant measure) *Let $(X(t))_{t\geq 0}$ be an irreducible, recurrent continuous-time Markov chain with infinitesimal generator Q.*

1. *Then there is an invariant measure ν with $\nu(i) > 0$ for all $i \in E$ and any other invariant measure is proportional to ν.*
2. *The invariant measure ν is a solution to*
$$\nu^\top Q = 0$$
and it is given by
$$\nu(i) = \begin{cases} E_{\mathbf{0}}\left[\int_0^{R_{\mathbf{0}}} 1_{\{X(s)=i\}} ds\right] \\ \frac{1}{q_i} E_{\mathbf{0}}\left[\sum_{n=1}^{T_{\mathbf{0}}} 1_{\{X_n=i\}}\right] \end{cases} \tag{25}$$
3. *The Markov chain $(X(t))_{t\geq 0}$ is positive recurrent if and only if the invariant measure is finite and in this case the invariant distribution of $(X(t))_{t\geq 0}$ is given by*
$$\pi(i) = \frac{1}{E_i[R_i]q_i} \qquad i \in E.$$

Proof We will only prove the result for uniformisable chains; see Brémaud, proof of Theorems 8.5.1 and 8.5.2, for a proof of the generel result.

If ν is invariant for $(P(t))_{t\geq 0}$, then $\nu^\top P(1) = \nu^\top$. Hence ν is invariant for the skeleton chain $(X(n))_{n\in\mathbb{N}_0}$. Since irreducibility and recurrence of $(X(t))_{t\geq 0}$ implies irreducibility and recurrence of $(X(n))_{n\in\mathbb{N}_0}$ (see Brémaud, Problem 8.5.3), ν is unique up to multiplication by a constant.

Let
$$N_i^{(\mathbf{0})} = \sum_{n=1}^{T_{\mathbf{0}}} 1_{\{X_n=i\}}$$
be the number of visits to state i for the subordinated chain before the first return to $\mathbf{0}$. Then
$$\int_0^{R_{\mathbf{0}}} 1_{\{X(s)=i\}} ds = \sum_{k=1}^{N_i^{(\mathbf{0})}} Y_k$$
where $(Y_k)_{k\in\mathbb{N}}$ are iid, exponentially distributed random variables with intensity q_i. It follows that
$$E_{\mathbf{0}}\left[\int_0^{R_{\mathbf{0}}} 1_{\{X(s)=i\}} ds\right] = E_{\mathbf{0}}\left[\sum_{k=1}^{N_i^{(\mathbf{0})}} Y_k\right] = E_{\mathbf{0}}\left[N_i^{(\mathbf{0})}\right] E_{\mathbf{0}}\left[Y_1\right] = \frac{1}{q_i} E_{\mathbf{0}}\left[\sum_{n=1}^{T_{\mathbf{0}}} 1_{\{X_n=i\}}\right]$$

We conclude that the two expression (25) for $\nu(i)$ are the same. Moreover, we see that with $\nu(i)$ given by either of these two expressions we obtain
$$\sum_{i\in E} \nu(i) q_{ij} = \sum_{i\neq j} \mu(i)\frac{q_{ij}}{q_i} + \mu(j)\frac{q_{jj}}{q_j} = \mu(j) - \mu(j) = 0$$

as $(\mu(i))_{i\in E} = (\nu(i)q_i)_{i\in E}$ is an invariant measure for the embedded chain. We conclude that ν is a solution to $\nu^\top Q = 0$.

Now write $Q = \lambda(K - I)$. We see that $\nu^\top Q = 0$ if and only if $\nu^\top K = \nu^\top$, i.e. $\nu^\top Q = 0$ if and only if ν is an invariant measure for K, the transition matrix for the subordinated chain. From (10) we see that if $\nu^\top Q = 0$ then ν is invariant for $(P(t))_{t\geq 0}$. In particular, an invariant measure exists and is given by (25).

