

16.16: Transition Matrices and Generators of Continuous-Time Chains

16. Transition Matrices and Generators of Continuous-Time Chains

Preliminaries

This is the second of the three introductory sections on continuous-time Markov chains. Thus, suppose that $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is a continuous-time Markov chain defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with state space (S, \mathcal{S}) . By the very meaning of *Markov chain*, the set of states S is countable and the σ -algebra \mathcal{S} is the collection of all subsets of S . So every subset of S is measurable, as is every function from S to another measurable space. Recall that \mathcal{S} is also the Borel σ algebra corresponding to the discrete topology on S . With this topology, every function from S to another topological space is continuous. Counting measure $\#$ is the natural measure on (S, \mathcal{S}) , so in the context of the general introduction, integrals over S are simply sums. Also, kernels on S can be thought of as matrices, with rows and sums indexed by S . The left and right kernel operations are generalizations of matrix multiplication.

A space of functions on S plays an important role. Let \mathcal{B} denote the collection of bounded functions $f : S \rightarrow \mathbb{R}$. With the usual pointwise definitions of addition and scalar multiplication, \mathcal{B} is a vector space. The *supremum norm* on \mathcal{B} is given by

$$\|f\| = \sup\{|f(x)| : x \in S\}, \quad f \in \mathcal{B} \quad (16.16.1)$$

Of course, if S is finite, \mathcal{B} is the set of all real-valued functions on S , and $\|f\| = \max\{|f(x)| : x \in S\}$ for $f \in \mathcal{B}$.

In the last section, we studied \mathbf{X} in terms of *when* and *how* the state changes. To review briefly, let $\tau = \inf\{t \in (0, \infty) : X_t \neq X_0\}$. Assuming that \mathbf{X} is right continuous, the Markov property of \mathbf{X} implies the memoryless property of τ , and hence the distribution of τ given $X_0 = x$ is exponential with parameter $\lambda(x) \in [0, \infty)$ for each $x \in S$. The assumption of right continuity rules out the pathological possibility that $\lambda(x) = \infty$, which would mean that x is an *instantaneous state* so that $\mathbb{P}(\tau = 0 | X_0 = x) = 1$. On the other hand, if $\lambda(x) \in (0, \infty)$ then 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, if $\lambda(x) = 0$ then x is an *absorbing state*, so that $\mathbb{P}(\tau = \infty | X_0 = x) = 1$. Next we define a sequence of stopping times: First $\tau_0 = 0$ and $\tau_1 = \tau$. Recursively, if $\tau_n < \infty$ then $\tau_n = \inf\{t > \tau_n : X_t \neq X_{\tau_n}\}$, while if $\tau_n = \infty$ then $\tau_{n+1} = \infty$. With $M = \sup\{n \in \mathbb{N} : \tau_n < \infty\}$ we define $Y_n = X_{\tau_n}$ if $n \in \mathbb{N}$ with $n \leq M$ and $Y_n = Y_M$ if $n \in \mathbb{N}$ with $n > M$. The sequence $\mathbf{Y} = (Y_0, Y_1, \dots)$ is a discrete-time Markov chain on S with one-step transition matrix Q given by $Q(x, y) = \mathbb{P}(X_\tau = y | X_0 = x)$ if $x, y \in S$ with x stable, and $Q(x, x) = 1$ if $x \in S$ is absorbing. Assuming that \mathbf{X} is *regular*, which means that $\tau_n \rightarrow \infty$ as $n \rightarrow \infty$ with probability 1 (ruling out the *explosion* event of infinitely many transitions in finite time), the structure of \mathbf{X} is completely determined by the sequence of stopping times $\boldsymbol{\tau} = (\tau_0, \tau_1, \dots)$ and the discrete-time jump chain $\mathbf{Y} = (Y_0, Y_1, \dots)$. Analytically, the distribution \mathbf{X} is determined by the exponential parameter function λ and the one-step transition matrix Q of the jump chain.

In this section, we will study the Markov chain \mathbf{X} in terms of the transition matrices in continuous time and a fundamentally important matrix known as the *generator*. Naturally, the connections between the two points of view are particularly interesting.

