

## 16.15: Introduction to Continuous-Time Markov Chains

This section begins our study of Markov processes in continuous time and with discrete state spaces. Recall that a Markov process with a discrete state space is called a *Markov chain*, so we are studying *continuous-time Markov chains*. It will be helpful if you review the section on general Markov processes, at least briefly, to become familiar with the basic notation and concepts. Also, discrete-time chains plays a fundamental role, so you will need review this topic also.

We will study continuous-time Markov chains from different points of view. Our point of view in this section, involving holding times and the embedded discrete-time chain, is the most intuitive from a probabilistic point of view, and so is the best place to start. In the next section, we study the transition probability matrices in continuous time. This point of view is somewhat less intuitive, but is closest to how other types of Markov processes are treated. Finally, in the third introductory section we study the Markov chain from the view point of potential matrices. This is the least intuitive approach, but analytically one of the best. Naturally, the interconnections between the various approaches are particularly important.

### Preliminaries

As usual, we start with a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , so that  $\Omega$  is the set of outcomes,  $\mathcal{F}$  the  $\sigma$ -algebra of events, and  $\mathbb{P}$  the probability measure on the sample space  $(\Omega, \mathcal{F})$ . The time space is  $([0, \infty), \mathcal{T})$  where as usual,  $\mathcal{T}$  is the Borel  $\sigma$ -algebra on  $[0, \infty)$  corresponding to the standard Euclidean topology. The state space is  $(S, \mathcal{S})$  where  $S$  is countable and  $\mathcal{S}$  is the power set 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  $\sigma$  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.

Suppose now that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is stochastic process with state space  $(S, \mathcal{S})$ . For  $t \in [0, \infty)$ , let  $\mathcal{F}_t^0 = \sigma\{X_s : s \in [0, t]\}$ , so that  $\mathcal{F}_t^0$  is the  $\sigma$ -algebra of events defined by the process up to time  $t$ . The collection of  $\sigma$ -algebras  $\mathfrak{F}^0 = \{\mathcal{F}_t^0 : t \in [0, \infty)\}$  is the natural filtration associated with  $\mathbf{X}$ . For technical reasons, it's often necessary to have a filtration  $\mathfrak{F} = \{\mathcal{F}_t : t \in [0, \infty)\}$  that is slightly finer than the natural one, so that  $\mathcal{F}_t^0 \subseteq \mathcal{F}_t$  for  $t \in [0, \infty)$  (or in equivalent jargon,  $\mathbf{X}$  is adapted to  $\mathfrak{F}$ ). See the general introduction for more details on the common ways that the natural filtration is refined. We will also let  $\mathcal{G}_t = \sigma\{X_s : s \in [t, \infty)\}$ , the  $\sigma$ -algebra of events defined by the process from time  $t$  onward. If  $t$  is thought of as the present time, then  $\mathcal{F}_t$  is the collection of events in the past and  $\mathcal{G}_t$  is the collection of events in the future.

It's often necessary to impose assumptions on the continuity of the process  $\mathbf{X}$  in time. Recall that  $\mathbf{X}$  is *right continuous* if  $t \mapsto X_t(\omega)$  is right continuous on  $[0, \infty)$  for every  $\omega \in \Omega$ , and similarly  $\mathbf{X}$  has left limits if  $t \mapsto X_t(\omega)$  has left limits on  $(0, \infty)$  for every  $\omega \in \Omega$ . Since  $S$  has the discrete topology, note that if  $\mathbf{X}$  is right continuous, then for every  $t \in [0, \infty)$  and  $\omega \in \Omega$ , there exists  $\epsilon$  (depending on  $t$  and  $\omega$ ) such that  $X_{t+s}(\omega) = X_t(\omega)$  for  $s \in [0, \epsilon)$ . Similarly, if  $\mathbf{X}$  has left limits, then for every  $t \in (0, \infty)$  and  $\omega \in \Omega$  there exists  $\delta$  (depending on  $t$  and  $\omega$ ) such that  $X_{t-s}(\omega)$  is constant for  $s \in (0, \delta)$ .

### The Markov Property

There are a number of equivalent ways to state the Markov property. At the most basic level, the property states that the *past and future are conditionally independent, given the present*.

The process  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a *Markov chain* on  $S$  if for every  $t \in [0, \infty)$ ,  $A \in \mathcal{F}_t$ , and  $B \in \mathcal{G}_t$ ,

$$\mathbb{P}(A \cap B \mid X_t) = \mathbb{P}(A \mid X_t) \mathbb{P}(B \mid X_t) \quad (16.15.1)$$

Another version is that the conditional distribution of a state in the future, given the past, is the same as the conditional distribution just given the present state.

The process  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$  if for every  $s, t \in [0, \infty)$ , and  $x \in S$ ,

$$\mathbb{P}(X_{s+t} = x \mid \mathcal{F}_s) = \mathbb{P}(X_{s+t} = x \mid X_s) \quad (16.15.2)$$

Technically, in the last two definitions, we should say that  $\mathbf{X}$  is a Markov process *relative to* the filtration  $\mathfrak{F}$ . But recall that if  $\mathbf{X}$  satisfies the Markov property relative to a filtration, then it satisfies the Markov property relative to any coarser filtration, and in particular, relative to the natural filtration. For the natural filtration, the Markov property can also be stated without explicit reference to  $\sigma$ -algebras, although at the cost of additional clutter:

The process  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$  if and only if for every  $n \in \mathbb{N}_+$ , time sequence  $(t_1, t_2, \dots, t_n) \in [0, \infty)^n$  with  $t_1 < t_2 < \dots < t_n$ , and state sequence  $(x_1, x_2, \dots, x_n) \in S^n$ ,

$$\mathbb{P}(X_{t_n} = x_n \mid X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_{n-1}} = x_{n-1}) = \mathbb{P}(X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}) \quad (16.15.3)$$

