SMSTC (2007/08)

Probability

www.smstc.ac.uk

Contents

12 Markov chains in continuous time **12–1**

12.1 Markov property and the Kolmogorov equations 12–2

12.1.1 Finite state space 12–6

12.1.2 Infinite state space 12–7

12.2 Jump chains 12–10

12.3 Classification of states 12–13

12.4 Stationary distribution 12–18

12.5 Ergodic theorem 12–22

12.6 Applications 12–25

12.7 Exercises 12–29

SMSTC (2007/08)

Probability

Lecture 12: Markov chains in continuous time

Xuerong Mao, Strathclyde University[a]

www.smstc.ac.uk

Contents

12.1 Markov property and the Kolmogorov equations **12–2**
12.1.1 Finite state space 12–6
12.1.2 Infinite state space 12–7
12.2 Jump chains **12–10**
12.3 Classification of states **12–13**
12.4 Stationary distribution **12–18**
12.5 Ergodic theorem **12–22**
12.6 Applications **12–25**
12.7 Exercises **12–29**

[a] xuerong@stams.strath.ac.uk

12.1 Markov property and the Kolmogorov equations

Let S be a countable space. A *continuous-time stochastic process* $\{X(t)\}_{t\geq 0}$ with values in S is a family of random variables $X(t) : \Omega \to S$. As in the study of discrete-time Markov chain, we are going to consider ways in which we might specify the probabilistic behaviour (or law) of $\{X(t)\}_{t\geq 0}$. These should enable us to find, at least in principle, any probability connected with the process e.g. $\mathbf{P}(X(t_1) = i_1, \ldots, X(t_n) = i_n)$. There are subtleties in this problem not present in the discrete-time case. They arise because, for a countable disjoint union

$$\mathbf{P}(\coprod_n A_n) = \sum_n \mathbf{P}(A_n),$$

whereas for an uncountable union $\cup_{t\geq 0} A_t$ there is no such rule. To avoid these subtleties as far as possible we shall restrict our attention to processes $X(t) : \Omega \to S$ which are *right-continuous*, namely

$$\lim_{s\downarrow t} X(t,\omega) = X(t,\omega) \quad \text{for all } t \geq 0,\ \omega \in \Omega.$$

In this chapter we are concerned with Markovian stochastic processes with continuous time and countable state space S. We define the *Markov property* of a process $\{X(t)\}_{t\geq 0}$ by

$$\mathbf{P}(X(t) = j \mid X(t_1) = i_1, \ldots, X(t_n) = i_n, X(s) = i) = \mathbf{P}(X(t) = j \mid X(s) = i) \tag{12.1}$$

for any $0 \leq t_1 \leq \ldots \leq t_n \leq s \leq t$ and $i_1, \ldots, i_n, i, j \in S$. Denote

$$p_{ij}(s,t) := \mathbf{P}(X(t) = j \mid X(s) = i).$$

If $p_{ij}(s,t) = p_{ij}(t-s)$ the process has stationary transition probabilities. Unless otherwise stated this property will be assumed henceforth. There is a possibility that a process may reach infinity in finite time. We will assume that if this happens the process stays at infinity for ever (infinity is then called a *coffin state*). This is called the *minimal construction*. The following proposition gives some basic properties of the transition probabilities.

Proposition 12.1.

$$0 \le p_{ij}(t) \le 1. \tag{12.2}$$

$$\sum_{j\in S} p_{ij}(t) \le 1. \tag{12.3}$$

$$p_{ij}(s+t) = \sum_{k\in S} p_{ik}(s)p_{kj}(t). \tag{12.4}$$

$$p_{ij}(0) = 1 \quad (i=j). \tag{12.5}$$

Proof The first and forth statements are trivial, while the second follows from

$$\sum_{j\in S} p_{ij}(t) = \mathbf{P}(X(t)\in S \mid X(0)=i).$$

If the inequality is strict, the process is *dishonest*. To show the third equation, the *Chapman–*

Kolmogorov equation, we compute

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

noticing that the coffin state is ruled out by our construction. □

As in the case of discrete time it is convenient to express things in matrix notation. Let $P(t) = (p_{ij}(t))_{i,j\in S}$. Then (12.4) can be written

$$P(s+t) = P(s)P(t). \tag{12.6}$$

Hence $\{P(t)\}_{t\geq 0}$ is a *semigroup*. It is *stochastic* if there is equality in (12.3), and *sub-stochastic* otherwise.