The Transition Semigroup

Definition and basic Properties

The first part of our discussion is very similar to the treatment for a general Markov processes, except for simplifications caused by the discrete state space. We assume that $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is a Markov chain on S .

The *transition probability matrix* P_t of \mathbf{X} corresponding to $t \in [0, \infty)$ is

$$P_t(x, y) = \mathbb{P}(X_t = y | X_0 = x), \quad (x, y) \in S^2 \quad (16.16.2)$$

In particular, $P_0 = I$, the identity matrix on S

Proof

The mapping $y \mapsto P_t(x, y)$ is the PDF of X_t given $X_0 = x$. Hence P_t is a probability matrix. That is, $P_t(x, y) \geq 0$ for $(x, y) \in S^2$ and $\sum_{y \in S} P_t(x, y) = 1$ for $x \in S$. Trivially, $P_0 = I$ by definition.

Note that since we are assuming that the Markov chain is homogeneous,

$$P_t(x, y) = \mathbb{P}(X_{s+t} = y | X_s = x), \quad (x, y) \in S^2 \quad (16.16.3)$$

for every $s, t \in [0, \infty)$. The *Chapman-Kolmogorov equation* given next is essentially yet another restatement of the Markov property. The equation is named for Andrei Kolmogorov and Sydney Chapman,

Suppose that $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ is the collection of transition matrices for the chain \mathbf{X} . Then $P_s P_t = P_{s+t}$ for $s, t \in [0, \infty)$. Explicitly,

$$P_{s+t}(x, z) = \sum_{y \in S} P_s(x, y) P_t(y, z), \quad x, z \in S \quad (16.16.4)$$

Proof

We condition on X_s .

$$P_{s+t}(x, z) = \mathbb{P}(X_{s+t} = z \mid X_0 = x) = \sum_{y \in S} \mathbb{P}(X_{s+t} = z \mid X_s = y, X_0 = x) \mathbb{P}(X_s = y \mid X_0 = x) \quad (16.16.5)$$

But by the Markov and time homogeneous properties,

$$\mathbb{P}(X_{s+t} = z \mid X_s = y, X_0 = x) = \mathbb{P}(X_{s+t} = z \mid X_s = y) = P_t(y, z) \quad (16.16.6)$$

Of course by definition, $\mathbb{P}(X_s = y \mid X_0 = x) = P_s(x, y)$. So the first displayed equation above becomes

$$P_{s+t}(x, z) = \sum_{y \in S} P_s(x, y) P_t(y, z) = P_s P_t(x, z) \quad (16.16.7)$$

Restated in another form of jargon, the collection $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ is a *semigroup* of probability matrices. The semigroup of transition matrices \mathbf{P} , along with the initial distribution, determine the finite-dimensional distributions of \mathbf{X} .

Suppose that X_0 has probability density function f . If $(t_1, t_2, \dots, t_n) \in [0, \infty)^n$ is a time sequence with $0 < t_1 < \dots < t_n$ and $(x_0, x_1, \dots, x_n) \in S^{n+1}$ is a state sequence, then

$$\mathbb{P}(X_0 = x_0, X_{t_1} = x_1, \dots, X_{t_n} = x_n) = f(x_0) P_{t_1}(x_0, x_1) P_{t_2 - t_1}(x_1, x_2) \cdots P_{t_n - t_{n-1}}(x_{n-1}, x_n) \quad (16.16.8)$$

Proof

To simplify the notation, we will just give the cases $n = 1$ and $n = 2$, which capture the essence of the proof. First suppose $x, y \in S$ and $t \in [0, \infty)$. Then

$$\mathbb{P}(X_0 = x, X_t = y) = \mathbb{P}(X_0 = x) \mathbb{P}(X_t = y \mid X_0 = x) = f(x) P_t(x, y) \quad (16.16.9)$$

Next suppose that $x, y, z \in S$ and $s, t \in [0, \infty)$ with $s < t$. Then

$$\mathbb{P}(X_0 = x, X_s = y, X_t = z) = \mathbb{P}(X_t = z \mid X_0 = x, X_s = y) \mathbb{P}(X_0 = x, X_s = y) \quad (16.16.10)$$

But by the Markov and time homogeneous properties, $\mathbb{P}(X_t = z \mid X_0 = x, X_s = y) = P_{t-s}(y, z)$. By the $n = 1$ case, $\mathbb{P}(X_0 = x, X_s = y) = f(x) P_s(x, y)$. Hence

$$\mathbb{P}(X_0 = x, X_s = y, X_t = z) = f(x) P_s(x, y) P_{t-s}(y, z) \quad (16.16.11)$$

As with any matrix on S , the transition matrices define left and right operations on functions which are generalizations of matrix multiplication. For a transition matrix, both have natural interpretations.