As usual, we also assume that our Markov chain  $\mathbf{X}$  is *time homogeneous*, so that  $\mathbb{P}(X_{s+t} = y \mid X_s = x) = \mathbb{P}(X_t = y \mid X_0 = x)$  for  $s, t \in [0, \infty)$  and  $x, y \in S$ . So, for a homogeneous Markov chain on  $S$ , the process  $\{X_{s+t} : t \in [0, \infty)\}$  given  $X_s = x$ , is independent of  $\mathcal{F}_s$  and equivalent to the process  $\{X_t : t \in [0, \infty)\}$  given  $X_0 = x$ , for every  $s \in [0, \infty)$  and  $x \in S$ . That is, if the chain is in state  $x \in S$  at a particular time  $s \in [0, \infty)$ , it does not matter how the chain *got to*  $x$ ; the chain essentially starts over in state  $x$ .

### The Strong Markov Property

Random times play an important role in the study of continuous-time Markov chains. It's often necessary to allow random times to take the value  $\infty$ , so formally, a random time  $\tau$  is a random variable on the underlying sample space  $(\Omega, \mathcal{F})$  taking values in  $[0, \infty]$ . Recall also that a random time  $\tau$  is a stopping time (also called a *Markov time* or an *optional time*) if  $\{\tau \leq t\} \in \mathcal{F}_t$  for every  $t \in [0, \infty)$ . If  $\tau$  is a stopping time, the  $\sigma$ -algebra associated with  $\tau$  is

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \in [0, \infty)\} \quad (16.15.4)$$

So  $\mathcal{F}_\tau$  is the collection of events up to the random time  $\tau$  in the same way that  $\mathcal{F}_t$  is the collection of events up to the deterministic time  $t \in [0, \infty)$ . We usually want the Markov property to extend from deterministic times to stopping times.

The process  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a *strong Markov chain* on  $S$  if for every stopping time  $\tau, t \in [0, \infty)$ , and  $x \in S$ ,

$$\mathbb{P}(X_{\tau+t} = x \mid \mathcal{F}_\tau) = \mathbb{P}(X_{\tau+t} = x \mid X_\tau) \quad (16.15.5)$$

So, for a homogeneous strong Markov chain on  $S$ , the process  $\{X_{\tau+t} : t \in [0, \infty)\}$  given  $X_\tau = x$ , is independent of  $\mathcal{F}_\tau$  and equivalent to the process  $\{X_t : t \in [0, \infty)\}$  given  $X_0 = x$ , for every stopping time  $\tau$  and  $x \in S$ . That is, if the chain is in state  $x \in S$  at a stopping time  $\tau$ , then the chain essentially starts over at  $x$ , independently of the past.

### Holding Times and the Jump Chain

For our first point of view, we still study *when* and *how* our Markov chain  $\mathbf{X}$  changes state. The discussion depends heavily on properties of the exponential distribution, so we need a quick review.

#### The Exponential Distribution

A random variable  $\tau$  has the *exponential distribution* with rate parameter  $r \in (0, \infty)$  if  $\tau$  has a continuous distribution on  $[0, \infty)$  with probability density function  $f$  given by  $f(t) = re^{-rt}$  for  $t \in [0, \infty)$ . Equivalently, the right distribution function  $F^c$  is given by

$$F^c(t) = \mathbb{P}(\tau > t) = e^{-rt}, \quad t \in [0, \infty) \quad (16.15.6)$$

The mean of the distribution is  $1/r$  and the variance is  $1/r^2$ . The exponential distribution has an amazing number of characterizations. One of the most important is the *memoryless property* which states that a random variable  $\tau$  with values in  $[0, \infty)$  has an exponential distribution if and only if the conditional distribution of  $\tau - s$  given  $\tau > s$  is the same as the distribution of  $\tau$  itself, for every  $s \in [0, \infty)$ . It's easy to see that the memoryless property is equivalent to the *law of exponents* for right distribution function  $F^c$ , namely  $F^c(s+t) = F^c(s)F^c(t)$  for  $s, t \in [0, \infty)$ . Since  $F^c$  is right continuous, the only solutions are exponential functions.

For our study of continuous-time Markov chains, it's helpful to extend the exponential distribution to two degenerate cases,  $\tau = 0$  with probability 1, and  $\tau = \infty$  with probability 1. In terms of the parameter, the first case corresponds to  $r = \infty$  so that  $F(t) = \mathbb{P}(\tau > t) = 0$  for every  $t \in [0, \infty)$ , and the second case corresponds to  $r = 0$  so that  $F(t) = \mathbb{P}(\tau > t) = 1$  for every  $t \in [0, \infty)$ . Note that in both cases, the function  $F$  satisfies the law of exponents, and so corresponds to a memoryless distribution in a general sense. In all cases, the mean of the exponential distribution with parameter  $r \in [0, \infty]$  is  $1/r$ , where we interpret  $1/0 = \infty$  and  $1/\infty = 0$ .

#### Holding Times

The Markov property implies the memoryless property for the random time when a Markov process first leaves its initial state. It follows that this random time must have an exponential distribution.

Suppose that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$ , and let  $\tau = \inf\{t \in [0, \infty) : X_t \neq X_0\}$ . For  $x \in S$ , the conditional distribution of  $\tau$  given  $X_0 = x$  is exponential with parameter  $\lambda(x) \in [0, \infty]$ .