To proceed we need to assume some regularity. We call the process (or the semigroup) *standard* if the transition probabilities are continuous at 0, i.e. if

$$\lim_{t\downarrow 0} p_{ij}(t) = p_{ij}(0). \tag{12.7}$$

Lemma 12.1. *Let $\{P(t)\}_{t\geq 0}$ be a standard semigroup. Then $p_{ij}(t)$ is a continuous function of t for all i, j.*

Proof We shall show that for any i, j,

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(h), \quad h > 0.$$

From (12.4) we have that

$$p_{ij}(t+h) = \sum_{k\in S} p_{ik}(h)p_{kj}(t)$$

so

$$p_{ij}(t+h) - p_{ij}(t) = (P_{ii}(h)-1)p_{ij}(t) + \sum_{k\neq i} p_{ik}(h)p_{kj}(t).$$

Noting $p_{kj}(t) \leq 1$, we have

$$\sum_{k\neq i} p_{ik}(h)p_{kj}(t) \leq \sum_{k\neq i} p_{ik}(h) = 1 - p_{ii}(h).$$

Hence

$$|p_{ij}(t+h) - p_{ij}(t)| \leq (1-p_{ii}(h))(1-p_{ij}(t)),$$

whence the claim follows. □

The following proposition describes the differentiability of the transition probabilities.

Proposition 12.2. *For a standard stochastic semigroup $\{P(t)\}_{t\geq 0}$ we have*
(i) $\dot{p}_{ii}(0) := dp_{ii}(0)/dt$ exists and is non-positive (but not necessarily finite);
(ii) $\dot{p}_{ij}(0) := dp_{ij}(0)/dt$ exists and is finite for $i \neq j$.

We omit the proof but refer the reader to [2]. Let $Q = (q_{ij}) = (\dot{p}_{ij}(0))$, which is called the *generator* of $\{P(t)\}_{t\geq 0}$ or the Markov chain. The following result is left as Exercise 12-1.

Lemma 12.2. *If S is finite then $\sum_{j\in S} q_{ij} = 0$, while if S is countably infinite we have $\sum_{j\in S} q_{ij} \leq 0$.*

12.1.1 Finite state space

When the state space is finite, using Proposition 12.2 and Lemma 12.2 we have $-q_{ii} = \sum_{j \neq i} q_{ij}$. We will find that it is convenient to set $q_i = -q_{ii}$. For any t write the Taylor expansions

$$p_{ij}(h) = p_{ij}(t, t+h) = q_{ij}h + o(h)$$

and

$$p_{ii}(h) = p_{ii}(t, t+h) = 1 - q_i h + o(h).$$

We hence call q_{ij} the *intensity* of the transition $i \to j$. The following theorem is the key result in the theory of continuous time Markov chains.

Theorem 12.1. *If the state space is finite, then the transition probabilities obey the* Kolmogorov forward equation

$$\frac{dP(t)}{dt} = P(t)Q \tag{12.8}$$

and the Kolmogorov backward equation

$$\frac{dP(t)}{dt} = QP(t). \tag{12.9}$$

Proof By (12.4),

$$p_{ij}(t+h) = \sum_{k \in S} p_{ik}(t)p_{kj}(h) = p_{ij}(t)p_{jj}(h) + \sum_{k \neq j} p_{ik}(t)p_{kj}(h),$$

so

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h} = -p_{ij}(t)\,\frac{1-p_{jj}(h)}{h} + \sum_{k\neq j} p_{ik}(t)\,\frac{p_{kj}(h)}{h}.$$

Letting $h \to 0$ yields

$$\frac{dp_{ij}(t)}{dt} = -p_{ij}(t)q_j + \sum_{k\neq j} p_{ik}(t)q_{kj} = \sum_{k\in S} p_{ik}(t)q_{kj}.$$

This proves the forward equation (12.8). To get the backward equation (12.9) we proceed in a similar fashion, but now writing

$$p_{ij}(t+h) = \sum_{k\in S} p_{ik}(h)p_{kj}(t).$$

□

12.1.2 Infinite state space

Theorem 12.1 is also true for a large class of Markov chains with infinite state space. The necessary assumptions have to do with assuring smoothness of the transition probabilities $p_{ij}(t)$. The semigroup $\{P(t)\}_{t\geq 0}$ is said to be *uniform* if

$$\lim_{t\downarrow 0} p_{ii}(t) = 1 \quad \text{uniformly in } i \in S.$$

The following lemma gives an easy criterion for uniformity (for a proof please see e.g. [2]).

Lemma 12.3. *(i)* $\{P(t)\}_{t\geq 0}$ *is uniform if*

$$\sup_{i\in S} q_i < \infty.$$

(ii) $\{P(t)\}_{t\geq 0}$ *is uniform if and only if*

$$\sum_{j\in S} q_{ij} = 0 \quad i \in S.$$

Theorem 12.2. *If* $\{P(t)\}_{t\geq 0}$ *is uniform then it obeys the Kolmogorov forward equation (12.8) as well as the backward equation (12.9).*

From now on we only consider the Markov chain whose generator Q *obeys* $\sup_{i\in S} q_i < \infty$*. Hence the transition probability matrix* $\{P(t)\}_{t\geq 0}$ *is uniform.*

By Proposition 12.2 and Lemma 12.3, we recall that the generator $Q = (q_{ij})_{i,j\in S}$ of the Markov chain has the following properties:

- $0 \leq -q_{ii} < \infty$ for all i;
- $q_{ij} \geq 0$ for all $i \neq j$;
- $\sum_{j\in S} q_{ij} = 0$ for all i.

