MATH 447 - Worked Examples 7 - Continuous-Time Markov Chains

Tasmin Chu

March 11 2023

Abstract

Hi everyone! The point of this set of worked examples is to deepen your understanding of continuous-time Markov chains, based on what you've learned so far. If you notice any typos or errors, or have any questions or feedback, feel free to email me at tasmin.chu@mail.mcgill.ca.

1 Review

Definition 1.1 (Markov Property)**.** *A continuous-time stochastic process $(X_t)_{t\geq 0}$ with discrete state space S is a* ***continuous-time Markov chain*** *if*

$$P(X_{t+s} = j | X_s = i, X_u = x_u, 0 \leq u < s) = P(X_{t+s} = j | X_s = i)$$

for all $s, t \geq 0, i, j, x_u \in S$ and $0 \leq u < s$. We call the property above the Markov property.

The definition above can seem somewhat unwieldly. But it really doesn't say anything that much different than what we saw previously for Markov chains indexed by $\mathbb{N}$. Basically, this generalizes the same idea of memorylessness to the case where $t \geq 0$ takes values in the continuum.

Example 1.1.1. Suppose we are given information about the values $(X_t)_{t\geq 0}$ takes at $t = 1.2$ and $t = 3$; in particular, $X_{1.2} = a, X_3 = b$. We are interested in the value X_t takes when $t = 12.5$. Let the state space $S = \mathbb{N}$. Then for any $k \in \mathbb{N}$

$$P(X_{12.5} = k | X_3 = b, X_{1.2} = a) = P(X_{12.5} = j | X_3 = b)$$

That is, given the entire history of the process, we can reduce it to the information given by the most recent state.

Definition 1.2. *A continuous-time stochastic process $(X_t)_{t\geq 0}$ with discrete state space S is a* ***time-homogenous continuous-time Markov chain*** *if it is a continuous-time Markov chain and*

$$P(X_{t+s} = j | X_s = i) = P(X_t = j | X_0 = i)$$

for all $s \geq 0$.

Essentially all of the material below deals with time-homogenous continuous-time Markov chains.

Example 1.2.1 (Holding-Time Construction)**.** How can we create a continuous-time Markov chain given a stochastic matrix whose diagonal entries are zero? Suppose we are given an $n \times n$ stochastic matrix $\tilde{P}$ which model transitions between states $\{1, \cdots, n\}$. Suppose also that $\tilde{P}_{ii} = 0$ for all states i, so we have zeroes on the diagonal. Take a stochastic row vector $\lambda = (\lambda_1, \cdots, \lambda_n)$. (That is, all the entries in the row vector add up to one.) Then we can construct a continuous-time Markov chain like so. Let the chain be at some state i. Then the chain stays there for some time t, where t is exponentially distributed with parameter λ_i. Once the alarm clock rings, the chain changes to one of the other states $\{1, \cdots, i-1, i+1, \cdots, n\}$ with probability given by our transition matrix $\tilde{P}_{i,j}$. Suppose it transitions to state j. Then the chain stays there for some time t' where t' is exponentially distributed with parameter λ_j, and so on.

Then $(X_t)_{t\geq 0}$, where X_t denotes the state at time t, is a continuous-time Markov chain.

For a concrete example, take the weather example. Here a stochastic matrix could be

$$\tilde{P}(t) = \begin{array}{c} \\ \text{Rain} \\ \text{Snow} \\ \text{Clear} \end{array} \begin{array}{c} \begin{array}{ccc} 0 & \text{Rain} & \text{Snow} \quad \text{Clear} \end{array} \\ \begin{pmatrix} 0 & 1/2 & 1/2 \\ 3/4 & 0 & 1/4 \\ 1/4 & 3/4 & 0 \end{pmatrix} \end{array}$$

Let's also suppose that rainfall lasts, on average, 3 hours at a time; snowfall lasts about 6 hours; and the weather stays clear for 12 hours on average. Then we can take our row vector to be

$$\lambda = (\lambda_r, \lambda_s, \lambda_c) = (1/3, 1/6, 1/12)$$

(Confused by the reciprocrals? Recall that the expectation of an exponential variable X with parameter β is $E(X) = 1/\beta$.)

Using the construction above, we get a continuous time Markov chain which models the behaviour of the weather.

Definition 1.3 (Transition Function)**.** *For each $t \geq 0$, we can arrange our transition-probabilities in a matrix function $\boldsymbol{P}(t)$ called the **transition function** with*

$$\boldsymbol{P}_{ij}(t) = P(X_t = j | X_0 = i)$$

Theorem 1.1 (Chapman-Kolomogorov)**.** *For a continuous-time Markov chain $(X_t)_{t\geq 0}$ with transition matrix $\boldsymbol{P}(t)$,*

$$\boldsymbol{P}(s+t) = \boldsymbol{P}(s)\boldsymbol{P}(t)$$

for $s, t \geq 0$. More concretely,

$$P_{ij}(s+t) = [P(s)P(t)]_{ij} = \sum_k P_{ik}(s)P_{kj}(t)$$

for states i, j and $s, t \geq 0$.