**Proof**

Let  $x \in S$  and  $s \in [0, \infty)$ . The events  $X_0 = x$  and  $\tau > s$  imply  $X_s = x$ . By the Markov property, given  $X_s = x$ , the chain starts over at time  $s$  in state  $x$ , independent of  $\{X_0 = x\}$  and  $\{\tau > s\}$ , since both events are in  $\mathcal{F}_s$ . Hence for  $t \in [0, \infty)$ ,

$$\mathbb{P}(\tau > t+s \mid X_0 = x, \tau > s) = \mathbb{P}(\tau > t+s \mid X_0 = x, X_s = x, \tau > s) = \mathbb{P}(\tau > t \mid X_0 = x) \quad (16.15.7)$$

It follows that  $\tau$  has the memoryless property, and hence has an exponential distribution with parameter  $\lambda(x) \in [0, \infty]$ .

So, associated with the Markov chain  $\mathbf{X}$  on  $S$  is a function  $\lambda : S \rightarrow [0, \infty]$  that gives the exponential parameters for the holding times in the states. Considering the ordinary exponential distribution, and the two degenerate versions, we are led to the following classification of states:

Suppose again that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$  with exponential parameter function  $\lambda$ . Let  $x \in S$ .

1. If  $\lambda(x) = 0$  then  $\mathbb{P}(\tau = \infty \mid X_0 = x) = 1$ , and  $x$  is said to be an *absorbing state*.
2. If  $\lambda(x) \in (0, \infty)$  then  $\mathbb{P}(0 < \tau < \infty \mid X_0 = x) = 1$  and  $x$  is said to be an *stable state*.
3. If  $\lambda(x) = \infty$  then  $\mathbb{P}(\tau = 0 \mid X_0 = x) = 1$ , and  $x$  is said to be an *instantaneous state*.

As you can imagine, an instantaneous state corresponds to weird behavior, since the chain starting in the state leaves the state at times arbitrarily close to 0. While mathematically possible, instantaneous states make no sense in most applications, and so are to be avoided. Also, the proof of the last result has some technical holes. We did not really show that  $\tau$  is a valid random time, let alone a stopping time. Fortunately, one of our standard assumptions resolves these problems.

Suppose again that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$ . If the process  $\mathbf{X}$  and the filtration  $\mathfrak{F}$  are right continuous, then

1.  $\tau$  is a stopping time.
2.  $\mathbf{X}$  has no instantaneous states.
3.  $\mathbb{P}(X_\tau \neq x \mid X_0 = x) = 1$  if  $x \in S$  is stable.
4.  $\mathbf{X}$  is a strong Markov process.

Proof

1. Let  $t \in [0, \infty)$ . By right continuity,

$$\{\tau < t\} = \{X_s \neq X_0 \text{ for some } s \in (0, t)\} = \{X_s \neq X_0 \text{ for some rational } s \in (0, t)\} \quad (16.15.8)$$

But for  $s \in (0, t)$ ,  $\{X_s \neq X_0\} \in \mathcal{F}_s \subseteq \mathcal{F}_t$ . The last event in the displayed equation is a countable union, so  $\{\tau < t\} \in \mathcal{F}_t$ . Since  $\mathfrak{F}$  is right continuous,  $\tau$  is a stopping time.

2. Suppose that  $\omega \in \Omega$  and  $X_0(\omega) = x$ . Since  $\mathbf{X}$  is right continuous, there exists  $\epsilon > 0$  such that  $X_t(\omega) = x$  for  $0 \leq t < \epsilon$  and hence  $\tau(\omega) \geq \epsilon > 0$ . So  $\mathbb{P}(\tau > 0 \mid X_0 = x) = 1$ .
3. Similarly, suppose that  $\omega \in \Omega$  and that  $X_0(\omega) = x$  and  $X_{\tau(\omega)}(\omega) = y$ . Since  $\mathbf{X}$  is right continuous, there exists  $\epsilon > 0$  such that  $X_t(\omega) = y$  for  $\tau(\omega) \leq t < \tau(\omega) + \epsilon$ . But by definition of  $\tau(\omega)$ , there exists  $t \in (\tau(\omega), \tau(\omega) + \epsilon)$  with  $X_t(\omega) \neq x$ . Hence  $\mathbb{P}(X_\tau \neq x \mid X_0 = x) = 1$ .

There is actually a converse to part (b) that states that if  $\mathbf{X}$  has no instantaneous states, then there is a version of  $\mathbf{X}$  that is right continuous. From now on, we will assume that our Markov chains are right continuous with probability 1, and hence have no instantaneous states. On the other hand, absorbing states are perfectly reasonable and often do occur in applications. Finally, if the chain enters a stable state, it will stay there for a (proper) exponentially distributed time, and then leave.

### The Jump Chain

Without instantaneous states, we can now construct a sequence of stopping times. Basically, we let  $\tau_n$  denote the  $n$ th time that the chain changes state for  $n \in \mathbb{N}_+$ , unless the chain has previously been caught in an absorbing state. Here is the formal construction:

Suppose again that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$ . Let  $\tau_0 = 0$  and  $\tau_1 = \inf\{t \in [0, \infty) : X_t \neq X_0\}$ . Recursively, suppose that  $\tau_n$  is defined for  $n \in \mathbb{N}_+$ . If  $\tau_n = \infty$  let  $\tau_{n+1} = \infty$ . Otherwise, let

$$\tau_{n+1} = \inf\{t \in [\tau_n, \infty) : X_t \neq X_{\tau_n}\} \quad (16.15.9)$$

Let  $M = \sup\{n \in \mathbb{N} : \tau_n < \infty\}$ .

In the definition of  $M$ , of course,  $\sup(\mathbb{N}) = \infty$ , so  $M$  is the number of changes of state. If  $M < \infty$ , the chain was sucked into an absorbing state at time  $\tau_M$ . Since we have ruled out instantaneous states, the sequence of random times is strictly increasing up until the (random) term  $M$ . That is, with probability 1, if  $n \in \mathbb{N}$  and  $\tau_n < \infty$  then  $\tau_n < \tau_{n+1}$ . Of course by construction, if  $\tau_n = \infty$  then  $\tau_{n+1} = \infty$ . The increments  $\tau_{n+1} - \tau_n$  for  $n \in \mathbb{N}$  with  $n < M$  are the times spent in the states visited by  $\mathbf{X}$ . The process at the random times when the state changes forms an embedded discrete-time Markov chain.