Suppose that $f : S \rightarrow \mathbb{R}$, and that either f is nonnegative or $f \in \mathcal{B}$. Then for $t \in [0, \infty)$,

$$P_t f(x) = \sum_{y \in S} P_t(x, y) f(y) = \mathbb{E}[f(X_t) \mid X_0 = x], \quad x \in S \quad (16.16.12)$$

The mapping $f \mapsto P_t f$ is a bounded, linear operator on \mathcal{B} and $\|P_t\| = 1$.

Proof

Since $P_t(x, \cdot)$ is the conditional probability density function of X_t given $X_0 = x$, it follows that $P_t f(x) = \mathbb{E}[f(X_t) \mid X_0 = x]$. The statement about $f \mapsto P_t f$ follows from general results on probability kernels.

If f is nonnegative and S is infinite, then it's possible that $P_t f(x) = \infty$. In general, the left operation of a positive kernel acts on positive measures on the state space. In the setting here, if μ is a positive (Borel) measure on (S, \mathcal{S}) , then the function $f : S \rightarrow [0, \infty)$ given by

$f(x) = \mu\{x\}$ for $x \in S$ is the density function of μ with respect to counting measure $\#$ on (S, \mathcal{S}) . This simply means that $\mu(A) = \sum_{x \in A} f(x)$ for $A \subseteq S$. Conversely, given $f : S \rightarrow [0, \infty)$, the set function $\mu(A) = \sum_{x \in A} f(x)$ for $A \subseteq S$ defines a positive measure on (S, \mathcal{S}) with f as its density function. So for the left operation of P_t , it's natural to consider only nonnegative functions.

If $f : S \rightarrow [0, \infty)$ then

$$fP_t(y) = \sum_{x \in S} f(x)P_t(x, y), \quad y \in S \quad (16.16.13)$$

If X_0 has probability density function f then X_t has probability density function fP_t .

Proof

If X_0 has PDF f , then conditioning gives

$$\mathbb{P}(X_t = y) = \sum_{x \in S} \mathbb{P}(X_t = y \mid X_0 = x) \mathbb{P}(X_0 = x) = \sum_{x \in S} P_t(x, y) f(x) = fP_t(y), \quad y \in S \quad (16.16.14)$$

More generally, if f is the density function of a positive measure μ on (S, \mathcal{S}) then fP_t is the density function of the measure μP_t , defined by

$$\mu P_t(A) = \sum_{x \in S} \mu\{x\} P_t(x, A) = \sum_{x \in S} f(x) P_t(x, A), \quad A \subseteq S \quad (16.16.15)$$

A function $f : S \rightarrow [0, \infty)$ is *invariant* for the Markov chain \mathbf{X} (or for the transition semigroup \mathbf{P}) if $fP_t = f$ for every $t \in [0, \infty)$.

It follows that if X_0 has an invariant probability density function f , then X_t has probability density function f for every $t \in [0, \infty)$, so \mathbf{X} is identically distributed. Invariant and limiting distributions are fundamentally important for continuous-time Markov chains.

Standard Semigroups

Suppose again that $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is a Markov chain on S with transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$. Once again, continuity assumptions need to be imposed on \mathbf{X} in order to rule out strange behavior that would otherwise greatly complicate the theory. In terms of the transition semigroup \mathbf{P} , here is the basic assumption:

The transition semigroup \mathbf{P} is *standard* if $P_t(x, x) \rightarrow 1$ as $t \downarrow 0$ for each $x \in S$.