Such a matrix is also known as a Q-matrix. In addition to the Q-matrix, we let the initial distribution be $\lambda = (\lambda_j)_{j\in S}$, thought as a row-vector. From now on we will say that $\{X(t)\}_{t\geq 0}$ is Markov (λ, Q). To close this section, let us state the strong Markov property without proof (the proof can be found in e.g. [4]). As defined in the discrete-case, a random variable T with values in $[0, \infty]$ is called a stopping time of $\{X(t)\}_{t\geq 0}$ if for each $t \in [0, \infty]$ the event $\{T \leq t\}$ depends only on $\{X(s) : s \leq t\}$.

Theorem 12.3. *Let* $\{X(t)\}_{t\geq 0}$ *be Markov* (λ, Q) *and let* T *be a stopping time. Then, conditional on* $T < \infty$ *and* $X_T = i$, $\{X(T+t)\}_{t\geq 0}$ *is Markov* (δ_i, Q) *and independent of* $\{X(s) : s \leq t\}$.

12.2 Jump chains

As the state space S is countable, the right-continuity of the Markov chain $\{X(t)\}_{t\geq 0}$ means that for all $\omega \in \Omega$ and $t \geq 0$ there exists $\epsilon > 0$ such that

$$X(s, \omega) = X(t, \omega) \quad \text{for } t \leq s \leq t + \epsilon.$$

Thus, every path of the process must remain constant for a while in each new state. We may have three possibilities for the sorts of path: (i) the path makes finitely many jumps and then becomes stuck in some state forever; (ii) the path makes infinitely many jumps but only finitely many in any interval $[0, t]$; (iii) the path makes infinitely many jumps in a finite interval $[0, \zeta)$. In case (iii), ζ is called the *explosion time* and we set $X(t) = \infty$ for $t \geq \zeta$ as assumed before. Define

$$J_0 = 0, \quad J_{n+1} = \inf\{t \geq J_n : X(t) \neq X(J_n)\}$$

for $n = 0, 1, \ldots$, and

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

for $n = 1, 2, \ldots$. We call $J_0, J_1, \ldots$ the *jump times* and $S_1, S_2, \ldots$ the *holding times*. Note that the right-continuity forces $S_n > 0$ for all n. If $J_{n+1} = \infty$ for some n, we define $X(\infty) = X(J_n)$, the final value, otherwise $X(\infty)$ is undefined. The (first) explosion time ζ is defined by

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

The discrete-time process $\{Y_n\}_{n\geq 0}$ given by $Y_n = X(J_n)$ is called the *jump chain*. This is simply the sequence of values taken by $X(t)$ up to explosion.
Associated to the matrix Q we define a *jump matrix* $\Pi = (\pi_{ij})_{i,j\in S}$ by

$$\pi_{ij} = \begin{cases} q_{ij}/q_i & \text{if } j \neq i \text{ and } q_i \neq 0 \ , \\ 0 & \text{if } j \neq i \text{ and } q_i = 0; \end{cases}$$
$$\pi_{ii} = \begin{cases} 0 & \text{if } q_i \neq 0 \ , \\ 1 & \text{if } q_i = 0 \ . \end{cases}$$

Note that Π is a stochastic matrix.
In the previous section, the Markov chain $\{X(t)\}_{t\geq 0}$ is described in terms of the semigroup $\{P(t)\}_{t\geq 0}$. Another equivalent way to describe how the process evolves is in terms of the jump chain and holding times as described in the following theorem.

Theorem 12.4. *Let $\{X(t)\}_{t\geq 0}$ be a right-continuous Markov chain with generator Q. Let Q be a Q-matrix with jump matrix Π and semigroup $\{P(t)\}_{t\geq 0}$. Then the following statements are equivalent:*

(i) Conditional on $X(0) = i$, the jump chain $\{Y_n\}_{n\geq 0}$ is discrete-time Markov $(\delta_i, \Pi)$and for each $n \geq 1$, conditional on $Y_0 = i_0, \ldots, Y_{n-1} = i_{n-1}$, the holding times $S_1, \ldots, S_n$ are independent exponential random variables of parameters $q_{i_0}, \ldots, q_{i_{n-1}}$ respectively.

