

16.18: Stationary and Limiting Distributions of Continuous-Time Chains

In this section, we study the limiting behavior of continuous-time Markov chains by focusing on two interrelated ideas: invariant (or stationary) distributions and limiting distributions. In some ways, the limiting behavior of continuous-time chains is simpler than the limiting behavior of discrete-time chains, in part because the complications caused by periodicity in the discrete-time case do not occur in the continuous-time case. Nonetheless as we will see, the limiting behavior of a continuous-time chain is closely related to the limiting behavior of the embedded, discrete-time jump chain.

Review

Once again, our starting point is a time-homogeneous, continuous-time Markov chain $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with discrete state space (S, \mathcal{S}) . By definition, this means that S is countable with the discrete topology, so that \mathcal{S} is the σ -algebra of all subsets of S .

Let's review what we have so far. We assume that the Markov chain \mathbf{X} is *regular*. Among other things, this means that the basic structure of \mathbf{X} is determined by the *transition times* $\tau = (\tau_0, \tau_1, \tau_2, \dots)$ and the *jump chain* $\mathbf{Y} = (Y_0, Y_1, Y_2, \dots)$. First, $\tau_0 = 0$ and $\tau_1 = \tau = \inf\{t > 0 : X_t \neq X_0\}$. The time-homogeneous and Markov properties imply that the distribution of τ given $X_0 = x$ is exponential with parameter $\lambda(x) \in [0, \infty)$. Part of regularity is that \mathbf{X} is right continuous so that there are no *instantaneous states* where $\lambda(x) = \infty$, which would mean $\mathbb{P}(\tau = 0 | X_0 = x) = 1$. On the other hand, $\lambda(x) \in (0, \infty)$ means that x is a *stable state* so that τ has a proper exponential distribution given $X_0 = x$, with $\mathbb{P}(0 < \tau < \infty | X_0 = x) = 1$. Finally, $\lambda(x) = 0$ means that x is an *absorbing state* so that $\mathbb{P}(\tau = \infty | X_0 = x) = 1$. The remaining transition times are defined recursively: $\tau_{n+1} = \inf\{t > \tau_n : X_t \neq X_{\tau_n}\}$ if $\tau_n < \infty$ and $\tau_{n+1} = \infty$ if $\tau_n = \infty$. Another component of regularity is that with probability 1, $\tau_n \rightarrow \infty$ as $n \rightarrow \infty$, ruling out the *explosion event* of infinitely many jumps in finite time. The jump chain \mathbf{Y} is formed by sampling \mathbf{X} at the transition times (until the chain is sucked into an absorbing state, if that happens). That is, with $M = \sup\{n : \tau_n < \infty\}$ and for $n \in \mathbb{N}$, we define $Y_n = X_{\tau_n}$ if $n \leq M$ and $Y_n = X_{\tau_M}$ if $n > M$. Then \mathbf{Y} is a discrete-time Markov chain with one-step transition matrix Q given $Q(x, y) = \mathbb{P}(X_{\tau} = y | X_0 = x)$ if $(x, y) \in S^2$ with x stable and $Q(x, x) = 1$ if $x \in S$ is absorbing.

The transition matrix P_t at time $t \in [0, \infty)$ is given by $P_t(x, y) = \mathbb{P}(X_t = y | X_0 = x)$ for $(x, y) \in S^2$. The time-homogeneous and Markov properties imply that the collection of transition matrices $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ satisfies the *Chapman-Kolmogorov equations* $P_s P_t = P_{s+t}$ for $s, t \in [0, \infty)$, and hence is a *semigroup* of transition matrices. The transition semigroup \mathbf{P} and the initial distribution of X_0 determine all of the finite-dimensional distributions of \mathbf{X} . Since there are no instantaneous states, P is *standard* which means that $P_t \rightarrow I$ as $t \downarrow 0$ (as matrices, and so pointwise). The fundamental relationship between \mathbf{P} on the one hand, and λ and Q on the other, is

$$P_t(x, y) = I(x, y)e^{-\lambda(x)t} + \int_0^t \lambda(x)e^{-\lambda(x)s} Q P_{t-s}(x, y) ds, \quad (x, y) \in S^2 \quad (16.18.1)$$

From this, it follows that the matrix function $t \mapsto P_t$ is differentiable (again, pointwise) and satisfies the *Kolmogorov backward equation* $\frac{d}{dt} P_t = G P_t$, where the *infinitesimal generator matrix* G is given by $G(x, y) = -\lambda(x)I(x, y) + \lambda(x)Q(x, y)$ for $(x, y) \in S^2$. If we impose the stronger assumption that \mathbf{P} is *uniform*, which means that $P_t \rightarrow I$ as $t \downarrow 0$ as operators on \mathcal{B} (so with respect to the supremum norm), then the backward equation as well as the companion *Kolmogorov forward equation* $\frac{d}{dt} P_t = P_t G$ hold as operators on \mathcal{B} . In addition, we have the matrix exponential representation $P_t = e^{tG}$ for $t \in [0, \infty)$. The uniform assumption is equivalent to the exponential parameter function being bounded.