The final claim of the theorem follows by the observation that $\sum_{i\in E}\nu(i) = E_{\mathbf{0}}[R_{\mathbf{0}}]$. Hence positive recurrence (of the arbitrary state $\mathbf{0}$) is equivalent to the existence of a finite invariant measure. Moreover, when positive recurrent, the invariant distribution is given by

$$\pi(\mathbf{0}) = \frac{\nu(\mathbf{0})}{E_{\mathbf{0}}[R_{\mathbf{0}}]} = \frac{1}{q_{\mathbf{0}}E_{\mathbf{0}}[R_{\mathbf{0}}]} \qquad \square$$

Note that the theorem gives the relationship

$$\mu(i) = q_i\nu(i)$$

between the invariant measure μ of the embedded process and the invariant measure ν of the continuous-time Markov chain.

As in discrete time, we see that positive and null recurrence are also class properties so that in the state space decomposition

$$E = T \bigcup \cup_k R_k$$

each R_k is either positive or null recurrent for the continuous-time chain. But we repeat that the embedded process does not in itself determined whether a recurrent communication class is positive or null recurrent.

Finally, we have the following useful result:

Theorem 30 (Stationary distribution criteria) *An irreducible continuous-time Markov chain with infinitesimal generator Q is positive recurrent if and only if there is a distribution π such that $\pi^\top Q = 0$.*

Proof Again we will only give a proof for uniformisable chains. By Theorem 29, if a continuous-time Markov chain is positive recurrent, then it has an invariant distribution satifying $\pi^\top Q = 0$.

Suppose conversely that there is a distribution π satisfying $\pi^\top Q = 0$. Then $\pi^\top K = \pi^\top$, where K is the transition matrix for the subordinated chain. If follows that K is positive recurrent. If we let $T_{\mathbf{0}}$ denote the return time to $\mathbf{0}$ for the subordinated chain, then

$$R_{\mathbf{0}} = \sum_{k=1}^{T_{\mathbf{0}}} Y_k$$

where the Y_k's (the holding times for the clock) are iid exponential random variables with intensity λ. By the independence of the clock and the subordinated chain

$$E_{\mathbf{0}}[R_{\mathbf{0}}] = E_{\mathbf{0}}\left[\sum_{k=1}^{T_{\mathbf{0}}} Y_k\right] = E_{\mathbf{0}}[T_{\mathbf{0}}]/\lambda < \infty$$

implying that the continuous-time Markov chain is positive recurrent. $\square$

2.6.3 Stationarity and reversibility

As in the discrete-time case we see that if π is an invariant distribution for a continuous-time Markov chain $(X(t))_{t\geq 0}$ then the Markov chain is *stationary* in the sense that

$$P\{X(t_0) = i_0, X(t_1) = i_1 \ldots, X(t_n) = i_n\}$$
$$= P\{X(t_0 + s) = i_0, X(t_1 + s) = i_1, \ldots, X(t_n + s) = i_n\}$$

for all $n \in \mathbb{N}_0, 0 < t_0 < t_1 < \ldots < t_n, i_0, i_1, \ldots, i_n$ and all $s > 0$. Hence $(X(t))_{t\geq 0}$ and the time-shifted chain $(X(t+s))_{t\geq 0}$ have the same distribution.

If the Markov chain is stationary, then

$$\tilde{p}_{ij}(s) = P\{X(t) = j|X(t+s) = i\} = \frac{p_{ji}(s)\pi(j)}{\pi(i)} \quad i, j \in E, t \geq 0$$

are the transition probabilities for the *time reversed chain*

$$\tilde{X}(s) = X(T - s) \qquad s \in [0; T]$$

where T is fixed but arbitrary. The infinitesimal generator for this chain $\tilde{Q}$ is given by

$$\tilde{q}_{ij} = \frac{\pi(j)q_{ji}}{\pi(i)} \quad i, j \in$$

It follows that if $Q = \tilde{Q}$ or equivalently $P(t) = \tilde{P}(t)$ for all $t \geq 0$ then the reversed chain has the same distribution as the original chain and the Markov chain is said to be *reversible*.

Theorem 31 (Reversal test) *If $(X(t))_{t\geq 0}$ is a continuous-time Markov chain with infinitesimal generator Q and there is a strictly positive distribution π such that for all $i, j \in E$*

$$\pi(i)q_{ij} = \pi(j)q_{ji}$$

then $(X(t))_{t\geq 0}$ is positive recurrent with invariant distribution π.