Theorem 1.2 (Holding Times are Exponentially Distributed)**.** *Let $(X_t)_{t\geq 0}$ be a time-homogeneous Markov chain. Let T_i be the **holding time** at state i, that is, the length of time that a continuous-time Markov chain started in i stays at i before transition to a new state. Then T_i has an exponential distribution.*

Now, for each state i, let q_i be the parameter of the exponential distribution for the holding time T_i. We assume that $0 < q_i < \infty$, but in fact we can define a continuous-time Markov chain where $q_i = 0$ or $q_i = \infty$.

Case 1: $q_i = 0$ for some i. In this case, when the Markov chain visits i, it never leaves, and i is called an absorbing state.

Case 2: $q_i = \infty$ for some i. In this case, when the Markov chain visits i, it leaves i immediately after entering i, which means that we can have infinitely many transitions in a finite interval. A continuous-time Markov chain which has

a positive probability of making infinitely many transitions in finite time is called **explosive**.

Case 3 $0 < q_i < \infty$, for all i.

Suppose $0 < q_i < \infty$ for all states i. Then starting from i, the process stays in i for an exponentially distributed length of time, on average a time of $1/q_i$. Then it hits a new state $j, j \neq i$ with probability p_{ij} and stays there for an exponentially distributed length of time. And so on.

Notice this is identical to Example 1.2.1! (The only difference is the state space might be infinite, for a generic continuous-time Markov chain. In the example there were only n states.)

These transition probabilities p_{ij} describe the discrete transitions from state to state. If we ignore time, and just watch the state-to-state transitions, we get a sequence $Y_0, Y_1, \cdots, Y_n$, where Y_n is the nth state visited. This gives us a discrete-time Markov chain, called the **embedded chain**.

Let $\tilde{P}$ be the transition matrix of this embedded chain. I.e., $\tilde{P}_{ij} = p_{ij}$. Then this is a stochastic matrix whose diagonal entries are zero.

The following way of thinking about non-explosive, non-absorbing continuous-time time-homogeneous Markov chains is probably the most fundamental way of understanding them.

Alarm Clocks and Transition Rates Imagine that for each state i, we set an independent alarm clock for each of the states that the process can visit after i. In particular, the alarm clock associated with (i, j) rings after an exponentially distributed length of time with parameter q_{ij}. After the chain transitions to state i, we start all the clocks simultaneously. The first one to ring determines the next state to visit. The q_{ij} are called the transition rates of the continuous-time Markov chain.

How long does the chain stay at state i? The time of the first alarm is equal to the minimum of all the exponential variables with parameters q_{ik}, for $k \in S$. In particular, this minimum is exponentially distributed with parameter $\sum_k q_{ik}$.

Thus,

$$q_i = \sum_k q_{ik}$$

Moreover, we can calculate the transition probabilities of the embedded chain.

From i, the chain moves to j if the (i, j) clock rings first, which occurs with probability

$$p_{ij} = \frac{q_{ij}}{\sum_k q_{ik}} = \frac{q_{ij}}{q_i}$$

Definition 1.4 (Generator Matrix)**.** *Let $(X_t)_{t\geq 0}$ be a continuous-time Markov chain with transition function $\boldsymbol{P}(t)$. Assume that the transition function is differentiable. If $X_t = i$, then the instantaneous transition rate of hitting some other state j, for $j \neq i$, is*

$$\begin{aligned}
&\lim_{h\to 0^+} \frac{E(\textit{Number of transitions in } (t, t+h] \textit{ to } j | X_t = i)}{h} \\
&= \lim_{h\to 0^+} \frac{P(X_{t+h} = j | X_t = i)}{h} \\
&= \lim_{h\to 0^+} \frac{P(X_h = j | X_0 = i)}{h} \textit{ by time homogeneity} \\
&= \lim_{h\to 0^+} \frac{P_{ij}(h)}{h} \\
&= P'_{ij}(0)
\end{aligned}$$

This motivates the following definition. Let $\boldsymbol{Q} = \boldsymbol{P}'(0)$. This is a matrix whose non-diagonal entries are the instantaneous transition rates, which are the transition rates q_{ij} introduced in the example above. The diagonal entries of $\boldsymbol{Q}$ are

$$Q_{ii} = P'_{ii}(0) = -\sum_{j\neq i} q_{ij} = -q_i$$

$\boldsymbol{Q}$ is called the generator matrix or infinitesimal generator. Notice that it's not a stochastic matrix–some of its entries are negative, and rows sum to 0.