Finally, for $\alpha \in [0, \infty)$, the α potential matrix U_α of \mathbf{X} is $U_\alpha = \int_0^\infty e^{-\alpha t} P_t dt$. The *resolvent* $\mathbf{U} = \{U_\alpha : \alpha \in (0, \infty)\}$ is the Laplace transform of \mathbf{P} and hence gives the same information as \mathbf{P} . From this point of view, the time-homogeneous and Markov properties lead to the *resolvent equation* $U_\alpha = U_\beta + (\beta - \alpha)U_\alpha U_\beta$ for $\alpha, \beta \in (0, \infty)$ with $\alpha \leq \beta$. For $\alpha \in (0, \infty)$, the α potential matrix is related to the generator by the fundamental equation $\alpha U_\alpha = I + G U_\alpha$. If \mathbf{P} is uniform, then this equation, as well as the companion $\alpha U_\alpha = I + U_\alpha G$ hold as operators on \mathcal{B} , which leads to $U_\alpha = (\alpha I - G)^{-1}$.

Basic Theory

Relations and Classification

We start our discussion with relations among states and classifications of states. These are the same ones that we studied for discrete-time chains in our study of recurrence and transience, applied here to the jump chain \mathbf{Y} . But as we will see, the relations and classifications make sense for the continuous-time chain \mathbf{X} as well. The discussion is complicated slightly when there are absorbing states. Only when \mathbf{X} is in an absorbing state can we not interpret the values of \mathbf{Y} as the values of \mathbf{X} at the transition times (because of course, there are no transitions when \mathbf{X} is in an absorbing state). But $x \in S$ is absorbing for the continuous-time chain \mathbf{X} if and only if x is absorbing for the jump chain \mathbf{Y} , so this trivial exception is easily handled.

For $y \in S$ let $\rho_y = \inf\{n \in \mathbb{N}_+ : Y_n = y\}$, the (discrete) hitting time to y for the jump chain \mathbf{Y} , where as usual, $\inf(\emptyset) = \infty$. That is, ρ_y is the first positive (discrete) time that \mathbf{Y} is in state y . The analogous random time for the continuous-time chain \mathbf{X} is τ_{ρ_y} , where naturally we take $\tau_\infty = \infty$. This is the first time that \mathbf{X} is in state y , not counting the possible initial period in y . Specifically, suppose $X_0 = x$. If $x \neq y$ then $\tau_{\rho_y} = \inf\{t > 0 : X_t = y\}$. If $x = y$ then $\tau_{\rho_y} = \inf\{t > \tau_1 : X_t = y\}$.

Define the *hitting matrix* H by

$$H(x, y) = \mathbb{P}(\rho_y < \infty \mid Y_0 = x), \quad (x, y) \in S^2 \quad (16.18.2)$$

Then $H(x, y) = \mathbb{P}(\tau_{\rho_y} < \infty \mid X_0 = x)$ except when x is absorbing and $y = x$.

So for the continuous-time chain, if $x \in S$ is stable then $H(x, x)$ is the probability that, starting in x , the chain \mathbf{X} returns to x after its initial period in x . If $x, y \in S$ are distinct, then $H(x, y)$ is simply the probability that \mathbf{X} , starting in x , eventually reaches y . It follows that the basic relation among states makes sense for either the continuous-time chain \mathbf{X} as well as its jump chain \mathbf{Y} .

Define the relation \rightarrow on S^2 by $x \rightarrow y$ if $x = y$ or $H(x, y) > 0$.

The *leads to* relation \rightarrow is *reflexive* by definition: $x \rightarrow x$ for every $x \in S$. From our previous study of discrete-time chains, we know it's also *transitive*: if $x \rightarrow y$ and $y \rightarrow z$ then $x \rightarrow z$ for $x, y, z \in S$. We also know that $x \rightarrow y$ if and only if there is a directed path in the state graph from x to y , if and only if $Q^n(x, y) > 0$ for some $n \in \mathbb{N}$. For the continuous-time transition matrices, we have a stronger result that in turn makes a stronger case that the *leads to* relation is fundamental for \mathbf{X} .

Suppose $(x, y) \in S^2$.

1. If $x \rightarrow y$ then $P_t(x, y) > 0$ for all $t \in (0, \infty)$.
2. If $x \nrightarrow y$ then $P_t(x, y) = 0$ for all $t \in (0, \infty)$.

Proof

This result is proved in the section on transition matrices and generators.

This result is known as the *Lévy dichotomy*, and is named for Paul Lévy. Let's recall a couple of other definitions:

Suppose that A is a nonempty subset of S .

1. A is *closed* if $x \in A$ and $x \rightarrow y$ imply $y \in A$.
2. A is *irreducible* if A is closed and has no proper closed subset.

If S is irreducible, we also say that the chain \mathbf{X} itself is irreducible.

If A is a nonempty subset of S , then $\text{cl}(A) = \{y \in S : x \rightarrow y \text{ for some } x \in A\}$ is the smallest closed set containing A , and is called the *closure* of A .

Suppose that $A \subseteq S$ is closed. Then

1. P_t^A , the restriction of P_t to $A \times A$, is a transition probability matrix on A for every $t \in [0, \infty)$.
2. \mathbf{X} restricted to A is a continuous-time Markov chain with transition semigroup $\mathbf{P}^A = \{P_t^A : t \in [0, \infty)\}$.

Proof

1. If $x \in A$ and $y \notin A$, then x does not lead to y so in particular $P_t(x, y) = 0$. It follows that $\sum_{y \in A} P_t(x, y) = 1$ for $x \in A$ so P_t^A is a transition probability matrix.
2. This follows from (a). If the chain starts in A , then the chain remains in A for all time, and of course, the Markov property still holds.