Suppose again that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$ . Let  $\{\tau_n : n \in \mathbb{N}\}$  denote the stopping times and  $M$  the random index, as defined above. For  $n \in \mathbb{N}$ , let  $Y_n = X_{\tau_n}$  if  $n \leq M$  and  $Y_n = X_{\tau_M}$  if  $n > M$ . Then  $\mathbf{Y} = \{Y_n : n \in \mathbb{N}\}$  is a (homogenous) discrete-time Markov chain on  $S$ , known as the *jump chain* of  $\mathbf{X}$ .

Proof

For  $n \in \mathbb{N}$  let  $\mathcal{G}_n = \sigma\{Y_0, Y_1, \dots, Y_n\}$ , the  $\sigma$ -algebra of events for the process  $\mathbf{Y}$ , up to the discrete time  $n$ . Let  $x \in S$ . If  $x$  is stable, then given  $Y_n = x$ , the random times  $\tau_n$  and  $\tau_{n+1}$  are finite with probability 1. (Note that we cannot get to  $x$  from an absorbing state.) So

$$\mathbb{P}(Y_{n+1} = y \mid Y_n = x, \mathcal{G}_n) = \mathbb{P}(X_{\tau_{n+1}} = y \mid X_{\tau_n} = x, \mathcal{G}_n), \quad y \in S \quad (16.15.10)$$

But by the strong Markov property, given  $X_{\tau_n} = x$ , the chain starts over at time  $\tau_n$  in state  $x$ , independent of  $\mathcal{G}_n \subseteq \mathcal{F}_{\tau_n}$ . Hence

$$\mathbb{P}(Y_{n+1} = y \mid Y_n = x, \mathcal{G}_n) = \mathbb{P}(X_\tau = y \mid X_0 = x), \quad y \in S \quad (16.15.11)$$

On the other hand, if  $x$  is an absorbing state, then by construction,

$$\mathbb{P}(Y_{n+1} = y \mid Y_n = x, \mathcal{G}_n) = I(x, y), \quad y \in S \quad (16.15.12)$$

where  $I$  is the identity matrix on  $S$ .

As noted in the proof, the one-step transition probability matrix  $Q$  for the jump chain  $\mathbf{Y}$  is given for  $(x, y) \in S^2$  by

$$Q(x, y) = \begin{cases} \mathbb{P}(X_\tau = y \mid X_0 = x), & x \text{ stable} \\ I(x, y), & x \text{ absorbing} \end{cases} \quad (16.15.13)$$

where  $I$  is the identity matrix on  $S$ . Of course  $Q$  satisfies the usual properties of a probability matrix on  $S$ , namely  $Q(x, y) \geq 0$  for  $(x, y) \in S^2$  and  $\sum_{y \in S} Q(x, y) = 1$  for  $x \in S$ . But  $Q$  satisfies another interesting property as well. Since the state actually *changes* at time  $\tau$  starting in a stable state, we must have  $Q(x, x) = 0$  if  $x$  is stable and  $Q(x, x) = 1$  if  $x$  is absorbing.

Given the initial state, the holding time and the next state are independent.

If  $x, y \in S$  and  $t \in [0, \infty)$  then  $\mathbb{P}(Y_1 = y, \tau_1 > t \mid Y_0 = x) = Q(x, y)e^{-\lambda(x)t}$

**Proof**

Suppose that  $x$  is a stable state, so that given  $Y_0 = X_0 = x$ , the stopping time  $\tau_1 = \tau$  has a proper exponential distribution with parameter  $\lambda(x) \in (0, \infty)$ . Note that

$$\mathbb{P}(Y_1 = y, \tau_1 > t \mid Y_0 = x) = \mathbb{P}(X_\tau = y, \tau > t \mid X_0 = x) = \mathbb{P}(X_\tau = y \mid \tau > t, X_0 = x) \mathbb{P}(\tau > t \mid X_0 = x) \quad (16.15.14)$$

Note that if  $X_0 = x$  and  $\tau > t$  then  $X_t = x$  also. By the Markov property, given  $X_t = x$ , the chain starts over at time  $t$  in state  $x$ , independent of  $\{X_0 = x\}$  and  $\{\tau > t\}$ , both events in  $\mathcal{F}_t$ . Hence

$$\mathbb{P}(X_\tau = y \mid \tau > t, X_0 = x) = \mathbb{P}(X_\tau = y \mid X_t = x, \tau > t, X_0 = x) = \mathbb{P}(X_\tau = y \mid X_0 = x) = Q(x, y) \quad (16.15.15)$$

Of course  $\mathbb{P}(\tau > t \mid X_0 = x) = e^{-\lambda(x)t}$ .

If  $x$  is an absorbing state then  $\mathbb{P}(\tau = \infty \mid X_0 = x) = 1$ ,  $\mathbb{P}(Y_1 = x \mid Y_0 = x) = 1$ , and  $\lambda(x) = 0$ . Hence

$$\mathbb{P}(Y_1 = y, \tau_1 > t \mid Y_0 = x) = I(x, y) = Q(x, y)e^{-\lambda(x)t} \quad (16.15.16)$$

The following theorem is a generalization. The changes in state and the holding times are independent, given the initial state.

