Chapter 5. Continuous-Time Markov Chains

Prof. Shun-Ren Yang
Department of Computer Science,
National Tsing Hua University, Taiwan

Continuous-Time Markov Chains

- Consider a continuous-time stochastic process $\{X(t), t \geq 0\}$ taking on values in the set of nonnegative integers. We say that the process $\{X(t), t \geq 0\}$ is a *continuous-time Markov chain* if for all $s, t \geq 0$, and nonnegative integers $i, j, x(u), 0 \leq u \leq s$,

$$\begin{aligned} P\{X(t+s) = j | X(s) = i, X(u) = x(u), 0 \leq u < s\} \\ = P\{X(t+s) = j | X(s) = i\}. \end{aligned}$$

- In other words, a continuous-time Markov chain is a stochastic process having the Markovian property that the conditional distribution of the future state at time $t+s$, given the present state at s and all past states depends only on the present state and is independent of the past.
- Suppose that a continuous-time Markov chain enters state i at some time, say time 0, and suppose that the process does not leave state i

(that is, a transition does not occur) during the next s time units. What is the probability that the process will not leave state i during the following t time units?

- Note that as the process is in state i at time s, it follows, by the Markovian property, that the probability it remains in that state during the interval $[s, s+t]$ is just the (unconditional) probability that it stays in state i for at least t time units. That is, if we let τ_i denote the amount of time that the process stays in state i before making a transition into a different state, then

$$P\{\tau_i > s+t | \tau_i > s\} = P\{\tau_i > t\}$$

for all $s, t \geq 0$. Hence, the random variable τ_i is memoryless and must thus be exponentially distributed.

Continuous-Time Markov Chains

- In fact, the above gives us a way of constructing a continuous-time Markov chain. Namely, it is a stochastic process having the properties that each time it enters state i :
  1. the amount of time it spends in that state before making a transition into a different state is exponentially distributed with rate, say, v_i; and
  2. when the process leaves state i, it will next enter state j with some probability, call it P_{ij}, where $\sum_{j \neq i} P_{ij} = 1$.
- A state i for which $v_i = \infty$ is called an *instantaneous* state since when entered it is instantaneously left. If $v_i = 0$, then state i is called *absorbing* since once entered it is never left.
- Hence, a continuous-time Markov chain is a stochastic process that moves from state to state in accordance with a (discrete-time) Markov

chain, but is such that the amount of time it spends in each state, before proceeding to the next state, is exponentially distributed.

- In addition, the amount of time the process spends in state i, and the next state visited, must be independent random variables. For if the next state visited were dependent on τ_i, then information as to how long the process has already been in state i would be relevant to the prediction of the next state — and this would contradict the Markovian assumption.
- Let q_{ij} be defined by

$$q_{ij} = v_i P_{ij}, \qquad \text{all } i \neq j.$$

Since v_i is the rate at which the process leaves state i and P_{ij} is the probability that it then goes to j, it follows that q_{ij} is the rate when in state i that the process makes a transition into state j; and in fact we

Continuous-Time Markov Chains

call q_{ij} the *transition rate* from i to j.

- Let us denote by $P_{ij}(t)$ the probability that a Markov chain, presently in state i, will be in state j after an additional time t. That is,

$$P_{ij}(t) = P\{X(t+s) = j | X(s) = i\}.$$

Birth and Death Processes

- A continuous-time Markov chain with states $0, 1, \ldots$ for which $q_{ij} = 0$ whenever $|i - j| > 1$ is called a *birth and death process*.
- Thus a birth and death process is a continuous-time Markov chain with states $0, 1, \ldots$ for which transitions from state i can only go to either state $i - 1$ or state $i + 1$. The state of the process is usually thought of as representing the size of some population, and when the state increases by 1 we say that a birth occurs, and when it decreases by 1 we say that a death occurs.
- Let λ_i and μ_i be given by

$$\lambda_i = q_{i,i+1},$$

$$\mu_i = q_{i,i-1}.$$

The values $\{\lambda_i, i \geq 0\}$ and $\{\mu_i, i \geq 1\}$ are called respectively the birth rates and the death rates.

Birth and Death Processes

- Since $\sum_j q_{ij} = v_i$, we see that

$$\begin{aligned} v_i &= \lambda_i + \mu_i, \\ P_{i,i+1} &= \frac{\lambda_i}{\lambda_i + \mu_i} = 1 - P_{i,i-1}. \end{aligned}$$

- Hence, we can think of a birth and death process by supposing that whenever there are i people in the system the time until the next birth is exponential with rate λ_i and is independent of the time until the next death, which is exponential with rate μ_i.