Define the relation \leftrightarrow on S by $x \leftrightarrow y$ if $x \rightarrow y$ and $y \rightarrow x$ for $(x, y) \in S^2$.

The *to and from* relation \leftrightarrow defines an equivalence relation on S and hence partitions S into mutually disjoint equivalence classes. Recall from our study of discrete-time chains that a closed set is not necessarily an equivalence class, nor is an equivalence class necessarily closed. However, an irreducible set is an equivalence class, but an equivalence class may not be irreducible. The importance of the relation \leftrightarrow stems from the fact that many important properties of Markov chains (in discrete or continuous time) turn out to be *class properties*, shared by all states in an equivalence class. The following definition is fundamental, and once again, makes sense for either the continuous-time chain \mathbf{X} or its jump chain \mathbf{Y} .

Let $x \in S$.

1. State x is *transient* if $H(x, x) < 1$
2. State x is *recurrent* if $H(x, x) = 1$.

Recall from our study of discrete-time chains that if x is recurrent and $x \rightarrow y$ then y is recurrent and $y \rightarrow x$. Thus, recurrence and transience are class properties, shared by all states in an equivalence class.

Time Spent in a State

For $x \in S$, let N_x denote the number of visits to state x by the jump chain \mathbf{Y} , and let T_x denote the total time spent in state x by the continuous-time chain \mathbf{X} . Thus

$$N_x = \sum_{n=0}^{\infty} \mathbf{1}(Y_n = x), \quad T_x = \int_0^{\infty} \mathbf{1}(X_t = x) dt \quad (16.18.3)$$

The expected values $R(x, y) = \mathbb{E}(N_y | Y_0 = x)$ and $U(x, y) = \mathbb{E}(T_y | X_0 = x)$ for $(x, y) \in S^2$ define the *potential matrices* of \mathbf{Y} and \mathbf{X} , respectively. From our previous study of discrete-time chains, we know the distribution and mean of N_y given $Y_0 = x$ in terms of the hitting matrix H . The next two results give a review:

Suppose that $x, y \in S$ are distinct. Then

1. $\mathbb{P}(N_y = n | Y_0 = y) = H^{n-1}(y, y)[1 - H(y, y)]$ for $n \in \mathbb{N}_+$
2. $\mathbb{P}(N_y = 0 | Y_0 = x) = 1 - H(x, y)$ and $\mathbb{P}(N_y = n | Y_0 = x) = H(x, y)H^{n-1}(y, y)[1 - H(y, y)]$ for $n \in \mathbb{N}_+$

Let's take cases. First suppose that y is recurrent. In part (a), $\mathbb{P}(N_y = n | Y_0 = y) = 0$ for all $n \in \mathbb{N}_+$, and consequently $\mathbb{P}(N_y = \infty | Y_0 = y) = 1$. In part (b), $\mathbb{P}(N_y = n | Y_0 = x) = 0$ for $n \in \mathbb{N}_+$, and consequently $\mathbb{P}(N_y = 0 | Y_0 = x) = 1 - H(x, y)$ while $\mathbb{P}(N_y = \infty | Y_0 = x) = H(x, y)$. Suppose next that y is transient. Part (a) specifies a proper geometric distribution on \mathbb{N}_+ while in part (b), probability $1 - H(x, y)$ is assigned to 0 and the remaining probability $H(x, y)$ is geometrically distributed over \mathbb{N}_+ as in (a). In both cases, N_y is finite with probability 1. Next we consider the expected value, that is, the (discrete) potential. To state the results succinctly we will use the convention that $a/0 = \infty$ if $a > 0$ and $0/0 = 0$.

Suppose again that $x, y \in S$ are distinct. Then

1. $R(y, y) = 1/[1 - H(y, y)]$
2. $R(x, y) = H(x, y)/[1 - H(y, y)]$

Let's take cases again. If $y \in S$ is recurrent then $R(y, y) = \infty$, and for $x \in S$ with $x \neq y$, either $R(x, y) = \infty$ if $x \rightarrow y$ or $R(x, y) = 0$ if $x \nrightarrow y$. If $y \in S$ is transient, $R(y, y)$ is finite, as is $R(x, y)$ for every $x \in S$ with $x \neq y$. Moreover, there is an inverse relationship of sorts between the potential and the hitting probabilities.

Naturally, our next goal is to find analogous results for the continuous-time chain \mathbf{X} . For the distribution of T_y it's best to use the right distribution function.

Suppose that $x, y \in S$ are distinct. Then for $t \in [0, \infty)$

1. $\mathbb{P}(T_y > t \mid X_0 = y) = \exp\{-\lambda(y)[1 - H(y, y)]t\}$
2. $\mathbb{P}(T_y > t \mid X_0 = x) = H(x, y) \exp\{-\lambda(y)[1 - H(y, y)]t\}$

Proof

The proof is by conditioning on N_y .