Proof It suffices to show that $\pi^\top Q = 0$. But the ith element of this vector is

$$\sum_{j\in E} \pi(j)q_{ji} = \sum_{j\in E} \pi(i)q_{ij} = \pi(i) \sum_{j\in E} q_{ij} = 0$$

$\square$

2.7 Long-run behaviour

2.7.1 Ergodicity

Theorem 32 (Ergodic theorem) *If $(X(t))_{t\geq 0}$ is an irreducible positive recurrent Markov chain then*

$$p_{ij}(t) \to \pi(j) \quad \textit{as } t \to \infty \textit{ for any } i, j \in E$$

where π is the invariant distribution. Moreover,

$$\frac{1}{t}\int_0^t h(X(s))ds \xrightarrow{P} \sum_{i\in E} h(i)\pi(i)) \quad \textit{as } t \to \infty$$

for any $h : E \to \mathbb{R}$ for which $\sum_{i\in E} |h(i)|\pi(i) < \infty$.

Proof The proof of the first part may be found in Brémaud (Theorem 8.6.1); the proof of the second part is omitted. □

2.7.2 Absorption

In the non-irreducible case, a continuous-time Markov chain is absorbed in a recurrent class if and only if the embedded process is absorbed in the same class, and the probability of this happening is the same for the continuous-time Markov chain and its embedded process. Recall, however, that an absorbing state for a continuous-time Markov chain (as well as for the embedded process) is transient, so in order to discuss absorption in an absorbing state a little extra work is necessary.

Define a new discrete-time process, $(\breve{X}_n)_{n\in\mathbb{N}_0}$, by

$$\breve{X}_n = \begin{cases} X(\tau_n) & \text{if } \tau_n < \infty \\ \breve{X}_{n-1} & \text{if } \tau_n = \infty \end{cases}$$

Clearly $\breve{X}_n = \tilde{X}_n$ until the chain reaches an absorbing state; then $\breve{X}_n$ stays in the absorbing state, whereas $\tilde{X}_n$ moves on to Δ. It follows that $(\breve{X}_n)_{n\in\mathbb{N}_0}$ is a discrete-time Markov chain with transition probabilities

$$\breve{p}_{ij} = \begin{cases} \frac{q_{ij}}{q_i} & \text{if } q_i > 0 \\ 1 & \text{if } q_i = 0 \text{ and } i = j \\ 0 & \text{if } q_i = 0 \text{ and } i \neq j \end{cases}$$

Moreover, absorbing states for the continuous-time Markov chain $(X(t))_{t\geq 0}$ are absorbing for $(\breve{X}_n)_{n\in\mathbb{N}_0}$ and vice versa. Also, recurrent states for $(X(t))_{t\geq 0}$ are recurrent for $(\breve{X}_n)_{n\in\mathbb{N}_0}$ (but not vice versa since absorbing states are not recurrent in a continuous-time Markov chain). Hence to find the probability of absorption in an absorbing state or a recurrent class for $(X(t))_{t\geq 0}$ we may find the absorption probabilities for $(\breve{X}_n)_{n\in\mathbb{N}_0}$; the probabilities are the same. The probability of an infinite sojourn in a given set of states A is also the same for $(X(t))_{t\geq 0}$ and $(\breve{X}_n)_{n\in\mathbb{N}_0}$.

Note also that the if the continous-time Markov chain is uniformisable, then the subordinated chain and the $(\breve{X}_n)_{n\in\mathbb{N}_0}$-chain have the same absorbing states, recurrent classes and transient states. Hence, absorption probabilities for uniformisable chains can also be found by looking at the subordinated chain.