Suppose that  $n \in \mathbb{N}_+$  and that  $(x_0, x_1, \dots, x_n)$  is a sequence of stable states and  $(t_1, t_2, \dots, t_n)$  is a sequence in  $[0, \infty)$ . Then

$$\begin{aligned} & \mathbb{P}(Y_1 = x_1, \tau_1 > t_1, Y_2 = x_2, \tau_2 - \tau_1 > t_2, \dots, Y_n = x_n, \tau_n - \tau_{n-1} > t_n \mid Y_0 = x_0) \\ &= Q(x_0, x_1)e^{-\lambda(x_0)t_1} Q(x_1, x_2)e^{-\lambda(x_1)t_2} \dots Q(x_{n-1}, x_n)e^{-\lambda(x_{n-1})t_n} \end{aligned}$$

**Proof**

The proof is by induction, and the essence is captured in the case  $n = 2$ . So suppose that  $x_0, x_1, x_2$  are stable states and  $t_1, t_2 \in [0, \infty)$ . Then

$$\begin{aligned} & \mathbb{P}(Y_1 = x_1, \tau_1 > t_1, Y_2 = x_2, \tau_2 - \tau_1 > t_2 \mid Y_0 = x_0) \\ &= \mathbb{P}(Y_2 = x_2, \tau_2 - \tau_1 > t_2 \mid X_0 = x, Y_1 = x_1, \tau_1 > t_1) \mathbb{P}(Y_1 = x_1, \tau_1 > t_1 \mid Y_0 = x_0) \end{aligned}$$

But  $\mathbb{P}(Y_1 = x_1, \tau_1 > t_1 \mid Y_0 = x_0) = Q(x_0, x_1)e^{-\lambda(x_0)t_1}$  by the previous theorem. Next, by definition,

$$\mathbb{P}(Y_2 = x_2, \tau_2 - \tau_1 > t_2 \mid X_0 = x, Y_1 = x_1, \tau_1 > t_1) = \mathbb{P}(X_{\tau_2} = x_2, \tau_2 - \tau_1 > t_2 \mid X_0 = x_0, X_{\tau_1} = x_1, \tau_1 > t_1) \quad (16.15.17)$$

But by the strong Markov property, given  $X_{\tau_1} = x_1$ , the chain starts over at time  $\tau_1$  in state  $x$ , independent of the events  $\{X_0 = x_0\}$  and  $\{\tau_1 > t_1\}$  (both events in  $\mathcal{F}_{\tau_1}$ ). Hence using the previous theorem again,

$$\mathbb{P}(Y_2 = x_2, \tau_2 - \tau_1 > t_2 \mid X_0 = x, Y_1 = x_1, \tau_1 > t_1) = \mathbb{P}(X_\tau = x_2, \tau > t_2 \mid X_0 = x_1) = Q(x_1, x_2)e^{-\lambda(x_1)t_2} \quad (16.15.18)$$

## Regularity

We now know quite a bit about the structure of a continuous-time Markov chain  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  (without instantaneous states). Once the chain enters a given state  $x \in S$ , the holding time in state  $x$  has an exponential distribution with parameter  $\lambda(x) \in [0, \infty)$ , after which the next state  $y \in S$  is chosen, independently of the holding time, with probability  $Q(x, y)$ . However, we don't know everything about the chain. For the sequence  $\{\tau_n : n \in \mathbb{N}\}$  defined above, let  $\tau_\infty = \lim_{n \rightarrow \infty} \tau_n$ , which exists in  $(0, \infty]$  of course, since the sequence is increasing. Even though the holding time in a state is positive with probability 1, it's possible that  $\tau_\infty < \infty$  with positive probability, in which case we know nothing about  $X_t$

for  $t \geq \tau_\infty$ . The event  $\{\tau_\infty < \infty\}$  is known as *explosion*, since it means that the  $\mathbf{X}$  makes infinitely many transitions before the finite time  $\tau_\infty$ . While not as pathological as the existence of instantaneous states, explosion is still to be avoided in most applications.

A Markov chain  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  on  $S$  is *regular* if each of the following events has probability 1:

1.  $\mathbf{X}$  is right continuous.
2.  $\tau_n \rightarrow \infty$  as  $n \rightarrow \infty$ .

There is a simple condition on the exponential parameters and the embedded chain that is equivalent to condition (b).

Suppose that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a right-continuous Markov chain on  $S$  with exponential parameter function  $\lambda$  and embedded chain  $\mathbf{Y} = (Y_0, Y_1, \dots)$ . Then  $\tau_n \rightarrow \infty$  as  $n \rightarrow \infty$  with probability 1 if and only if  $\sum_{n=0}^{\infty} 1/\lambda(Y_n) = \infty$  with probability 1.

*Proof*

Given  $\mathbf{Y} = (y_0, y_1, \dots)$ , the distribution of  $\tau_\infty = \lim_{n \rightarrow \infty} \tau_n$  is the distribution of  $T_\infty = \sum_{n=0}^{\infty} T_n$  where  $(T_0, T_1, \dots)$  are independent, and  $T_n$  has the exponential distribution with parameter  $\lambda(y_n)$ . Note that  $\mathbb{E}(T_\infty) = \sum_{n=0}^{\infty} 1/\lambda(y_n)$ . In the section on the exponential distribution, it's shown that  $\mathbb{P}(T_\infty = \infty) = 1$  if and only if  $\mathbb{E}(T_\infty) = \infty$ .

If  $\lambda$  is bounded, then  $\mathbf{X}$  is regular.

Suppose that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$  with exponential parameter function  $\lambda$ . If  $\lambda$  is bounded, then  $\mathbf{X}$  is regular.

*Proof*

Suppose that  $\lambda(x) \leq r$  for  $x \in S$ , where  $r \in (0, \infty)$ . Then in particular,  $\mathbf{X}$  has no instantaneous states and so is right continuous. Moreover,  $1/\lambda(x) \geq 1/r$  for  $x \in S$  so  $\sum_{n=0}^{\infty} 1/\lambda(Y_n) = \infty$  with probability 1, where as usual,  $\mathbf{Y} = (Y_0, Y_1, \dots)$  is the jump chain of  $\mathbf{X}$ .