1. First, if $H(y, y) = 1$ (so that y is recurrent), then either y is absorbing with $\mathbb{P}(\tau_1 = \infty \mid X_0 = y) = 1$ or y is stable and recurrent, so that $\mathbb{P}(N_y = \infty \mid X_0 = y) = 1$. In the second case, starting in state y , T_y is the sum of infinitely many independent variables, each with the exponential distribution with parameter $\lambda(y) \in (0, \infty)$. In both cases, $\mathbb{P}(T_y = \infty \mid X_0 = y) = 1$ and so $\mathbb{P}(T_y > t \mid X_0 = y) = 1$ for every $t \in [0, \infty)$. So suppose that $H(y, y) < 1$ so that y is transient. Then

$$\mathbb{P}(T_y > t \mid X_0 = y) = \sum_{n=1}^{\infty} \mathbb{P}(T_y > t \mid X_0 = y, N_y = n) \mathbb{P}(N_y = n \mid X_0 = y) \quad (16.18.4)$$

Given $N_y = n$, T_y is the sum of n independent variables, each having the exponential distribution with parameter $\lambda(y)$. So T_y has the gamma distribution with parameters n and $\lambda(y)$ and hence

$$\mathbb{P}(T_y > t \mid X_0 = y, N_y = n) = \sum_{k=0}^{n-1} e^{-\lambda(y)t} \frac{[\lambda(y)t]^k}{k!} \quad (16.18.5)$$

From the previous result, $\mathbb{P}(N_y = n \mid X_0 = y) = \mathbb{P}(N_y = n \mid Y_0 = y) = H^{n-1}(y, y)[1 - H(y, y)]$. We substitute, change the order of summation, use geometric series and then exponential series:

$$\begin{aligned} \mathbb{P}(T_y > t \mid X_0 = y) &= \sum_{n=1}^{\infty} \left(\sum_{k=0}^{n-1} e^{-\lambda(y)t} \frac{[\lambda(y)t]^k}{k!} \right) H^{n-1}(y, y)[1 - H(y, y)] \\ &= e^{-\lambda(y)t} [1 - H(y, y)] \sum_{k=0}^{\infty} \frac{[\lambda(y)t]^k}{k!} \sum_{n=k+1}^{\infty} H^{n-1}(y, y) \\ &= e^{-\lambda(y)t} \sum_{k=0}^{\infty} \frac{[\lambda(y)t]^k}{k!} H^k(y, y) = e^{-\lambda(y)t} \exp[\lambda(y)H(y, y)t] \end{aligned}$$

Simplifying gives the result.

2. The proof is similar. If $H(y, y) = 1$ so that y is recurrent, then starting in state x , either $T_y = 0$ if $N_y = 0$, which occurs with probability $1 - H(x, y)$ or $T_y = \infty$ if $N_y = \infty$, which occurs with probability $H(x, y)$. If $H(y, y) < 1$ so that y is transient, then the result follows from conditioning on N_y as in (a), except that $\mathbb{P}(T_y = 0 \mid X_0 = x) = \mathbb{P}(N_y = 0 \mid Y_0 = x) = 1 - H(x, y)$.

Let's take cases as before. Suppose first that y is recurrent. In part (a), $\mathbb{P}(T_y > t \mid X_0 = y) = 1$ for every $t \in [0, \infty)$ and hence $\mathbb{P}(T_y = \infty \mid X_0 = y) = 1$. In part (b), $\mathbb{P}(T_y > t \mid X_0 = x) = H(x, y)$ for every $t \in [0, \infty)$ and consequently $\mathbb{P}(T_y = 0 \mid X_0 = x) = 1 - H(x, y)$ while $\mathbb{P}(T_y = \infty \mid X_0 = x) = H(x, y)$. Suppose next that y is transient. From part (a), the distribution of T_y given $X_0 = y$ is exponential with parameter $\lambda(y)[1 - H(y, y)]$. In part (b), the distribution assigns probability $1 - H(x, y)$ to 0 while the remaining probability $H(x, y)$ is exponentially distributed over $(0, \infty)$ as in (a). Taking expected value, we get a very nice relationship between the potential matrix U of the continuous-time chain \mathbf{X} and the potential matrix R of the discrete-time jump chain \mathbf{Y} :

For every $(x, y) \in S^2$,

$$U(x, y) = \frac{R(x, y)}{\lambda(y)} \quad (16.18.6)$$

Proof

If y is recurrent, then $U(x, y) = R(x, y)$ and the common value is either 0 if $H(x, y) = 0$ or ∞ if $H(x, y) = 1$. So suppose that y is transient. We can compute the expected value of T_y by integrating the right distribution function in the previous theorem. In case $x = y$, we have

$$U(y, y) = \int_0^\infty \exp\{-\lambda(y)[1 - H(y, y)]t\} dt = \frac{1}{\lambda(y)[1 - H(y, y)]} = \frac{R(y, y)}{\lambda(y)} \quad (16.18.7)$$

In the case that x and y are distinct,

$$U(x, y) = \int_0^\infty H(x, y) \exp\{-\lambda(y)[1 - H(y, y)]t\} dt = \frac{H(x, y)}{\lambda(y)[1 - H(y, y)]} = \frac{R(x, y)}{\lambda(y)} \quad (16.18.8)$$

In particular, $y \in S$ is transient if and only if $R(x, y) < \infty$ for every $x \in S$, if and only if $U(x, y) < \infty$ for every $x \in S$. On the other hand, y is recurrent if and only if $R(x, y) = U(x, y) = \infty$ if $x \rightarrow y$ and $R(x, y) = U(x, y) = 0$ if $x \nrightarrow y$.