(ii) For all $n = 0, 1, 2, \ldots$, all times $0 \leq t_0 \leq \ldots \leq t_n \leq t_{n+1}$ and all states $i_0, i_1, \ldots, i_{n+1}$,

$$\mathbf{P}(X(t_{n+1}) = i_{n+1} \mid X(t_0) = i_0, \ldots, X(t_n) = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n).$$

Again we refer the reader to [4] for the proof. This theorem shows that the process evolves in following way: given that the chain $\{X(t)\}_{t\geq 0}$ starts at i_0 at time $t = 0$, it waits there for an exponential time S_1 of parameter q_{i_0} and then jumps to a new state, choosing state i_1 with probability π_{i_0,i_1}. It then waits for an exponential time S_2 of parameter q_{i_1} and then jumps to a new state, choosing state i_2 with probability π_{i_1,i_2}. It further evolves in the same fashion.

12.3 Classification of states

As in the study of the discrete-time Markov chains, it is useful to identify the class structure of a continuous Markov chain. It is observed from the previous section that the continuous Markov chain can be described in terms of the jump chain. This indicates clearly that the class structure of the continuous Markov chain is simply the discrete-time class structure of the corresponding jump chain. That is, the notions of closed class, irreducible closed class, absorbing state are inherited from the jump chain. Similarly, we say that i reaches j if

$$\mathbf{P}_i(X(t) = j \text{ for some } t \geq 0\ ) > 0.$$

where, as before, we write $\mathbf{P}_i(A)$ for the conditional probability $\mathbf{P}(A|X(0) = i)$. We say that i and j communicate if they reach each other.

Theorem 12.5. *For distinct states i and j the following statements are equivalent:*

(i) *i reaches j;*

(ii) *i reaches j for the jump chain;*

(iii) *$q_{i_0 i_1} q_{i_1 i_2} \cdots q_{i_{n-1} i_n} > 0$ for some states $i_0, i_1, \ldots, i_n$ with $i_0 = i$ and $i_n = j$;*

(iv) *$p_{ij}(t) > 0$ for all $t > 0$;*

(v) *$p_{ij}(t) > 0$ for some $t > 0$.*

Proof Implications (iv) $\Rightarrow$ (v) $\Rightarrow$ (i) $\Rightarrow$ (ii) are clear. If (ii) holds, then by Exercise 10_8, there are some states $i_0, i_1, \ldots, i_n$ with $i_0 = i$ and $i_n = j$ such that $\pi_{i_0 i_1} \pi_{i_1 i_2} \cdots \pi_{i_{n-1} i_n} > 0$, which implies (iii).

We now note that if $q_{uv} > 0$, then

$$p_{uv}(t) \geq \mathbf{P}_u(J_1 \leq t,\ Y_1 = v,\ S_2 > t) = (1 - e^{q_u t})\pi_{uv} e^{-q_v t} > 0$$

for all $t > 0$. Thus, if (iii) holds, then

$$p_{ij}(t) \geq p_{i_0 i_1}(t/n) p_{i_1 i_2}(t/n) \cdots p_{i_{n-1} i_n}(t/n) > 0$$

for all $t > 0$, and (iv) holds. □

We observe that condition of Theorem 12.5 shows that the situation in continuous-time is simpler than in discrete-time, where it may be possible to reach a state, but only after a certain length of time, and then only periodically.

Just as in the discrete-time theory, we say that a state i is *recurrent* if

$$\mathbf{P}_i(\{t \geq 0 : X(t) = i\} \text{ is unbounded}) = 1,$$

while we say i is *transient* if

$$\mathbf{P}_i(\{t \geq 0 : X(t) = i\} \text{ is unbounded}) = 0.$$

The following theorem shows that recurrence and transience are determined by the jump chain again.

Theorem 12.6. *We have:*

(i) *If i is recurrent for the jump chain $\{Y_n\}_{n \geq 0}$, then i is recurrent for $\{X(t)\}_{t \geq 0}$.*

(ii) *If i is transient for the jump chain $\{Y_n\}_{n \geq 0}$, then i is transient for $\{X(t)\}_{t \geq 0}$.*

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

(iv) *Let C be an irreducible closed class. Then either all states in C are transient or all are recurrent.*

Proof (i) Suppose i is recurrent for $\{Y_n\}_{n\geq 0}$. If $X(0) = i$, then $J_n \to \infty$ and $X(J_n) = Y_n = i$ infinitely often with probability 1. We must therefore have

$$\mathbf{P}_i(\{t \geq 0 : X(t) = i\} \text{ is unbounded}) = 1.$$

(ii) Suppose i is transient for $\{Y_n\}_{n\geq 0}$. If $X(0) = i$, then

$$N = \sup\{n \geq 0 : Y_n = i\} < \infty,$$

so $\{t \geq 0 : X(t) = i\}$ is bounded by J_{n+1}, which is finite with probability 1, because $\{Y_n : n \leq N\}$ cannot include an absorbing state.
(iii) Apply Theorem 10.3 to the jump chain.
(iv) Apply Theorem 10.4 to the jump chain. □