Here is another sufficient condition that is useful when the state space is infinite.

Suppose that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a Markov chain on  $S$  with exponential parameter function  $\lambda : S \rightarrow [0, \infty)$ . Let  $S_+ = \{x \in S : \lambda(x) > 0\}$ . Then  $\mathbf{X}$  is regular if

$$\sum_{x \in S_+} \frac{1}{\lambda(x)} = \infty \quad (16.15.19)$$

*Proof*

By assumption,  $\lambda(x) < \infty$  for  $x \in S$ , so there are no instantaneous states and so we can take  $\mathbf{X}$  to be right continuous. Next,

$$\sum_{n=0}^{\infty} \frac{1}{\lambda(Y_n)} = \sum_{n=0}^{\infty} \sum_{x \in S} \frac{1}{\lambda(x)} \mathbf{1}(Y_n = x) = \sum_{x \in S} \frac{1}{\lambda(x)} \sum_{n=0}^{\infty} \mathbf{1}(Y_n = x) = \sum_{x \in S} \frac{N_x}{\lambda(x)} \quad (16.15.20)$$

where  $N_x = \sum_{n=0}^{\infty} \mathbf{1}(Y_n = x)$  is the number of times that the jump chain  $\mathbf{Y}$  is in state  $x$ . Suppose that  $\sum_{x \in S_+} 1/\lambda(x) = \infty$ . Note that it must be the case that  $S_+$ , and hence  $S$ , is infinite. With probability 1, either  $\mathbf{Y}$  enters an absorbing state (a state  $x \in S$  with  $\lambda(x) = 0$ ), or  $N_x = \infty$  for some  $x \in S_+$ , or  $N_x \geq 1$  for infinitely many  $x \in S_+$ . In any case,

$$\sum_{n=0}^{\infty} \frac{1}{\lambda(Y_n)} = \sum_{x \in S} \frac{N_x}{\lambda(x)} = \infty \quad (16.15.21)$$

As a corollary, note that if  $S$  is finite then  $\lambda$  is bounded, so a continuous-time Markov chain on a finite state space is regular. So to review, if the exponential parameter function  $\lambda$  is finite, the chain  $\mathbf{X}$  has no instantaneous states. Even better, if  $\lambda$  is bounded or if the conditions in the last theorem are satisfied, then  $\mathbf{X}$  is regular. A continuous-time Markov chain with bounded exponential parameter function  $\lambda$  is called *uniform*, for reasons that will become clear in the next section on transition matrices. As we will see in later section, a uniform continuous-time Markov chain can be constructed from a discrete-time chain and an independent Poisson process. For the next result, recall that to say that  $\mathbf{X}$  has *left limits* with probability 1 means that the random function  $t \mapsto X_t$  has limits from the left on  $(0, \infty)$  with probability 1.

If  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is regular then  $\mathbf{X}$  has left limits with probability 1.

*Proof*

Suppose first that there are no absorbing states. Under the assumptions, with probability 1,  $0 < \tau_n < \infty$  for each  $n \in \mathbb{N}$  and  $\tau_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Moreover,  $X_t = Y_n$  for  $t \in [\tau_n, \tau_{n+1})$  and  $n \in \mathbb{N}$ . So  $t \mapsto X_t$  has left limits on  $(0, \infty)$  with probability 1. The same basic argument works with absorbing states, except that possibly  $\tau_{n+1} = \infty$ .

Thus, our standard assumption will be that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a regular Markov chain on  $S$ . For such a chain, the behavior of  $\mathbf{X}$  is completely determined by the exponential parameter function  $\lambda$  that governs the holding times, and the transition probability matrix  $Q$  of the

jump chain  $\mathbf{Y}$ . Conversely, when modeling real stochastic systems, we often start with  $\lambda$  and  $Q$ . It's then relatively straightforward to construct the continuous-time Markov chain that has these parameters. For simplicity, we will assume that there are no absorbing states. The inclusion of absorbing states is not difficult, but mucks up the otherwise elegant exposition.

Suppose that  $\lambda : S \rightarrow (0, \infty)$  is bounded and that  $Q$  is a probability matrix on  $S$  with the property that  $Q(x, x) = 0$  for every  $x \in S$ . The regular, continuous-time Markov chain  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  with exponential parameter function  $\lambda$  and jump transition matrix  $Q$  can be constructed as follows:

1. First construct the jump chain  $\mathbf{Y} = (Y_0, Y_1, \dots)$  having transition matrix  $Q$ .
2. Next, given  $\mathbf{Y} = (x_0, x_1, \dots)$ , the transition times  $(\tau_1, \tau_2, \dots)$  are constructed so that the holding times  $(\tau_1, \tau_2 - \tau_1, \dots)$  are independent and exponentially distributed with parameters  $(\lambda(x_0), \lambda(x_1), \dots)$
3. Again given  $\mathbf{Y} = (x_0, x_1, \dots)$ , define  $X_t = x_0$  for  $0 \leq t < \tau_1$  and for  $n \in \mathbb{N}_+$ , define  $X_t = x_n$  for  $\tau_n \leq t < \tau_{n+1}$ .

Additional details

Using product sets and product measures, it's straightforward to construct a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with the following objects and properties:

1.  $\mathbf{Y} = (Y_0, Y_1, \dots)$  is a Markov chain on  $S$  with transition matrix  $Q$ .
2.  $\mathbf{T} = \{T_x : x \in S\}$  is a collection of independent random variables with values in  $[0, \infty)$  such that  $T_x$  has the exponential distribution with parameter  $\lambda(x)$  for each  $x \in S$ .
3.  $\mathbf{Y}$  and  $\mathbf{T}$  are independent.