Null and Positive Recurrence

Unlike transience and recurrence, the definitions of null and positive recurrence of a state $x \in S$ are different for the continuous-time chain \mathbf{X} and its jump chain \mathbf{Y} . This is because these definitions depend on the *expected* hitting time to x , starting in x , and not just the *finiteness* of this hitting time. For $x \in S$, let $\nu(x) = \mathbb{E}(\rho_x \mid Y_0 = x)$, the expected (discrete) return time to x starting in x . Recall that x is *positive recurrent* for \mathbf{Y} if $\nu(x) < \infty$ and x is *null recurrent* if x is recurrent but not positive recurrent, so that $H(x, x) = 1$ but $\nu(x) = \infty$. The definitions are similar for \mathbf{X} , but using the continuous hitting time τ_{ρ_x} .

For $x \in S$, let $\mu(x) = 0$ if x is absorbing and $\mu(x) = \mathbb{E}(\tau_{\rho_x} \mid X_0 = x)$ if x is stable. So if x is stable, $\mu(x)$ is the expected return time to x starting in x (after the initial period in x).

1. State x is *positive recurrent* for \mathbf{X} if $\mu(x) < \infty$.
2. State x is *null recurrent* for \mathbf{X} if x recurrent but not positive recurrent, so that $H(x, x) = 1$ but $\mu(x) = \infty$.

A state $x \in S$ can be positive recurrent for \mathbf{X} but null recurrent for its jump chain \mathbf{Y} or can be null recurrent for \mathbf{X} but positive recurrent for \mathbf{Y} . But like transience and recurrence, positive and null recurrence are *class properties*, shared by all states in an equivalence class under the *to and from* equivalence relation \leftrightarrow .

Invariant Functions

Our next discussion concerns functions that are invariant for the transition matrix Q of the jump chain \mathbf{Y} and functions that are invariant for the transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ of the continuous-time chain \mathbf{X} . For both discrete-time and continuous-time chains, there is a close relationship between invariant functions and the limiting behavior in time.

First let's recall the definitions. A function $f : S \rightarrow [0, \infty)$ is *invariant* for Q (or for the chain \mathbf{Y}) if $fQ = f$. It then follows that $fQ^n = f$ for every $n \in \mathbb{N}$. In continuous time we must *assume* invariance at each time. That is, a function $f : S \rightarrow [0, \infty)$ is invariant for \mathbf{P} (or for the chain \mathbf{X}) if $fP_t = f$ for all $t \in [0, \infty)$. Our interest is in nonnegative functions, because we can think of such a function as the density function, with respect to counting measure, of a positive measure on S . We are particularly interested in the special case that f is a *probability density function*, so that $\sum_{x \in S} f(x) = 1$. If Y_0 has a probability density function f that is invariant for Q , then Y_n has probability density function f for all $n \in \mathbb{N}$ and hence \mathbf{Y} is *stationary*. Similarly, if X_0 has a probability density function f that is invariant for \mathbf{P} then X_t has probability density function f for every $t \in [0, \infty)$ and once again, the chain \mathbf{X} is *stationary*.

Our first result shows that there is a one-to-one correspondence between invariant functions for Q and zero functions for the generator G .

Suppose $f : S \rightarrow [0, \infty)$. Then $fG = 0$ if and only if $(\lambda f)Q = \lambda f$, so that λf is invariant for Q .

Proof

This is a simple consequence of the definition of the generator:

$$fG(y) = \sum_{x \in S} f(x)G(x, y) = -\lambda(y)f(y) + \sum_{x \in S} f(x)\lambda(x)Q(x, y), \quad y \in S \quad (16.18.9)$$

or in functional form, $fG = -\lambda f + (\lambda f)Q$

If our chain \mathbf{X} has no absorbing states, then $f : S \rightarrow [0, \infty)$ is invariant for Q if and only if $(f/\lambda)G = 0$.

Suppose that $f : S \rightarrow [0, \infty)$. Then f is invariant for \mathbf{P} if and only if $fG = 0$.

Proof 1

Assume that λ is bounded, so that the transition semigroup \mathbf{P} is uniform. Then $P_t = e^{tG}$ for $t \in [0, \infty)$. So if $f : S \rightarrow [0, \infty)$ then

$$fP_t = f(e^{tG}) = f \sum_{n=0}^{\infty} \frac{t^n}{n!} G^n = f + \sum_{n=1}^{\infty} \frac{t^n}{n!} fG^n \quad (16.18.10)$$

Since f is nonnegative, $fP_t = f$ if and only if $fG = 0$ (in which case $fG^n = 0$ for every $n \in \mathbb{N}_+$).

Proof 2

Suppose that $fP_t = f$ for $t \in [0, \infty)$. Then $\frac{d}{dt}(fP_t) = 0$ for $t \in [0, \infty)$. But using the Kolmogorov backward equation, $\frac{d}{dt}(fP_t) = f \frac{d}{dt}P_t = fGP_t = 0$. Letting $t = 0$ we conclude that $fG = 0$. Conversely, if $fG = 0$ then $\frac{d}{dt}(fP_t) = f \frac{d}{dt}P_t = fGP_t = 0$ for $t \in [0, \infty)$. It follows that fP_t is constant in $t \in [0, \infty)$. Since $fP_0 = f$ it follows that $fP_t = f$ for all $t \in [0, \infty)$.