Since $P_0(x, x) = 1$ for $x \in S$, the standard assumption is clearly a continuity assumption. It actually implies much stronger smoothness properties that we will build up by stages.

If the transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ is standard, then the function $t \mapsto P_t(x, y)$ is right continuous for each $(x, y) \in S^2$.

Proof

First note that if $(x, y) \in S^2$ with $x \neq y$ then $P_h(x, y) \leq 1 - P_h(x, x) \rightarrow 0$ as $h \downarrow 0$. Hence $P_h(x, y) \rightarrow I(x, y)$ as $h \downarrow 0$ for all $(x, y) \in S^2$. Suppose next that $t \in (0, \infty)$ and $(x, y) \in S^2$. By the semigroup property,

$$P_{t+h}(x, y) = P_t P_h(x, y) = \sum_{z \in S} P_t(x, z) P_h(z, y) \quad (16.16.16)$$

But $P_h(z, y) \rightarrow I(z, y)$ as $h \downarrow 0$ so by the bounded convergence theorem, $P_{t+h}(x, y) \rightarrow P_t(x, y)$ as $h \downarrow 0$.

Our next result connects one of the basic assumptions in the section on transition times and the embedded chain with the standard assumption here.

If the Markov chain \mathbf{X} has no instantaneous states then the transition semigroup \mathbf{P} is standard.

Proof

Given $X_0 = x \in S$ note that $\tau > t$ implies $X_t = x$. Hence

$$P_t(x, x) = \mathbb{P}(X_t = x \mid X_0 = x) \geq \mathbb{P}(\tau > t \mid X_0 = x) = e^{-\lambda(x)t} \quad (16.16.17)$$

Since \mathbf{X} has no instantaneous states, $0 \leq \lambda(x) < \infty$ so $e^{-\lambda(x)t} \rightarrow 1$ as $t \downarrow 0$.

Recall that the non-existence of instantaneous states is essentially equivalent to the right continuity of \mathbf{X} . So we have the nice result that if \mathbf{X} is right continuous, then so is \mathbf{P} . For the remainder of our discussion, we assume that $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is a regular Markov chain on S with transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$, exponential function λ and one-step transition matrix Q for the jump chain. Our next result is the fundamental integral equations relating \mathbf{P} , λ , and Q .

For $t \in [0, \infty)$,

$$P_t(x, y) = I(x, y)e^{-\lambda(x)t} + \int_0^t \lambda(x)e^{-\lambda(x)s} Q P_{t-s}(x, y) ds, \quad (x, y) \in S^2 \quad (16.16.18)$$

Proof

If x is an absorbing state, then the equation trivially holds, since $\lambda(x) = 0$ and $P_t(x, y) = I(x, y)$. So suppose that x is a stable state, and as above, let $\tau = \inf\{t \in [0, \infty) : X_t \neq X_0\}$. Given $X_0 = x$, τ has a proper exponential distribution with parameter $\lambda(x) \in (0, \infty)$. Taking cases,

$$P_t(x, y) = \mathbb{P}(X_t = y \mid X_0 = x) = \mathbb{P}(X_t = y, \tau > t \mid X_0 = x) + \mathbb{P}(X_t = y, \tau \leq t \mid X_0 = x) \quad (16.16.19)$$

The first term on the right is 0 if $y \neq x$ and is $\mathbb{P}(\tau > t \mid X_0 = x) = e^{-\lambda(x)t}$ if $y = x$. In short,

$$\mathbb{P}(X_t = y, \tau > t \mid X_0 = x) = I(x, y)e^{-\lambda(x)t} \quad (16.16.20)$$

For the second term on the right in the displayed equation, we condition on τ and $Y_1 = X_\tau$. By a result in the last section on transition times and the embedded chain, the joint PDF of (τ, Y_1) at $s \in [0, \infty)$ and $z \in S$, given $X_0 = x$, is $\lambda(x)e^{-\lambda(x)s} Q(x, z)$ (continuous in time, discrete in space). Also, given $\tau = s \in [0, t]$ and $Y_1 = z \in S$, we can use the strong Markov property to “restart the clock” at s giving

$$\mathbb{P}(X_t = y \mid X_0 = x, \tau = s, Y_1 = z) = \mathbb{P}(X_{t-s} = y \mid X_0 = z) = P_{t-s}(z, y) \quad (16.16.21)$$