The next result gives the conditions for recurrence and transience. We denote by T_i the *first passage time* of $\{X(t)\}_{t\geq 0}$ to state i, defined by

$$T_i = \inf\{t \geq J_1 : X(t) = i\}.$$

Theorem 12.7. *We have:*

(i) *If $q_i = 0$ or $\mathbf{P}_i(T_i < \infty) = 1$, then i is recurrent and $\int_0^\infty p_{ii}(t)dt = \infty$.*

(ii) *If $q_i > 0$ and $\mathbf{P}_i(T_i < \infty) < 1$, then i is transient and $\int_0^\infty p_{ii}(t)dt < \infty$.*

Proof If $q_i = 0$ then $\{X(t)\}_{t\geq 0}$ cannot leave i, so i is recurrent and $p_{ii}(t) = 1$ for all t which yields $\int_0^\infty p_{ii}(t)dt = \infty$. Suppose then $q_i > 0$. Let N_i denote the first passage time of $\{Y_n\}_{n\geq 0}$ to state i. Then

$$\mathbf{P}_i(T_i < \infty) = \mathbf{P}_i(N_i < \infty).$$

By Theorem 12.6 and the corresponding result of the discrete-time Markov chains, we see that i is recurrent (resp. transient) if and only if $\mathbf{P}_i(T_i < \infty) = 1$ (resp. $\mathbf{P}_i(T_i < \infty) < 1$).
Write $\pi_{ij}^{(n)}$ for the (i, j) entry in Π^n. Compute

$$\begin{aligned} \int_0^\infty p_{ii}(t)dt &= \int_0^\infty \mathbf{E}_i(1_{\{X(t)=i\}})dt = \mathbf{E}_i \int_0^\infty 1_{\{X(t)=i\}}dt \\ &= \mathbf{E}_i \sum_{n=0}^\infty S_{n+1}1_{\{Y_n=i\}} = \sum_{n=0}^\infty \mathbf{E}_i(S_{n+1}|Y_n = i)\mathbf{P}_i(Y_n = i) \\ &= \frac{1}{q_i}\sum_0^\infty \pi_{ii}^{(n)}. \end{aligned}$$

The assertions now follow from Theorems 12.6 and 10.3. □

The following result shows that the recurrence and transience are determined by any discrete-time sampling of $\{X(t)\}_{t\geq 0}$.

Theorem 12.8. *Let $\Delta > 0$ be any given step-size and set $Z_n = X(n\Delta)$.*

(i) *If i is recurrent for $\{X(t)\}_{t\geq 0}$, then it is recurrent for $\{Z_n\}_{n\geq 0}$.*

(ii) *If i is transient for $\{X(t)\}_{t\geq 0}$, then it is transient for $\{Z_n\}_{n\geq 0}$.*

Proof Assertion (i) is obvious. To show (ii), we estimate, for $n\Delta \le t < (n+1)\Delta$,

$$\begin{aligned} p_{ii}((n+1)\Delta) &\ge p_{ii}(t)\mathbf{P}(X(t+s) = i \text{ for } s \in [0,\Delta] \mid X(t) = i) \\ &= p_{ii}(t)\mathbf{P}_i(X(s) = i \text{ for } s \in [0,\Delta]) \\ &= p_{ii}(t)\mathbf{P}_i(S_1 > \Delta) \\ &= p_{ii}(t)e^{-q_i\Delta}. \end{aligned}$$

Then

$$\int_0^\infty p_{ii}(t)dt = \sum_{n=0}^\infty \int_{n\Delta}^{(n+1)\Delta} p_{ii}(t)dt \le e^{q_i\Delta} \sum_{n=0}^\infty p_{ii}((n+1)\Delta)$$

and the result follows from Theorems 12.7 and 10.3. □

12.4 Stationary distribution

In this section and next one, we let $\{X(t)\}_{t\geq 0}$ be a Markov chain with the state space S and the generator Q as well as the semigroup $\{P(t)\}_{t\geq 0}$. We always assume that Q is a Q-matrix.

Definition 12.1. A *stationary distribution* of the Markov chain is a probability distribution $\lambda = (\lambda_i)_{i\in S}$ on S, thought of as a row-vector, such that $\lambda Q = 0$. It can also be called an *invariant distribution.*

To see why λ is also called an invariant distribution, we observe that if the state space is finite, the the backward equation shows

$$\frac{d}{dt}\lambda P(t) = \lambda \frac{d}{dt} P(t) = \lambda Q P(t),$$

so $\lambda Q = 0$ implies $\lambda P(t) = \lambda P(0)$ for all t. It is more difficult to show this if the state space is infinite as the interchange of differentiation with the summation involved in multiplication by λ is not justified. In this case, an entirely different proof is needed but not presented here.

Theorem 12.9. *Let Q be a Q-matrix with the jump matrix Π and let λ be a probability distribution on S. The following statements are equivalent:*

(i) λ *is stationary;*

(ii) $\mu\Pi = \mu$, *where* $\mu = (\mu_i)_{i\in S}$ *with* $\mu_i = \lambda_i q_i$.