Similarly to the discrete time, if τ denotes the time to absorption, then for any $i \in \breve{T}$

$$P_i\{\tau > t\} = P_i\{X(t) \in \breve{T}\} = \sum_{j\in\breve{T}} p_{ij}(t)$$

Here $\breve{T}$ denotes the non-absorbing, transient states of the continuous-time Markov chain $(X(t))_{t\geq 0}$; these are the transient states of the discrete-time process $(\breve{X}_n)_{n\in\mathbb{N}_0}$. The continuous-time Markov chain $(X(t))_{t\geq 0}$ and the discrete-time process $(\breve{X}_n)_{n\in\mathbb{N}_0}$ will of course have different times until absorption, but if the chain is uniformisable, then we may write

$$\begin{aligned}\sum_{j\in\breve{T}} p_{ij}(t) &= \sum_{j\in\breve{T}}\sum_{n=0}^{\infty} k_{ij}(n)\frac{(\lambda t)^n}{n!}e^{-\lambda t} \; = \sum_{n=0}^{\infty}\sum_{j\in T} k_{ij}(n)\frac{(\lambda t)^n}{n!}e^{-\lambda t} \\ &= \sum_{n=0}^{\infty} P_i\{\hat{\tau} > n\}\frac{(\lambda t)^n}{n!}e^{-\lambda t}\end{aligned}$$

where $\hat{\tau}$ is the absorption time for the subordinated chain. Putting $K_{\breve{T}} = [k_{ij}]_{i,j\in\breve{T}}$ we know (from Brémaud, Theorem 4.5.2) that

$$P_i\{\hat{\tau} > n\} = e_i^\top K_{\breve{T}}^n 1_{\breve{T}}$$

where e_i is the ith unit vector and 1_T is a vector of 1s. Thus, we can write

$$\begin{aligned}P_i\{\tau > t\} &= \sum_{n=0}^{\infty} e_i^\top K_{\breve{T}}^n 1_{\breve{T}}\frac{(\lambda t)^n}{n!}e^{-\lambda t} = e_i^\top \sum_{n=0}^{\infty} K_{\breve{T}}^n \frac{(\lambda t)^n}{n!}e^{-\lambda t} 1_{\breve{T}} \\ &= e_i^\top e^{t\lambda(I-K_{\breve{T}})}1_{\breve{T}} = e_i^\top e^{tQ_{\breve{T}}}1_{\breve{T}}\end{aligned}$$

where $Q_{\breve{T}} = [q_{ij}]_{i,j\in\breve{T}}$ is the part of the infinitesimal generator corresponding to the transient, non-absorbing states.

2.8 Birth-and-death processes

A much used class of continuous-time Markov chains are the so-called *birth-and-death processes*. These Markov chains have state space $\mathbb{N}_0$ and infinitesimal generators of the form

$$q_{i,i-1} = \delta_i \quad (i > 0) \qquad q_{ii} = -(\delta_i + \beta_i) \qquad q_{i,i+1} = \beta_i$$

(all other q_{ij}s are 0). The β_is are the *birth intensities* and the δ_is are the *death intensities*. Birth-and-death processes are used to describe populations, where births and deaths never happen at the same time. Thus given a population of size i at time t, i.e. $X(t) = i$, there is an exponential waiting time with intensity β_i to the next "birth" and an exponential waiting time with intensity δ_i to the next "death". Hence the next transition happens after an exponential waiting time with intensity $\beta_i + \delta_i$ and it is a birth with probability $\beta_i/(\beta_i + \delta_i)$. Naturally if $X(t) = 0$, no death can take place. In many cases we would want 0 to be an absorbing state (in which case $\beta_0 = 0$), but in some cases we may want to let 0 be a reflecting barrier ($\beta_0 > 0$).

2.8.1 Explosion

In the example on page 19 we see that a birth-and-death process (actually a pure birth-process; see below) may explode if births occur with shorter and shorter time between them as the population grows. Generally, whether a birth-or-death process explodes or not is a question of whether the birth intensities increase too fast for the death intensities to compensate.