Putting the pieces together we have

$$\mathbb{P}(X_t = y, \tau \leq t \mid X_0 = x) = \int_0^t \lambda(x)e^{-\lambda(x)s} \sum_{z \in S} Q(x, z) P_{t-s}(z, y) ds = \int_0^t \lambda(x)e^{-\lambda(x)s} Q P_{t-s}(x, y) ds \quad (16.16.22)$$

We can now improve on the continuity result that we got earlier. First recall the leads to relation for the jump chain \mathbf{Y} : For $(x, y) \in S^2$, x leads to y if $Q^n(x, y) > 0$ for some $n \in \mathbb{N}$. So by definition, x leads to x for each $x \in S$, and for $(x, y) \in S^2$ with $x \neq y$, x leads to y if and only if the discrete-time chain starting in x eventually reaches y with positive probability.

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

1. $t \mapsto P_t(x, y)$ is continuous.
2. If x leads to y then $P_t(x, y) > 0$ for every $t \in (0, \infty)$.
3. If x does not lead to y then $P_t(x, y) = 0$ for every $t \in (0, \infty)$.

Proof

For $t \in [0, \infty)$, we can use the change of variables $r = t - s$ in the fundamental integral equation to get

$$P_t(x, y) = I(x, y)e^{-\lambda(x)t} + \lambda(x)e^{-\lambda(x)t} \int_0^t e^{\lambda(x)r} Q P_r(x, y) dr, \quad (x, y) \in S^2 \quad (16.16.23)$$

1. In the displayed equation, $r \mapsto P_r(x, y)$ is right continuous for every $(x, y) \in S^2$, and hence by the bounded convergence theorem again, so is $r \mapsto Q P_r(x, y)$. Since the integrand in the displayed equation is bounded and right continuous, the integral is a continuous function of t . Hence $t \mapsto P_t(x, y)$ is continuous for $(x, y) \in S^2$.
2. For $x \in S$, note that $P_t(x, x) \geq e^{-\lambda(x)t} > 0$ for $t \in [0, \infty)$. If x leads to y and $x \neq y$ then there exists $n \in \mathbb{N}_+$ and $(x_1, x_2, \dots, x_{n-1}) \in S^{n-1}$ such that $Q(x, x_1) > 0, \dots, Q(x_{n-1}, y) > 0$. Then

$$P_t(x, y) = \mathbb{P}(X_t = y \mid X_0 = x) \geq \mathbb{P}(Y_1 = x_1, \dots, Y_{n-1} = x_{n-1}, Y_n = y, \tau_n \leq t < \tau_{n+1}) > 0 \quad (16.16.24)$$

3. This is clear from the definition of the embedded chain \mathbf{Y} .

Parts (b) and (c) are known as the *Lévy dichotomy*, named for Paul Lévy. It's possible to prove the Lévy dichotomy just from the semigroup property of \mathbf{P} , but this proof is considerably more complicated. In light of the dichotomy, the *leads to* relation clearly makes

sense for the continuous-time chain \mathbf{X} as well as the discrete-time embedded chain \mathbf{Y} .

The Generator Matrix

Definition and Basic Properties

In this discussion, we assume again that $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is a regular Markov chain on S with transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$, exponential parameter function λ and one-step transition matrix Q for the embedded jump chain. The fundamental integral equation above now implies that the transition probability matrix P_t is differentiable in t . The derivative at 0 is particularly important.

The matrix function $t \mapsto P_t$ has a (right) derivative at 0:

$$\frac{P_t - I}{t} \rightarrow G \text{ as } t \downarrow 0 \quad (16.16.25)$$

where the *infinitesimal generator matrix* G is given by $G(x, y) = -\lambda(x)I(x, y) + \lambda(x)Q(x, y)$ for $(x, y) \in S^2$.