Proof We have $q_i(\pi_{ij} - \delta_{ij}) = q_{ij}$ for all $i, j \in S$, so

$$(\mu(\Pi - I))_j = \sum_{i\in S} \mu_i(\pi_{ij} - \delta_{ij}) = \sum_{i\in S} \pi_i q_{ij} = (\lambda Q)_j.$$

□

Recall that a state i is recurrent if $q_i = 0$ or $\mathbf{P}_i(T_i < \infty) = 1$. If $q_i = 0$ and the expected return time $m_i = \mathbf{E}_i(T_i) < \infty$, we then say i is positive recurrent. Otherwise a recurrent state i is called null recurrent. It can be showed that i is positive recurrent if and only if it is positive recurrent for the jump chain.

Theorem 12.10. *If the Markov chain $\{X(t)\}_{t\geq 0}$ is irreducible and positive recurrent, then it has a unique stationary distribution.*

Proof We assume the state space S consists at least two states; otherwise the theorem is trivial. Then the irreducibility forces $q_i > 0$ for all i. By Theorems 12.5 and 12.6, the jump chain $\{Y_n\}_{n\geq 0}$ is irreducible and positive recurrent. By Theorem 10.6, the jump chain has a unique stationary distribution $\mu = (\mu_i)_{i\in S}$ which obeys $\mu\Pi = \mu$. By Theorem 12.9, we can take $\lambda_i = \mu_i/q_i$ to obtain the stationary distribution for the Markov chain. The uniqueness follows from the unique stationary distribution for the jump chain and Theorem 12.9. □

We are now concerned with the limiting behavior of $p_{ij}(t)$ at $t \to \infty$ and its relation to stationary distribution. The situation is analogous to the case of discrete-time and is in fact simpler as there is no longer any possibility of periodicity. Let us prepare a lemma.

Lemma 12.4. *Let Q be a Q-matrix with semigroup $\{P(t)\}_{t\geq 0}$. Then for all $t, h \geq 0$,*

$$|p_{ij}(t+h) - p_{ij}(t)| = 1 - e^{-q_i h}.$$

Proof By Lemma 12.1

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(h) \leq \mathbf{P}_i(J_1 \leq h) \leq 1 - e^{-q_i h}.$$

□

Theorem 12.11. *Assume that the Markov chain $\{X(t)\}_{t\geq 0}$ is irreducible and positive recurrent, and its unique stationary distribution is $\lambda = (\lambda_j)_{j\in S}$. Then for all states i, j we have*

$$\lim_{t\to\infty} p_{ij}(t) = \lambda_j$$

Proof Let $\{X(t)\}_{t\geq 0}$ be Markov (δ_i, Q). Fix $h > 0$ and consider the h-skeleton $Z_n = X(nh)$. By the Markov property

$$\mathbf{P}(Z_{n+1} = i_{n+1} | Z_0 = i_0, \ldots, Z_n = i_n) = p_{i_n i_{n+1}}(h)$$

so $\{Z_n\}_{n\geq 0}$ is discrete-time Markov $(\delta_i, P(h))$. By Theorem 12.3, the irreducibility implies $p_{ij}(h) > 0$ for all i, j so $\{Z_n\}_{n\geq 0}$ is irreducible and aperiodic. By Exercise 12-3, λ is a stationary distribution for $\{Z_n\}_{n\geq 0}$. So, by Theorem 10.7,

$$\lim_{n\to\infty} p_{ij}(nh) = \lambda_j.$$

Now, fix state i. For any $\epsilon > 0$ we can find $h > 0$ so that

$$1 - e^{-q_i s} \leq \epsilon/2 \quad \text{for } 0 \leq s \leq h,$$

and then find N such that

$$|p_{ij}(nh) - \lambda_i| \leq \epsilon/2 \quad \text{for } n \geq N.$$

Hence for $t \geq Nh$, we have $nh \leq t < (n+1)h$ for some $n \geq N$ and

$$|p_{ij}(t) - \lambda_i| \leq |p_{ij}(t) - p_{ij}(nh)| + |p_{ij}(nh) - \lambda_i| \leq \epsilon$$

by Lemma 12.4. This completes the proof. □

To close this section, let us introduce the concept of detailed balance for the continuous-time chains. A Q-matrix Q and a probability distribution λ are said to be in *detailed balance* if

$$\lambda_i q_{ij} = \lambda_j q_{ji} \quad \text{for all } i, j.$$

It is easy to see that if Q and λ are in detailed balance, then λ is a stationary distribution. We leave the proof as an exercise.

12.5 Ergodic theorem

The following ergodic theorem gives the long-run proportion of time spent by a continuous-time chain in each state as in the discrete-time case.