Define  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  as follows: First,  $\tau_1 = T_{Y_0}$  and  $X_t = Y_0$  for  $0 \leq t < \tau_1$ . Recursively, if  $X_t$  is defined on  $[0, \tau_n)$ , let  $\tau_{n+1} = \tau_n + T_{Y_n}$  and then let  $X_t = Y_n$  for  $\tau_n \leq t < \tau_{n+1}$ . Since  $\lambda$  is bounded,  $\tau_n \rightarrow \infty$  as  $n \rightarrow \infty$ , so  $X_t$  is well defined for  $t \in [0, \infty)$ . By construction,  $t \mapsto X_t$  is right continuous and has left limits. The Markov property holds by the memoryless property of the exponential distribution and the fact that  $\mathbf{Y}$  is a Markov chain. Finally, by construction,  $\mathbf{X}$  has exponential parameter function  $\lambda$  and jump chain  $\mathbf{Y}$ .

Often, particularly when  $S$  is finite, the essential structure of a standard, continuous-time Markov chain can be succinctly summarized with a graph.

Suppose again that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a regular Markov chain on  $S$ , with exponential parameter function  $\lambda$  and embedded transition matrix  $Q$ . The *state graph* of  $\mathbf{X}$  is the graph with vertex set  $S$  and directed edge set  $E = \{(x, y) \in S^2 : Q(x, y) > 0\}$ . The graph is *labeled* as follows:

1. Each vertex  $x \in S$  is labeled with the exponential parameter  $\lambda(x)$ .
2. Each edge  $(x, y) \in E$  is labeled with the transition probability  $Q(x, y)$ .

So except for the labels on the vertices, the state graph of  $\mathbf{X}$  is the same as the state graph of the discrete-time jump chain  $\mathbf{Y}$ . That is, there is a directed edge from state  $x$  to state  $y$  if and only if the chain, when in  $x$ , can move to  $y$  after the random holding time in  $x$ . Note that the only loops in the state graph correspond to absorbing states, and for such a state there are no outward edges.

Let's return again to the [construction above](#) of a continuous-time Markov chain from the jump transition matrix  $Q$  and the exponential parameter function  $\lambda$ . Again for simplicity, assume there are no absorbing states. We assume that  $Q(x, x) = 0$  for all  $x \in S$ , so that the state really does *change* at the transition times. However, if we drop this assumption, the construction still produces a continuous-time Markov chain, but with an altered jump transition matrix and exponential parameter function.

Suppose that  $Q$  is a transition matrix on  $S \times S$  with  $Q(x, x) < 1$  for  $x \in S$ , and that  $\lambda : S \rightarrow (0, \infty)$  is bounded. The stochastic process  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  [constructed above](#) from  $Q$  and  $\lambda$  is a regular, continuous-time Markov chain with exponential parameter function  $\tilde{\lambda}$  and jump transition matrix  $\tilde{Q}$  given by

$$\begin{aligned}\tilde{\lambda}(x) &= \lambda(x)[1 - Q(x, x)], \quad x \in S \\ \tilde{Q}(x, y) &= \frac{Q(x, y)}{1 - Q(x, x)}, \quad (x, y) \in S^2, x \neq y\end{aligned}$$

Proof 1

As before, the fact that  $\mathbf{X}$  is a continuous-time Markov chain follows from the memoryless property of the exponential distribution and the Markov property of the jump chain  $\mathbf{Y}$ . By construction,  $t \mapsto X_t$  is right continuous and has left limits. The main point, however, is that  $(\tau_1, \tau_2, \dots)$  is not necessarily the sequence of transition times, when the state actually changes. So we just need to determine the parameters. Suppose  $X_0 = x \in S$  and let  $\tau = \tau_1$  have the exponential distribution with parameter  $\lambda(x)$ , as in the construction. Let  $T$  denote the time when the state actually does change. For  $t \in [0, \infty)$ , the event  $T > t$  can happen in two ways: either  $\tau > t$  or  $\tau = s$  for some  $s \in [0, t]$ , the chain jumps back into state  $x$  at time  $s$ , and the process then stays in  $x$  for a period of at least  $t - s$ . Thus let  $F_x(t) = \mathbb{P}(T > t \mid X_0 = x)$ . Taking the two cases, conditioning on  $\tau$ , and using the Markov property gives



$$F_x(t) = e^{-\lambda(x)t} + \int_0^t \lambda(x) e^{-\lambda(x)s} Q(x, x) F_x(t-s) ds \quad (16.15.22)$$

Using the change of variables  $u = t - s$  and simplifying gives

$$F_x(t) = e^{-\lambda(x)t} \left[ 1 + \lambda(x) Q(x, x) \int_0^t e^{\lambda(x)u} F_x(u) du \right] \quad (16.15.23)$$

Differentiating with respect to  $t$  then gives

$$F'_x(t) = -\lambda(x)[1 - Q(x, x)]F_x(t) \quad (16.15.24)$$

with the initial condition  $F_x(0) = 1$ . The solution of course is  $F_x(t) = \exp\{-\lambda(x)[1 - Q(x, x)]t\}$  for  $t \in [0, \infty)$ . When the state does change, the new state  $y \neq x$  is chosen with probability

$$\mathbb{P}(Y_1 = y \mid Y_0 = x, Y_1 \neq x) = \frac{Q(x, y)}{1 - Q(x, x)} \quad (16.15.25)$$

Proof 2

As in the first proof, we just need to determine the parameters. Given  $X_0 = Y_0 = x$ , the discrete time  $N$  when  $Y$  first changes state has the geometric distribution on  $\mathbb{N}_+$  with success parameter  $1 - Q(x, x)$ . Hence the time until  $X$  actually changes state has the distribution of  $T = \sum_{i=1}^N U_i$  where  $U = (U_1, U_2, \dots)$  is a sequence of independent variables, each exponentially distributed with parameter  $\lambda(x)$  and with  $U$  independent of  $N$ . In the section on the exponential distribution, it is shown that  $T$  also has the exponential distribution, but with parameter  $\lambda(x)[1 - Q(x, x)]$ . (The proof is simple using generating functions.) As in the first proof, when the state does change, the new state  $y \neq x$  is chosen with probability

$$\mathbb{P}(Y_1 = y \mid Y_0 = x, Y_1 \neq x) = \frac{Q(x, y)}{1 - Q(x, x)} \quad (16.15.26)$$

This construction will be important in our study of chains subordinate to the Poisson process.