Reuter's criterion (Theorem 21) tells us that we have to look at whether the system of equations

$$\lambda x = Qx$$

has a non-trivial, non-negative bounded solution for any/all $\lambda > 0$. Writing this system of equations out in more detail, we get

$$\begin{aligned} \lambda x_0 &= -\beta_0 x_0 + \beta_0 x_1 \\ \lambda x_i &= \delta_i x_{i-1} - (\beta_i + \delta_i) x_i + \beta_i x_{i+1} \qquad i \in \mathbb{N} \end{aligned} \tag{26}$$

Recall that this system of equations have a bounded non-trivial non-negative solution for all $\lambda > 0$ if and only if it does for some $\lambda > 0$. Hence we may look at the special case $\lambda = 1$. In fact this is clear from (26): Assuming that $\lambda = 1$ corresponds to replacing δ_i and β_i by δ_i/λ and β_i/λ, which is just a transformation of the time scale (as in measuring time in weeks rather that in months); clearly if the process exploded on one time-scale it explodes on every time scale.

Letting $\lambda = 1$ we can calculate the only solution for a fixed value of x_0 by finding first x_1, then x_2 and so on. We see that all solutions are obtained in this way and may therefore assume that $x_0 = 1$ as all other solutions are proportional to the one obtained with $x_0 = 1$. From (26) we obtain

$$\begin{aligned} y_1 &= x_1 - x_0 = 1/\beta_0 \\ y_{i+1} &= x_{i+1} - x_i = \frac{\delta_i}{\beta_i} y_i + \frac{x_i}{\beta_i} \quad i = 2, 3, \dots \end{aligned}$$

from which it follows that the $(x_i)_{i\in\mathbb{N}}$-sequence is strictly increasing and that

$$\begin{aligned} y_{i+1} &= \frac{x_i}{\beta_i} + \frac{\delta_i}{\beta_i}\left(\frac{x_{i-1}}{\beta_{i-1}} + \frac{\delta_{i-1}}{\beta_{i-1}} y_{i-1}\right) = \dots \\ &= \frac{1}{\beta_i} x_i + \frac{\delta_i}{\beta_i \beta_{i-1}} x_{i-1} + \cdots + \frac{\delta_i \cdots \delta_1}{\beta_i \cdots \beta_0} x_0 \\ &\geq \frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \cdots + \frac{\delta_i \cdots \delta_1}{\beta_i \cdots \beta_0} \end{aligned}$$

Since

$$\sum_{i=1}^{\infty} y_i = \lim_{i\to\infty} x_i - 1$$

the solution $(x_i)_{i\in\mathbb{N}_0}$ is unbounded and the process does not explode if

$$\sum_{i=1}^{\infty} \left(\frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \cdots + \frac{\delta_i \cdots \delta_1}{\beta_i \cdots \beta_0}\right) = \infty$$

On the other hand, we also have that

$$y_i \leq \left(\frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \cdots + \frac{\delta_i \cdots \delta_1}{\beta_i \cdots \beta_0}\right) x_i$$

so that

$$
\begin{aligned}
x_{i+1} &\leq \left(1 + \frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \cdots + \frac{\delta_i \cdots \delta_1}{\beta_i \cdots \beta_0}\right) x_i \\
&\leq \exp\left(\frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \cdots + \frac{\delta_i \cdots \delta_1}{\beta_i \cdots \beta_0}\right) x_i \leq \ldots \\
&\leq \exp\left(\sum_{k=0}^{i} \frac{1}{\beta_k} + \frac{\delta_k}{\beta_k \beta_{k-1}} + \cdots + \frac{\delta_k \cdots \delta_1}{\beta_k \cdots \beta_0}\right) x_0
\end{aligned}
$$

We see that if

$$\sum_{i=1}^{\infty} \left(\frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \cdots + \frac{\delta_i \cdots \delta_1}{\beta_i \cdots \beta_0}\right) < \infty$$

then $(x_i)_{i \in \mathbb{N}_0}$ is a bounded solution and the process may explode.