We now turn to the problems. Try working them out on your own, then compare your answers with the solution on the next page.

2 Problems

Question 1

Consider the following figure, depicting a continuous-time Markov chain $(X_t)_{t\geq 0}$ with states $\{1, 2, 3\}$. Here, q_{ij} describes the transition rates.

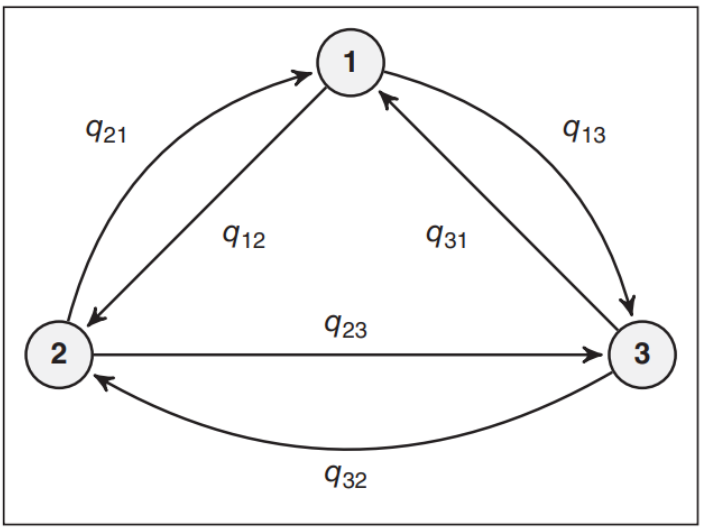


Figure 7.3 Transition rates for three-state chain.

a) Determine the holding time parameters of the continuous-time Markov chain.

b) Write down the transition matrix $\tilde{P}$ of the embedded chain.

Question 2

Understanding the (1-dimensional) Poisson process deeply is a really good way of understanding continuous-time Markov chains. With that in mind, consider the following questions.

Below, let $(N_t)_{t\geq 0}$ be a 1-dimensional Poisson process with parameter λ.

a) Calculate the transition function $\mathbf{P}_{ij}(t) = P(X_t = j | X_0 = i)$ of the Poisson process, considering carefully the cases where $j \geq i$ and $j < i$. (Hint: you may have seen this in class.) Write down the resulting matrix.

b) Using your results from part a), write the matrix **P**(t), the transition function.

c) Write down the matrix $\tilde{P}$ of the embedded chain of the Markov chain.

d) Write down the matrix **Q**(t), the infinitesimal generator of the Poisson process.

Solutions

Question 1

Consider the following figure, depicting a continuous-time Markov chain $(X_t)_{t\geq 0}$ with states $\{1, 2, 3\}$. Here, q_{ij} describes the transition rates.

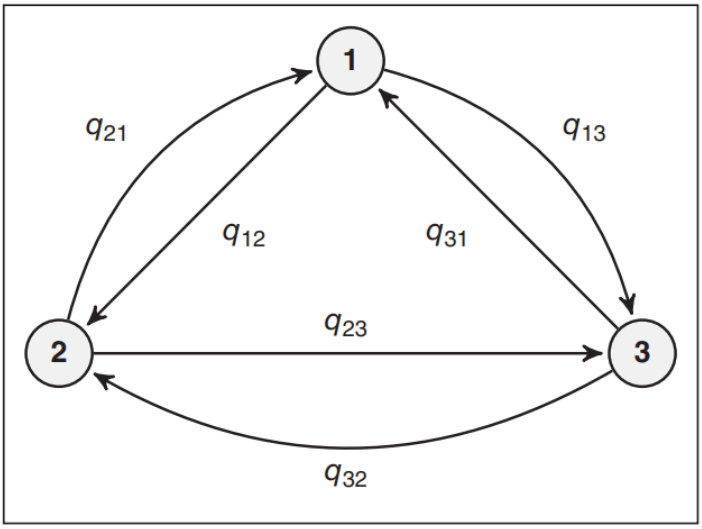


Figure 7.3 Transition rates for three-state chain.

a) Determine the holding time parameters of the continuous-time Markov chain.

b) Write down the transition matrix $\tilde{P}$ of the embedded chain.