Theorem 12.12. *Let $\{X(t)\}_{t\geq 0}$ be Markov (ρ, Q), where ρ is any probability distribution on S. If it is irreducible, then*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t 1_{\{X(s)=i\}} ds = \frac{1}{m_i q_i} \quad a.s. \tag{12.10}$$

where $m_i = \mathbf{E}_i(T_i)$ is the expected return time to state i. Moreover, if the chain is positive recurrent, then for any bounded function $f : S \to \mathbb{R}$, we have

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(X(s))ds = \bar{f} := \sum_{i\in S} \lambda_i f(i) \quad a.s \tag{12.11}$$

where $(\lambda_i)_{i\in S}$ is the unique stationary distribution, namely $\lambda_i = 1/(m_i q_i)$.

Proof If the chain is transient then the total time spent in any state i is finte, so

$$0 \leq \frac{1}{t}\int_0^t 1_{\{X(s)=i\}} ds \leq \frac{1}{t}\int_0^\infty 1_{\{X(s)=i\}} ds \to 0 = \frac{1}{m_i q_i}$$

as $m_i = \infty$. Suppose then that the chain is recurrent. Fix any state i. Then $X(t)$ hits i with probability 1 and the long-run proportion of time in i equals the long-run proportion of time in i after first hitting i. So, by the stonge Markov property (of the jump chain), it is sufficient to consider the case when $\rho = \delta_i$. Set $T_i^0 = 0$ and define, for $n = 0, 1, 2, \ldots$,

$$\begin{aligned} M_i^{n+1} &= \inf\{t > T_i^n : X(t) \neq i\} - T_i^n, \\ T_i^{n+1} &= \inf\{t > T_i^n + M_i^{n+1} : X(t) = i\}, \\ L_i^{n+1} &= T_i^{n+1} - T_i^n. \end{aligned}$$

By the strong Markov property (of the jump chain) at the stopping times T_i^n for $n \geq 0$ we see that $L_i^1, L_i^2, \ldots$ are independent and identically distributed with mean m_i, and that $M_i^1, M_i^2, \ldots$ are independent and identically distributed with mean $1/q_i$. Hence, by the strong law of large numbers,

$$\lim_{n\to\infty} \frac{L_i^1 + \cdots + L_i^n}{n} = m_i \quad a.s.$$

and

$$\lim_{n\to\infty} \frac{M_i^1 + \cdots + M_i^n}{n} = \frac{1}{q_i} \quad a.s.$$

Hence

$$\lim_{n\to\infty} \frac{M_i^1 + \cdots + M_i^n}{L_i^1 + \cdots + L_i^n} = \frac{1}{m_i q_i} \quad a.s.$$

Moreover, we note that $T_i^n = L_i^1 + \cdots + L_i^n$ and $T_i^n / T_i^{n+1} \to 1$ as $n \to \infty$ with probability 1. Now, for $T_i^n \leq t < T_i^{n+1}$, we have

$$\frac{T_i^n}{T_i^{n+1}} \frac{M_i^1 + \cdots + M_i^n}{L_i^1 + \cdots + L_i^n} \leq \frac{1}{t} \int_0^t 1_{\{X(s)=i\}} ds \leq \frac{T_i^{n+1}}{T_i^n} \frac{M_i^1 + \cdots + M_i^{n+1}}{L_i^1 + \cdots + L_i^{n+1}}.$$

Letting $t \to \infty$ we obtain the assertion (12.10). In the positive recurrent case, by writting

$$\frac{1}{t}\int_0^t f(X(s))ds - \bar{f} = \sum_{i\in S}\Big(\frac{1}{t}\int_0^t 1_{\{X(s)=1\}}ds - \lambda_i\Big)f(i) \tag{12.12}$$

we can show another assertion (12.11) in the same way as Theorem 10.8 was proved. We leave the details as an exercise. □

12.6 Applications

Example 12.1. (General birth process) A general birth process is a Markov chain on the state space $S = \{0, 1, 2, \ldots\}$ with the generator

$$Q = \begin{pmatrix} -q_0 & q_0 & 0 & 0 & \cdots \\ 0 & -q_1 & q_1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