Hence we have obtained

Theorem 33 (Reuter's criterion for birth-and-death processes) *Let Q be an infinitesimal generator for a birth-and-death process with all birth intensities strictly positive. Then the process is non-explosive if and only if*

$$\sum_{i=1}^{\infty} \left(\frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \cdots + \frac{\delta_i \cdots \delta_1}{\beta_i \cdots \beta_0}\right) = \infty \tag{27}$$

In the argument above we have used that all birth intensities are strictly positive. If some of the birth intensities are 0, then the process can only explode if it starts in a state larger than or equal to $i_0 = 1 + \sup\{i \in \mathbb{N}_0 : \beta_i = 0\}$ and then if and only if

$$\sum_{i=i_0}^{\infty} \left(\frac{1}{\beta_i} + \frac{\delta_i}{\beta_i \beta_{i-1}} + \cdots + \frac{\delta_i \cdots \delta_{i_0+1}}{\beta_i \cdots \beta_{i_0}}\right) < \infty$$

To see this, consider the process

$$Y(t) = \min(X(t) - i_0, 0)$$

where $(X(t))_{t \geq 0}$ is the original birth-and-death process. This is a birth-and-death process with infinitesimal generator $[q_{i+i_0, j+i_0}]_{i,j \in \mathbb{N}_0}$ and it explodes if and only if $(X(t))_{t \geq 0}$ does.

2.8.2 Linear birth-and-death processes

Consider a population of individuals where (at time t) the i individuals have independent exponentially distributed life times (=times to death) with intensity δ. Then the time until the next individual dies, is exponentially distributed with intensity $i\delta$ by the Competition theorem (Theorem 7). If they similarly have iid exponential waiting times until giving birth to a new individual and this is independent of the life times then we obtain a birth-and-death process with intensities

$$\beta_i = i\beta \qquad \delta_i = i\delta \qquad i \in \mathbb{N}_0$$

This is known as a *linear birth-and-death process*.

2.8.3 Birth-and-death processes with immigration

In some cases one would like the population size described by a birth and death process to be able to increase not only by individuals in the population giving birth to new individuals but also by new individuals from outside the population joining the population; this is called immigration. If immigration takes place independently of the population size, then we get birth intensities

$$q_{i,i+1} = \alpha + \beta_i \qquad i \in \mathbb{N}_0$$

where α is the rate of immigration (i.e. immigration happens independently of births and deaths after an exponential waiting time with intensity α) and β_i is the intensity at which new individuals are born when the population size is i.

2.8.4 Pure birth and death processes

Birth and death process, where all death intensities are 0, are called *(pure) birth processes*. Similarly, if all birth intensities are 0, we have a *(pure) death process*. A Poisson process is a birth process; another example of a birth process was given on page 19.

Whereas death processes cannot explode (the number of jumps in finite time is at most equal to the initial value), birth processes explode if and only if

$$\sum_{i=1}^{\infty} \frac{1}{\beta_i} < \infty$$

Compare this to the Poisson process and the birth process in the example on page 19.

Index

absorbing state, 22, 27
absorption, 27
accessible, 21

birth-and-death process, 28
    linear, 30
birth intensities, 28
birth process, 31

Chapman-Kolmogorov equations, 7
clock, 12
communicate, 22
competition theorem, 4
continuous-time stochastic process, 2
continuous-time Markov chains, 6

death intensities, 28
death process, 31

eigenvalue, 17
eigenvector, 17
embedded process, 8
ergodic theorem, 27
escape time, 22

holding times, 8
homogeneous Poisson process, 2

immigration, 31
infinitesimal generator, 10
intensity diagram, 11
inter-arrival times, 8
invariant distribution, 23
invariant measure, 23
    regenerative form, 24
irreducibility, 22

jump matrix, 10
jump process, 8

Kolmogorov's backward differential system, 15
Kolmogorov's forward differential system, 15
Kolmogorov's global differential system, 18

matrix exponentials, 17
memoryless property, 4

non-explosive, 19
null recurrent, 23

periodicity, 22
Poisson process, 2
    equivalent definition, 3
    is Markov, 8
    sum, 5
    thinning, 6
positive recurrent, 23

recurrent, 22
regular jump process, 8
return time, 22
Reuter's criterion, 19
    for birth-and-death processes, 30
reversal test, 26
reversible, 26

stationary, 26
stationary distribution, 23
stationary distribution criteria, 25
subordinated chain, 12
sum of Poisson processses, 5

thinning of Poisson processes, 6
time reversed chain, 26
transient, 22
transition semigroup, 7
transition times, 8

uniformisation, 14
uniform Markov chain, 12
    is Markov, 12