Proof

As before the change of variables $r = t - s$ in the fundamental integral equation gives

$$P_t(x, y) = I(x, y)e^{-\lambda(x)t} + \lambda(x)e^{-\lambda(x)t} \int_0^t e^{\lambda(x)r} Q P_r(x, y) dr \quad (16.16.26)$$

The first term is clearly differentiable in t , and the second term is also differentiable in t since we now know that the integrand is a continuous function of r . The result then follows from standard calculus.

Note that $\lambda(x)Q(x, x) = 0$ for every $x \in S$, since $\lambda(x) = 0$ if x is absorbing, while $Q(x, x) = 0$ if x is stable. So $G(x, x) = -\lambda(x)$ for $x \in S$, and $G(x, y) = \lambda(x)Q(x, y)$ for $(x, y) \in S^2$ with $y \neq x$. Thus, the generator matrix G determines the exponential parameter function λ and the jump transition matrix Q , and thus determines the distribution of the Markov chain \mathbf{X} .

Given the generator matrix G of \mathbf{X} ,

1. $\lambda(x) = -G(x, x)$ for $x \in S$
2. $Q(x, y) = -G(x, y)/G(x, x)$ if $x \in S$ is stable and $y \in S - \{x\}$

The infinitesimal generator has a nice interpretation in terms of our discussion in the last section. Recall that when the chain first enters a stable state x , we set independent, exponentially distributed “timers” on (x, y) , for each $y \in S - \{x\}$. Note that $G(x, y)$ is the exponential parameter for the timer on (x, y) . As soon as an alarm sounds for a particular (x, y) , the chain moves to state y and the process continues.

The generator matrix G satisfies the following properties for every $x \in S$:

1. $G(x, x) \leq 0$
2. $\sum_{y \in S} G(x, y) = 0$

The matrix function $t \mapsto P_t$ is differentiable on $[0, \infty)$, and satisfies the *Kolmogorov backward equation*: $P'_t = GP_t$. Explicitly,

$$P'_t(x, y) = -\lambda(x)P_t(x, y) + \sum_{z \in S} \lambda(x)Q(x, z)P_t(z, y), \quad (x, y) \in S^2 \quad (16.16.27)$$

Proof

The proof is just like before, and follows from standard calculus and the integral equation

$$P_t(x, y) = I(x, y)e^{-\lambda(x)t} + \lambda(x)e^{-\lambda(x)t} \int_0^t e^{\lambda(x)r} Q P_r(x, y) dr \quad (16.16.28)$$

The backward equation is named for Andrei Kolmogorov. In continuous time, the transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ can be obtained from the single, generator matrix G in a way that is reminiscent of the fact that in discrete time, the transition semigroup $\mathbf{P} = \{P^n : n \in \mathbb{N}\}$ can be obtained from the single, one-step matrix P . From a modeling point of view, we often start with the generator matrix G and then solve the backward equation, subject to the initial condition $P_0 = I$, to obtain the semigroup of transition matrices \mathbf{P} .

As with any matrix on S , the generator matrix G defines left and right operations on functions that are analogous to ordinary matrix multiplication. The right operation is defined for functions in \mathcal{B} .

If $f \in \mathcal{B}$ then Gf is given by

$$Gf(x) = -\lambda(x)f(x) + \sum_{y \in S} \lambda(x)Q(x, y)f(y), \quad x \in S \quad (16.16.29)$$

Proof

By definition,

$$Gf(x) = \sum_{y \in S} G(x, y)f(y) = -\lambda(x)f(x) + \sum_{y \in S - \{x\}} \lambda(x)Q(x, y)f(y) \quad (16.16.30)$$

In the second term, we can sum over all $y \in S$ since $\lambda(x) = 0$ if x is absorbing and $Q(x, x) = 0$ if x is stable. Note that Gf is well defined since

$$\sum_{y \in S - \{x\}} \lambda(x)Q(x, y) |f(y)| \leq \sum_{y \in S - \{x\}} \lambda(x)Q(x, y) \|f\| = \lambda(x) \|f\| \quad (16.16.31)$$

But note that Gf is not in \mathcal{B} unless $\lambda \in \mathcal{B}$. Without this additional assumption, G is a linear operator from the vector space \mathcal{B} of bounded functions from S to \mathbb{R} into the vector space of *all* functions from S to \mathbb{R} . We will return to this point in our next discussion.