Solution

Part a)

Recall that the holding time parameters q_i are equal to

$$q_i = \sum_k q_{ik}$$

From state 1, the chain can transition to state 2 at rate q_{12} and to state 3 at rate q_{13}. Thus,

$$q_1 = q_{12} + q_{13}$$

Using the same logic for the other two states, we get that

$$(q_1, q_2, q_3) = (q_{12} + q_{13}, q_{21} + q_{23}, q_{31} + q_{32})$$

Part b)

Recall that $p_{ij} = \dfrac{q_{ij}}{q_i}$. Thus, the transition matrix $\tilde{P}$ of the embedded chain is

$$\tilde{P} = \begin{pmatrix} 0 & q_{12}/q_1 & q_{13}/q_1 \\ q_{21}/q_2 & 0 & q_{23}/q_2 \\ q_{31}/q_3 & q_{32}/q_3 & 0 \end{pmatrix}$$

Question 2

Understanding the (1-dimensional) Poisson process deeply is a really good way of understanding continuous-time Markov chains. With that in mind, consider the following questions.

Below, let $(N_t)_{t\geq 0}$ be a 1-dimensional Poisson process with parameter λ.

a) Calculate the transition function $\mathbf{P}_{ij}(t) = P(X_t = j | X_0 = i)$ of the Poisson process, considering carefully the cases where $j \geq i$ and $j < i$. (Hint: you may have seen this in class.) Write down the resulting matrix.

b) Using your results from part a), write the matrix **P**(t), the transition function.

c) Write down the matrix $\tilde{P}$ of the embedded chain of the Markov chain.

d) Write down the matrix **Q**(t), the infinitesimal generator of the Poisson process.

Solution

Part a)

Notice if $j < i$, then $P_{ij}(t) = P(N_{t+s} = j | N_s = i) = 0$.

So assume $j \geq i$. Then

$$P_{ij}(t) = P(N_{t+s} = j | N_s = i)$$

$$
\begin{aligned}
&= \frac{P(N_{t+s} = j, N_s = i)}{P(N_s = i)} \\
&= \frac{P(N_{t+s} - N_s = j - i, N_s = i)}{P(N_s = i)} \\
&= P(N_{t+s} - N_s = j - i) \\
&= P(N_t = j - i) \\
&= \frac{e^{-\lambda t}(\lambda t)^{j-i}}{(j-i)!}
\end{aligned}
$$

Part b)

It follows from part $a)$ that

$$
\mathbf{P}(t) = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{array}
\begin{array}{c}
\begin{array}{ccccc} 0 & 1 & 2 & 3 & \cdots \end{array} \\
\begin{pmatrix}
e^{-\lambda t} & (\lambda t)e^{-\lambda t} & (\lambda t)^2 e^{-\lambda t}/2 & (\lambda t)^3 e^{-\lambda t}/6 & \cdots \\
0 & e^{-\lambda t} & (\lambda t)e^{-\lambda t} & (\lambda t)^2 e^{-\lambda t}/2 & \cdots \\
0 & 0 & e^{-\lambda t} & (\lambda t)e^{-\lambda t} & \cdots \\
0 & 0 & 0 & e^{-\lambda t} & \cdots \\
0 & 0 & 0 & 0 & \ddots
\end{pmatrix}
\end{array}
$$

Part c)

Recall that $(N_t)_{t\geq 0}$ is a counting process, that is, N_t increments up by 1. So if we don't consider time but just consider the states of $(N_t)_{t\geq 0}$, it's clear that N_t moves from 0 to 1 to 2, etc. It just counts upwards.

Thus, the transition matrix of the embedded chain is extremely simple:

$$
\tilde{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{array}
\begin{array}{c}
\begin{array}{ccccc} 0 & 1 & 2 & 3 & \cdots \end{array} \\
\begin{pmatrix}
0 & 1 & 0 & 0 & \cdots \\
0 & 0 & 1 & 0 & \cdots \\
0 & 0 & 0 & 1 & \cdots \\
0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & \ddots
\end{pmatrix}
\end{array}
$$

Another way of putting this is that, ignoring time, the probability of transitioning from state i to state $i+1$ is 1 and 0 otherwise.

Part d)

Recall that the interarrival times of a Poisson process are exponentially distributed with parameter λ. Thus, the holding time T_i – the time we expect that $(N_t)_{t\geq 0}$ will stay at a state i – is exponentially distributed with holding time parameter $q_i = \lambda$, for any i.

Since

$$p_{ij} = \frac{q_{ij}}{q_i}$$

we can rearrange this to find that

$$q_{ij} = q_i \cdot p_{ij} = \lambda \cdot p_{ij}$$

where

$$p_{ij} = \begin{cases} 1 & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

This gives us all the non-diagonal entries. Recall that the diagonal entries are equal to $-q_i$. Since we saw $q_i = \lambda$, all the diagonal entries of $\mathbf{Q}(t)$ are equal to $-\lambda$. Thus,

$$\mathbf{Q}(t) = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & \cdots \end{array} \\ \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \cdots \\ 0 & 0 & 0 & -\lambda & \cdots \\ 0 & 0 & 0 & 0 & \ddots \end{pmatrix} \end{array}$$