### Transition Times

The structure of a regular Markov chain on  $S$ , as described above, can be explained purely in terms of a family of independent, exponentially distributed random variables. The main tools are some additional special properties of the exponential distribution, that we need to restate in the setting of our Markov chain. Our interest is in how the process evolves among the stable states until it enters an absorbing state (if it does). Once in an absorbing state, the chain stays there forever, so the behavior from that point on is trivial.

Suppose that  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  is a regular Markov chain on  $S$ , with exponential parameter function  $\lambda$  and transition probability matrix  $Q$ . Define  $\mu(x, y) = \lambda(x)Q(x, y)$  for  $(x, y) \in S^2$ . Then

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

The main point is that the new parameters  $\mu(x, y)$  for  $(x, y) \in S^2$  determine the exponential parameters  $\lambda(x)$  for  $x \in S$ , and the transition probabilities  $Q(x, y)$  when  $x \in S$  is stable and  $y \in S$ . Of course we know that if  $\lambda(x) = 0$ , so that  $x$  is absorbing, then  $Q(x, x) = 1$ . So in fact, the new parameters, as specified by the function  $\mu$ , completely determine the old parameters, as specified by the functions  $\lambda$  and  $Q$ . But so what?

Consider the functions  $\mu$ ,  $\lambda$ , and  $Q$  as given in the previous result. Suppose that  $T_{x,y}$  has the exponential distribution with parameter  $\mu(x, y)$  for each  $(x, y) \in S^2$  and that  $\{T_{x,y} : (x, y) \in S^2\}$  is a set of independent random variables. Then

1.  $T_x = \inf\{T_{x,y} : y \in S\}$  has the exponential distribution with parameter  $\lambda(x)$  for  $x \in S$ .
2.  $\mathbb{P}(T_x = T_{x,y}) = Q(x, y)$  for  $(x, y) \in S^2$ .

Proof

These are basic results proved in the section on the exponential distribution.

So here's how we can think of a regular, continuous-time Markov chain on  $S$ : There is a *timer* associated with each  $(x, y) \in S^2$ , set to the random time  $T_{x,y}$ . All of the timers function independently. When the chain enters state  $x \in S$ , the timers on  $(x, y)$  for  $y \in S$  are started simultaneously. As soon as the first alarm goes off for a particular  $(x, y)$ , the chain immediately moves to state  $y$ , and the process repeats. Of course, if  $\mu(x, y) = 0$  then  $T_{x,y} = \infty$  with probability 1, so only the timers with  $\lambda(x) > 0$  and  $Q(x, y) > 0$  matter (these correspond to the non-loop edges in the state graph). In particular, if  $x$  is absorbing, then the timers on  $(x, y)$  are set to infinity for each  $y$ , and no alarm ever sounds.

The new collection of exponential parameters can be used to give an alternate version of the state graph. Again, the vertex set is  $S$  and the edge set is  $E = \{(x, y) \in S^2 : Q(x, y) > 0\}$ . But now each edge  $(x, y)$  is labeled with the exponential rate parameter  $\mu(x, y)$ . The exponential rate

parameters are closely related to the generator matrix, a matrix of fundamental importance that we will study in the next section.

## Examples and Exercises

### The Two-State Chain

The two-state chain is the simplest non-trivial, continuous-time Markov chain, but yet this chain illustrates many of the important properties of general continuous-time chains. So consider the Markov chain  $\mathbf{X} = \{X_t : t \in [0, \infty)\}$  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.

The transition matrix  $Q$  for the embedded chain is given below. Draw the state graph in each case.

1.  $Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  if  $a > 0$  and  $b > 0$ , so that both states are stable.
2.  $Q = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$  if  $a = 0$  and  $b > 0$ , so that  $a$  is absorbing and  $b$  is stable.
3.  $Q = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$  if  $a > 0$  and  $b = 0$ , so that  $a$  is stable and  $b$  is absorbing.
4.  $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  if  $a = 0$  and  $b = 0$ , so that both states are absorbing.

We will return to the two-state chain in subsequent sections.

### 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

$$Q = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 \end{bmatrix} \quad (16.15.27)$$

1. Draw the state graph and classify the states.
2. Find the matrix of transition rates.
3. Classify the jump chain in terms of recurrence and period.
4. Find the invariant distribution of the jump chain.

Answer

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

$$\begin{bmatrix} 0 & 2 & 2 \\ 1 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix} \quad (16.15.28)$$

3. The jump chain is irreducible, positive recurrent, and aperiodic.
4. The invariant distribution for the jump chain has PDF

$$f = \left[ \frac{6}{14} \quad \frac{5}{14} \quad \frac{3}{14} \right] \quad (16.15.29)$$

### Special Models

Read the introduction to chains subordinate to the Poisson process.

Read the introduction to birth-death chains.

Read the introduction to continuous-time queuing chains.

Read the introduction to continuous-time branching chains.

This page titled [16.15: Introduction to Continuous-Time Markov 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.