If $\sup_{i\geq 0} q_i < \infty$, then the semigroup $\{P(t)\}_{t\geq 0}$ is uniform. If $\sum_{i\geq 0} q_i^{-1} < \infty$, then $q_i \to \infty$ so $\sup_{i\geq 0} q_i = \infty$ and $\{P(t)\}_{t\geq 0}$ is not uniform.
Let us now consider a birth process which has constant intensity of births, namely $q_i = q$ for all $i \in S$. Clearly, $\{P(t)\}_{t\geq 0}$ is uniform (provided that $q < \infty$). The forward equation yields

$$\frac{dp_{jk}(t)}{dt} = -qp_{jk}(t) + qp_{j,k-1}(t).$$

In particular, if $j = 0$, we have

$$\frac{dp_{0k}(t)}{dt} = -qp_{0k}(t) + qp_{0,k-1}(t). \tag{12.13}$$

The initial conditions are assumed to be $p_{00}(0) = 1$ and $p_{0i}(0) = 0$ for $i \geq 1$, so that the process starts in state 0. In order to solve this differential equation, we will attempt to convert it into a partial differential equation for the probability generating function

$$G(s;t) = \mathbf{E}s^{X(t)} = \sum_{i\geq 0} p_{0i}(t)s^i.$$

Multiplying both sides of (12.13) by s^k and summing over k we get

$$\frac{\partial G(s;t)}{\partial t} = -qG(s;t) + qsG(s,t).$$

For a fixed value of s we see that

$$\frac{\partial G(s;t)}{\partial t} = -q(1-s)G(s;t)$$

so that

$$G(s;t) = A(s)e^{-q(1-s)t}.$$

From the initial condition $G(s;0) = 1$, we must have $A(s) = 1$, so $X(t)$ follows a Poisson distribution with mean qt. In general

$$p_{ij}(t) = p_{ij}(s, s+t) = e^{-qt}\frac{(qt)^{j-i}}{(j-i)!}, \quad j \geq i.$$

We observe that the process we just derived is the *Poisson process* discussed in the previous Chapter.

Example 12.2. (Birth and death process) A birth and death process is a Markov chain on the state space $S = \{0, 1, 2, \dots\}$ with the generator

$$Q = \begin{pmatrix} -b_0 & b_0 & 0 & 0 & \cdots \\ d_1 & -(b_1+d_1) & b_1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

This process was introduced by McKendrick in 1925, used by Feller in 1939 to describe biological population growth, and studied in detail by Kendall later. When it is used to describe biological population, the stationary distribution is particularly interested. To determine the stationary distribution $\lambda = (\lambda_j)_{j\in S}$, the equation $\lambda Q = 0$ yields

$$\begin{aligned} -b_0\lambda_0 + d_1\lambda_1 &= 0, \\ b_{j-1}\lambda_{j-1} - (b_j + d_j)\lambda_j + d_{j+1}\lambda_{j+1} &= 0, \quad j \geq 1, \end{aligned}$$

with solution

$$\lambda_j = \frac{b_{j-1}}{d_j}\lambda_{j-1} = \frac{b_{j-1}\cdots b_0}{d_j\cdots d_1}\lambda_0, \quad j \geq 1.$$

But $\sum_{j\geq 0}\lambda_j = 1$, whence

$$\Big(1 + \sum_{j=1}^{\infty}\frac{b_{j-1}\cdots b_0}{d_j\cdots d_1}\Big)\lambda_0 = 1.$$

Naturally we require

$$\sum_{j=1}^{\infty}\frac{b_{j-1}\cdots b_0}{d_j\cdots d_1} < \infty$$

to give

$$\lambda_0 = \Big(1 + \sum_{j=1}^{\infty}\frac{b_{j-1}\cdots b_0}{d_j\cdots d_1}\Big)^{-1}$$

and

$$\lambda_j = \frac{b_{j-1}\cdots b_0}{d_j\cdots d_1}\left(1+\sum_{j=1}^{\infty}\frac{b_{j-1}\cdots b_0}{d_j\cdots d_1}\right)^{-1}, \quad j \geq 1.$$

As a special case, let $b_j = b$ and $d_j = d$ and assume $r := b/d < 1$. Then

$$\lambda_0 = 1 - r \quad \text{and} \quad \lambda_j = (1-r)r^j, \quad j \geq 1,$$

which is a geometric distribution.

12.7 Exercises

12–1. Prove Lemma 12.2.

12–2. Show that the general birth process defined in Example 12.1 is dishonest if $\sum_{i\geq 1} q_i^{-1} < \infty$.

12–3. Assume that the Markov chain $\{X(t)\}_{t\geq 0}$ is irreducible and positive recurrent, and its unique stationary distribution is $\lambda = (\lambda)_{j\in S}$. Show

$$\lambda P(t) = \lambda \quad \forall t \geq 0.$$

12–4. If a Q-matrix Q and a probability distribution λ are in detailed balance, show that λ is a stationary distribution.

12–5. Show assertion (12.11) from (12.12).

References

[1] K.L. Chung, *Markov Chains with Stationary Transition Probabilities*, Springer, 2nd edition, 1967.

[2] D. Freedman, *Markov Chains*, Spinger, 1983.

[3] P. Guttorp, *Stochastic Modelling of Scientific Data*, Chapman and Hall, 1995.

[4] J.R. Norris, *Markov Chains*, C.U.P., 1997.

[5] D. Revus, *Markov Chains*, North-Holland, 1984.

[6] D. Williams, *Proability with Martingales*, C.U.P., 1991.