So putting the two main results together we see that f is invariant for the continuous-time chain \mathbf{X} if and only if λf is invariant for the jump chain \mathbf{Y} . Our next result shows how functions that are invariant for \mathbf{X} are related to the resolvent $\mathbf{U} = \{U_\alpha : \alpha \in (0, \infty)\}$. To appreciate the result, recall that for $\alpha \in (0, \infty)$ the matrix αU_α is a *probability* matrix, and in fact $\alpha U_\alpha(x, \cdot)$ is the conditional probability density function of X_T , given $X_0 = x$, where T is independent of \mathbf{X} and has the exponential distribution with parameter α . So αU_α is a transition matrix just as P_t is a transition matrix, but corresponding to the exponentially distributed random time T with parameter $\alpha \in (0, \infty)$ rather than the deterministic time $t \in [0, \infty)$.

Suppose that $f : S \rightarrow [0, \infty)$. If $fG = 0$ then $f(\alpha U_\alpha) = f$ for $\alpha \in (0, \infty)$. Conversely if $f(\alpha U_\alpha) = f$ for $\alpha \in (0, \infty)$ then $fG = 0$.

Proof

Recall that $I + GU_\alpha = \alpha U_\alpha$ for $\alpha \in (0, \infty)$. Hence if $fG = 0$ then

$$f(\alpha U_\alpha) = f + fGU_\alpha = f \quad (16.18.11)$$

Conversely, suppose that $f(\alpha U_\alpha) = f$. Then

$$fGU_\alpha = \int_0^\infty e^{-\alpha t} fGP_t dt = 0 \quad (16.18.12)$$

As a function of $\alpha \in (0, \infty)$, the integral on the right side is the Laplace transform of the time function $t \mapsto fGP_t$. Hence we must have $fGP_t = 0$ for $t \in (0, \infty)$, and letting $t \downarrow 0$ gives $fG = 0$.

So extending our summary, $f : S \rightarrow [0, \infty)$ is invariant for the transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ if and only if λf is invariant for jump transition matrices $\{Q^n : n \in \mathbb{N}\}$ if and only if $fG = 0$ if and only if f is invariant for the collection of probability matrices $\{\alpha U_\alpha : \alpha \in (0, \infty)\}$. From our knowledge of the theory for discrete-time chains, we now have the following fundamental result:

Suppose that \mathbf{X} is irreducible and recurrent.

1. There exists $g : S \rightarrow (0, \infty)$ that is invariant for \mathbf{X} .
2. If f is invariant for \mathbf{X} , then $f = cg$ for some constant $c \in [0, \infty)$.

Proof

The result is trivial if S consists of a single, necessarily absorbing, state. Otherwise, there are no absorbing states, since \mathbf{X} is irreducible and so $\lambda(x) > 0$ for $x \in S$. From the result above, f is invariant for \mathbf{X} if and only if λf is invariant for \mathbf{Y} . But \mathbf{Y} is also irreducible and recurrent, so we know that there exists a strictly positive function that is invariant for \mathbf{Y} , and every other function that is invariant for \mathbf{Y} is a nonnegative multiple of this one. Hence the same is true for \mathbf{X} .

Invariant functions have a nice interpretation in terms of *occupation times*, an interpretation that parallels the discrete case. The potential gives the expected total time in a state, starting in another state, but here we need to consider the expected time in a state during a *cycle* that starts and ends in another state.

For $x \in S$, define the function γ_x by

$$\gamma_x(y) = \mathbb{E} \left(\int_0^{\tau_{P_x}} \mathbf{1}(X_s = y) ds \mid X_0 = x \right), \quad y \in S \quad (16.18.13)$$

so that $\gamma_x(y)$ is the expected *occupation time* in state y before the first return to x , starting in x .

Suppose again that \mathbf{X} is irreducible and recurrent. For $x \in S$,

1. $\gamma_x : S \rightarrow (0, \infty)$
2. γ_x is invariant for \mathbf{X}
3. $\gamma_x(x) = 1/\lambda(x)$
4. $\mu(x) = \sum_{y \in S} \gamma_x(y)$

Proof

As is often the case, the proof is based on results that we already have for the embedded jump chain. For $x \in S$, define

$$\delta_x(y) = \mathbb{E} \left(\sum_{n=0}^{\rho_x-1} \mathbf{1}(Y_n = y) \mid Y_0 = x \right), \quad y \in S \quad (16.18.14)$$

so that $\delta_x(y)$ is the expected number of visits to y before the first return to x , starting in x , for the jump chain $\mathbf{Y} = (Y_0, Y_1, \dots)$. Since \mathbf{X} is irreducible and recurrent, so is \mathbf{Y} . From our results in the discrete case we know that

1. $\delta_x : S \rightarrow (0, \infty)$
2. δ_x is invariant for \mathbf{Y}
3. $\delta_x(x) = 1$

From our results above, it follows that the function $y \mapsto \delta_x(y)/\lambda(y)$ satisfies properties (a), (b), and (c) in the theorem. But each visit to y by the jump chain \mathbf{Y} has expected length $1/\lambda(y)$ for the continuous-time chain \mathbf{X} . It follows that $\gamma_x(y) = \delta_x(y)/\lambda(y)$ for $x, y \in S$. By definition, $\gamma_x(y)$ is the expected occupation time in y before the first return to x , starting in x . Hence, summing over $y \in S$ gives the expected return time to x , starting in x , so (d) holds.