Uniform Transition Semigroups

We can obtain stronger results for the generator matrix if we impose stronger continuity assumptions on \mathbf{P} .

The transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ is *uniform* if $P_t(x, x) \rightarrow 1$ as $t \downarrow 0$ uniformly in $x \in S$.

If \mathbf{P} is uniform, then the operator function $t \mapsto P_t$ is continuous on the vector space \mathcal{B} .

Proof

The statement means that for $f \in \mathcal{B}$, the function $t \mapsto P_t f$ is continuous with respect to the supremum norm on \mathcal{B} .

As usual, we want to look at this new assumption from different points of view.

The following are equivalent:

1. The transition semigroup \mathbf{P} is uniform.
2. The exponential parameter function λ is bounded.
3. The generator matrix G defines a bounded linear operator on \mathcal{B} .

Proof

From our remarks above we know that $\lambda \in \mathcal{B}$ if and only if the generator matrix G defines a bounded linear operator on \mathcal{B} . So we just need to show the equivalence of (a) and (b). If $\lambda \in \mathcal{B}$ then

$$P_t(x, x) = \mathbb{P}(X_t = x \mid X_0 = x) \geq \mathbb{P}(\tau > t \mid X_0 = x) = \exp[-\lambda(x)t] \geq \exp(-\|\lambda\|t) \quad (16.16.32)$$

The last term converges to 1 as $t \downarrow 0$ uniformly in x .

So when the equivalent conditions are satisfied, the Markov chain $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ is also said to be uniform. As we will see in a later section, a uniform, continuous-time Markov chain can be constructed from a discrete-time Markov chain and an independent Poisson process. For a uniform transition semigroup, we have a companion to the backward equation.

Suppose that \mathbf{P} is a uniform transition semigroup. Then $t \mapsto P_t$ satisfies the *Kolmogorov forward equation* $P'_t = P_t G$. Explicitly,

$$P'_t(x, y) = -\lambda(y)P_t(x, y) + \sum_{z \in S} P_t(x, z)\lambda(z)Q(z, y), \quad (x, y) \in S^2 \quad (16.16.33)$$

The backward equation holds with more generality than the forward equation, since we only need the transition semigroup \mathbf{P} to be *standard* rather than *uniform*. It would seem that we need stronger conditions on λ for the forward equation to hold, for otherwise it's not even obvious that $\sum_{z \in S} P_t(x, z)\lambda(z)Q(z, y)$ is finite for $(x, y) \in S$. On the other hand, the forward equation is sometimes easier to

solve than the backward equation, and the assumption that λ is bounded is met in many applications (and of course holds automatically if S is finite).

As a simple corollary, the transition matrices and the generator matrix commute for a uniform semigroup: $P_t G = G P_t$ for $t \in [0, \infty)$. The forward and backward equations formally look like the differential equations for the exponential function. This actually holds with the operator exponential.

Suppose again that $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ is a uniform transition semigroup with generator G . Then

$$P_t = e^{tG} = \sum_{n=0}^{\infty} \frac{t^n}{n!} G^n, \quad t \in [0, \infty) \quad (16.16.34)$$

Proof

First e^{tG} is well defined as a bounded linear operator on \mathcal{B} for $t \in [0, \infty)$ (and hence also simply as a matrix), since G is a bounded linear operator on \mathcal{B} . Trivially $e^{0G} = I$, and by basic properties of the matrix exponential,

$$\frac{d}{dt} e^{tG} = G e^{tG}, \quad t \in (0, \infty) \quad (16.16.35)$$

It follows that $P_t = e^{tG}$ for $t \in [0, \infty)$.

We can characterize the generators of uniform transition semigroups. We just need the minimal conditions that the diagonal entries are nonpositive and the row sums are 0.