So now we have some additional insight into positive and null recurrence for the continuous-time chain \mathbf{X} and the associated jump chain \mathbf{Y} . Suppose again that the chains are irreducible and recurrent. There exist $g : S \rightarrow (0, \infty)$ that is invariant for \mathbf{Y} , and then g/λ is invariant for \mathbf{X} . The invariant functions are unique up to multiplication by positive constants. The jump chain \mathbf{Y} is positive recurrent if and only if $\sum_{x \in S} g(x) < \infty$ while the continuous-time chain \mathbf{X} is positive recurrent if and only if $\sum_{x \in S} g(x)/\lambda(x) < \infty$. Note that if λ is bounded (which is equivalent to the transition semigroup \mathbf{P} being uniform), then \mathbf{X} is positive recurrent if and only if \mathbf{Y} is positive recurrent.

Suppose again that \mathbf{X} is irreducible and recurrent.

1. If \mathbf{X} is null recurrent then \mathbf{X} does not have an invariant probability density function.
2. If \mathbf{X} is positive recurrent then \mathbf{X} has a unique, positive invariant probability density function.

Proof

From the previous result, there exists $g : S \rightarrow (0, \infty)$ that is invariant for \mathbf{X} , and every other invariant function is a nonnegative multiple of this one. The function f given by

$$f(y) = \frac{g(y)}{\sum_{x \in S} g(x)}, \quad y \in S \quad (16.18.15)$$

is uniquely defined (that is, unchanged if we replace g by cg where $c > 0$).

1. If $\sum_{x \in S} g(x) = \infty$ then $f(y) = 0$ for every $y \in S$.
2. If $\sum_{x \in S} g(x) < \infty$ then $f(y) > 0$ for every $y \in S$ and $\sum_{y \in S} f(y) = 1$.

Limiting Behavior

Our next discussion focuses on the limiting behavior of the transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$. Our first result is a simple corollary of the result above for potentials.

If $y \in S$ is transient, then $P_t(x, y) \rightarrow 0$ as $t \rightarrow \infty$ for every $x \in S$.

Proof

This follows from the previous result. If $y \in S$ is transient, then for any $x \in S$,

$$U(x, y) = \int_0^\infty P_t(x, y) dt < \infty \quad (16.18.16)$$

and so we must have $P_t(x, y) \rightarrow 0$ as $t \rightarrow \infty$.

So we should turn our attention to the recurrent states. The set of recurrent states partitions into equivalent classes under \leftrightarrow , and each of these classes is irreducible. Hence we can assume without loss of generality that our continuous-time chain $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is irreducible and recurrent. To avoid trivialities, we will also assume that S has at least two states. Thus, there are no absorbing states and so $\lambda(x) > 0$ for $x \in S$. Here is the main result.

Suppose that $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is irreducible and recurrent. Then $f(y) = \lim_{t \rightarrow \infty} P_t(x, y)$ exists for each $y \in S$, independently of $x \in S$. The function f is invariant for \mathbf{X} and

$$f(y) = \frac{\gamma_x(y)}{\mu(x)}, \quad y \in S \quad (16.18.17)$$

1. If \mathbf{X} is null recurrent then $f(y) = 0$ for all $y \in S$.
2. If \mathbf{X} is positive recurrent then $f(y) > 0$ for all $y \in S$ and $\sum_{y \in S} f(y) = 1$.

Proof sketch

The basic idea is that

$$\lim_{t \rightarrow \infty} P_t(x, y) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t P_s(x, y) ds \quad (16.18.18)$$

The expression on the right is the limiting proportion of time spent in $y \in S$, starting in $x \in S$. This proportion is $\gamma_x(y)/\mu(x)$, so the results follow from the theorem [above](#).

The limiting function f can be computed in a number of ways. First we find a function $g : S \rightarrow (0, \infty)$ that is invariant for \mathbf{X} . We can do this by solving

- $gP_t = g$ for $t \in (0, \infty)$
- $gG = 0$
- $g(\alpha U_\alpha) = g$ for $\alpha \in (0, \infty)$
- $hQ = h$ and then $g = h/\lambda$

The function g is unique up to multiplication by positive constants. If $\sum_{x \in S} g(x) < \infty$, then we are in the positive recurrent case and so f is simply g normalized:

$$f(y) = \frac{g(y)}{\sum_{x \in S} g(x)}, \quad y \in S \quad (16.18.19)$$

The following result is known as the *ergodic theorem* for continuous-time Markov chains. It can also be thought of as a strong law of large numbers for continuous-time Markov chains.

Suppose that $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is irreducible and positive recurrent, with (unique) invariant probability density function f . If $h : S \rightarrow \mathbb{R}$ then

$$\frac{1}{t} \int_0^t h(X_s) ds \rightarrow \sum_{x \in S} f(x)h(x) \text{ as } t \rightarrow \infty \quad (16.18.20)$$

with probability 1, assuming that the sum on the right converges absolutely.

Notes

First, let $x, y \in S$ and let $h = \mathbf{1}_y$, the indicator function of y . Then given $X_0 = x$, $\frac{1}{t} \int_0^t h(X_s) ds$ is the average occupation time in state y , starting in state x , over the time interval $[0, t]$. In expected value, this is $\frac{1}{t} \int_0^t P_s(x, y) ds$ which we know converges to $f(y)$ as $t \rightarrow \infty$, independently of x . So in this special case, the ergodic theorem states that the convergence is with probability 1 also. A general function $h : S \rightarrow \mathbb{R}$ is a linear combination of the indicator functions of the points in S , so the ergodic theorem is plausible.

Note that no assumptions are made about X_0 , so the limit is independent of the initial state. By now, this should come as no surprise. After a long period of time, the Markov chain \mathbf{X} “forgets” about the initial state. Note also that $\sum_{x \in S} f(x)h(x)$ is the expected value of h , thought of as a random variable on S with probability measure defined by f . On the other hand, $\frac{1}{t} \int_0^t h(X_s) ds$ is the average of the time function $s \mapsto h(X_s)$ on the interval $[0, t]$. So the ergodic theorem states that the limiting time average on the left is the same as the spatial average on the right.

Applications and Exercises

The Two-State Chain

The continuous-time, two-state chain has been studied in the last several sections. The following result puts the pieces together and completes the picture.

Consider the continuous-time Markov chain $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ on $S = \{0, 1\}$ with transition rate $a \in (0, \infty)$ from 0 to 1 and transition rate $b \in (0, \infty)$ from 1 to 0. Give each of the following

1. The transition matrix Q^n for \mathbf{Y} at $n \in \mathbb{N}$.
2. The infinitesimal generator G .
3. The transition matrix P_t for \mathbf{X} at $t \in [0, \infty)$.
4. The invariant probability density function for \mathbf{Y} .
5. The invariant probability density function for \mathbf{X} .
6. The limiting behavior of Q^n as $n \rightarrow \infty$.
7. The limiting behavior of P_t as $t \rightarrow \infty$.

Answer

Note that since the transition rates a and b are positive, the chain is irreducible.

1. First, $Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and then for $n \in \mathbb{N}$, $Q^n = Q$ if n is odd and $Q^n = I$ if n is even.
2. $G = \begin{bmatrix} -a & a \\ b & -b \end{bmatrix}$.
3. $P_t = \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix} - \frac{1}{a+b} e^{-(a+b)t} \begin{bmatrix} -a & a \\ b & -b \end{bmatrix}$ for $t \in [0, \infty)$.
4. $f_d = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$
5. $f_c = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}$
6. As in (a), $Q^{2n} = I$ and $Q^{2n+1} = Q$ for $n \in \mathbb{N}$. So there are two sub-sequential limits. The jump chain \mathbf{Y} is periodic with period 2.
7. $P_t \rightarrow \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix}$ as $t \rightarrow \infty$. Each row is f_c .

Computational Exercises

The following continuous-time chain has also been studied in the previous three sections.

Consider the Markov chain $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ on $S = \{0, 1, 2\}$ with exponential parameter function $\lambda = (4, 1, 3)$ and jump transition matrix

$$Q = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 \end{bmatrix} \quad (16.18.21)$$

1. Recall the generator matrix G .
2. Find the invariant probability density function f_d for \mathbf{Y} by solving $f_d Q = f_d$.
3. Find the invariant probability density function f_c for \mathbf{X} by solving $f_c G = 0$.
4. Verify that λf_c is a multiple of f_d .
5. Describe the limiting behavior of Q^n as $n \rightarrow \infty$.
6. Describe the limiting behavior of P_t as $t \rightarrow \infty$.
7. Verify the result in (f) by recalling the transition matrix P_t for \mathbf{X} at $t \in [0, \infty)$.

Answer

1. $G = \begin{bmatrix} -4 & 2 & 2 \\ 1 & -1 & 0 \\ 1 & 2 & -3 \end{bmatrix}$
2. $f_d = \frac{1}{14} [6 \ 5 \ 3]$
3. $f_c = \frac{1}{15} [3 \ 10 \ 2]$
4. $\lambda f_c = \frac{1}{15} [12 \ 10 \ 6] = \frac{28}{15} f_d$
5. $Q^n \rightarrow \frac{1}{14} \begin{bmatrix} 6 & 5 & 3 \\ 6 & 5 & 3 \\ 6 & 5 & 3 \end{bmatrix}$ as $n \rightarrow \infty$
6. $P_t \rightarrow \frac{1}{15} \begin{bmatrix} 3 & 10 & 2 \\ 3 & 10 & 2 \\ 3 & 10 & 2 \end{bmatrix}$ as $t \rightarrow \infty$
7. $P_t = \frac{1}{15} \begin{bmatrix} 3 + 12e^{-5t} & 10 - 10e^{-3t} & 2 - 12e^{-5t} + 10e^{-3t} \\ 3 - 3e^{-5t} & 10 + 5e^{-3t} & 2 + 3e^{-5t} - 5e^{-3t} \\ 3 - 3e^{-5t} & 10 - 10e^{-3t} & 2 + 3e^{-5t} + 10e^{-3t} \end{bmatrix}$ for $t \in [0, \infty)$

Special Models

Read the discussion of stationary and limiting distributions for chains subordinate to the Poisson process.

Read the discussion of stationary and limiting distributions for continuous-time birth-death chains.

Read the discussion of classification and limiting distributions for continuous-time queuing chains.

This page titled [16.18: Stationary and Limiting Distributions of Continuous-Time Chains](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist \(Random Services\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.