Suppose that G a matrix on S with $\|G\| < \infty$. Then G is the generator of a uniform transition semigroup $\mathbf{P} = \{P_t : t \in [0, \infty)\}$ if and only if for every $x \in S$,

1. $G(x, x) \leq 0$
2. $\sum_{y \in S} G(x, y) = 0$

Proof

We know of course that if G is the generator of a transition semigroup, then conditions (a) and (b) hold. For the converse, we can use the previous result. Let

$$P_t = e^{tG} = \sum_{n=0}^{\infty} \frac{t^n}{n!} G^n, \quad t \in [0, \infty) \quad (16.16.36)$$

which makes sense since G is bounded in norm. Then $P_t(x, y) \geq 0$ for $(x, y) \in S^2$. By part (b), $\sum_{y \in S} G^n(x, y) = 0$ for every $x \in S$ and $n \in \mathbb{N}_+$, and hence $\sum_{y \in S} P_t(x, y) = \sum_{y \in S} I(x, y) = 1$ for $x \in S$. Finally, the semigroup property is a consequence of the law of exponents, which holds for the exponential of a matrix.

$$P_s P_t = e^{sG} e^{tG} = e^{(s+t)G} = P_{s+t} \quad (16.16.37)$$

Examples and Exercises

The Two-State Chain

Let $\mathbf{X} = \{X_t : t \in [0, \infty)\}$ be the Markov chain on the set of states $S = \{0, 1\}$, with transition rate $a \in [0, \infty)$ from 0 to 1 and transition rate $b \in [0, \infty)$ from 1 to 0. This *two-state Markov chain* was studied in the previous section. To avoid the trivial case with both states absorbing, we will assume that $a + b > 0$.

The generator matrix is

$$G = \begin{bmatrix} -a & a \\ b & -b \end{bmatrix} \quad (16.16.38)$$

Show that for $t \in [0, \infty)$,

$$P_t = \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix} - \frac{1}{a+b} e^{-(a+b)t} \begin{bmatrix} -a & a \\ b & -b \end{bmatrix} \quad (16.16.39)$$

1. By solving the Kolmogorov backward equation.

2. By solving the Kolmogorov forward equation.
3. By computing $P_t = e^{tG}$.

You probably noticed that the forward equation is easier to solve because there is less coupling of terms than in the backward equation.

Define the probability density function f on S by $f(0) = \frac{b}{a+b}$, $f(1) = \frac{a}{a+b}$. Show that

1. $P_t \rightarrow \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix}$ as $t \rightarrow \infty$, the matrix with f in both rows.
2. $fP_t = f$ for all $t \in [0, \infty)$, so that f is invariant for P .
3. $fG = 0$.

Computational Exercises

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 embedded transition matrix

$$Q = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 \end{bmatrix} \quad (16.16.40)$$

1. Draw the state graph and classify the states.
2. Find the generator matrix G .
3. Find the transition matrix P_t for $t \in [0, \infty)$.
4. Find $\lim_{t \rightarrow \infty} P_t$.

Answer

1. The edge set is $E = \{(0, 1), (0, 2), (1, 0), (2, 0), (2, 1)\}$. All states are stable.
2. The generator matrix is

$$G = \begin{bmatrix} -4 & 2 & 2 \\ 1 & -1 & 0 \\ 1 & 2 & -3 \end{bmatrix} \quad (16.16.41)$$

3. For $t \in [0, \infty)$,

$$P_t = \frac{1}{15} \begin{bmatrix} 3 + 12e^{-5t} & 10 - 10e^{-3t} & 2 - 12e^{-5t} + 10e^{-3t} \\ 3 - 3e^{-5t} & 10 + 5e^{-3t} & 2 + 3e^{-5t} - 5e^{-3t} \\ 3 - 3e^{-5t} & 10 - 10e^{-3t} & 2 + 3e^{-5t} + 10e^{-3t} \end{bmatrix} \quad (16.16.42)$$

- 4.

$$P_t \rightarrow \frac{1}{15} \begin{bmatrix} 3 & 10 & 2 \\ 3 & 10 & 2 \\ 3 & 10 & 2 \end{bmatrix} \quad (16.16.43)$$

Special Models

Read the discussion of generator and transition matrices for chains subordinate to the Poisson process.

Read the discussion of the infinitesimal generator for continuous-time birth-death chains.

Read the discussion of the infinitesimal generator for continuous-time queuing chains.

Read the discussion of the infinitesimal generator for continuous-time branching chains.

This page titled [16.16: Transition Matrices and Generators 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.