An Example: *The $M/M/s$ Queue*

- Suppose that customers arrive at an s-server service station in accordance with a Poisson process having rate λ.
- Each customer, upon arrival, goes directly into service if any of the servers are free, and if not, then the customer joins the queue (that is, he waits in line).
- When a server finishes serving a customer, the customer leaves the system, and the next customer in line, if there are any waiting, enters the service. The successive service times are assumed to be independent exponential random variables having mean $1/\mu$.
- If we let $X(t)$ denote the number in the system at time t, then

An Example: *The M/M/s Queue*

$\{X(t), t \geq 0\}$ is a birth and death process with

$$
\begin{aligned}
\mu_n &= \begin{cases} n\mu & 1 \leq n \leq s \\ s\mu & n > s, \end{cases} \\
\lambda_n &= \lambda, \qquad n \geq 0.
\end{aligned}
$$

The Kolmogorov Differential Equations

- Recall that
$$P_{ij}(t) = P\{X(t+s) = j | X(s) = i\}$$
represents the probability that a process presently in state i will be in state j a time t later.

- By exploiting the Markovian property, we will derive two sets of differential equations for $P_{ij}(t)$, which may sometimes be explicitly solved. However, before doing so we need the following lemmas.

- **Lemma 1.**

  1. $\displaystyle\lim_{t\to 0} \frac{1 - P_{ii}(t)}{t} = v_i.$
  2. $\displaystyle\lim_{t\to 0} \frac{P_{ij}(t)}{t} = q_{ij}, \qquad i \neq j.$

The Kolmogorov Differential Equations

- **Lemma 2.** For all s, t,

$$P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t)P_{kj}(s).$$

- From Lemma 2 we obtain

$$P_{ij}(t+h) = \sum_{k} P_{ik}(h)P_{kj}(t),$$

or, equivalently,

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

Dividing by h and then taking the limit as $h \to 0$ yields, upon

The Kolmogorov Differential Equations

application of Lemma 1,

$$\lim_{h\to 0}\frac{P_{ij}(t+h)-P_{ij}(t)}{h}=\lim_{h\to 0}\sum_{k\neq i}\frac{P_{ik}(h)}{h}P_{kj}(t)-v_iP_{ij}(t).$$

- Assuming that we can interchange the limit and summation on the right-hand side of the above equation, we thus obtain, again using Lemma 1, the following.
- **Theorem.** (Kolmogorov's Backward Equations)

  For all i, j, and $t \geq 0$,

$$P'_{ij}(t)=\sum_{k\neq i} q_{ik}P_{kj}(t)-v_iP_{ij}(t).$$

The Kolmogorov Differential Equations

- The set of differential equations for $P_{ij}(t)$ given in the above Theorem are known as the Kolmogorov *backward equations*.
- They are called the backward equations because in computing the probability distribution of the state at time $t+h$ we conditioned on the state (all the way) back at time h. That is, we started our calculation with

$$\begin{aligned} P_{ij}(t+h) &= \sum_k P\{X(t+h)=j|X(0)=i, X(h)=k\} \\ &\qquad\qquad \times P\{X(h)=k|X(0)=i\} \\ &= \sum_k P_{kj}(t)P_{ik}(h). \end{aligned}$$

The Kolmogorov Differential Equations

- We may derive another set of equations, known as the Kolmogorov's *forward equations*, by now conditioning on the state at time t. This yields

$$P_{ij}(t+h) = \sum_k P_{ik}(t)P_{kj}(h)$$

or

$$\begin{aligned} P_{ij}(t+h) - P_{ij}(t) &= \sum_k P_{ik}(t)P_{kj}(h) - P_{ij}(t) \\ &= \sum_{k \neq j} P_{ik}(t)P_{kj}(h) - [1 - P_{jj}(h)]P_{ij}(t). \end{aligned}$$

Therefore,

$$\lim_{h \to 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h} = \lim_{h \to 0} \{\sum_{k \neq j} P_{ik}(t)\frac{P_{kj}(h)}{h} - \frac{1 - P_{jj}(h)}{h}P_{ij}(t)\}.$$

The Kolmogorov Differential Equations

- Assuming that we can interchange limit with summation, we obtain by Lemma 1 that

$$P'_{ij}(t) = \sum_{k \neq j} q_{kj} P_{ik}(t) - v_j P_{ij}(t).$$

- **Theorem.** (Kolmogorov's Forward Equations)

  Under suitable regularity conditions,

$$P'_{ij}(t) = \sum_{k \neq j} q_{kj} P_{ik}(t) - v_j P_{ij}(t).$$

The Kolmogorov Differential Equations

Example. The Two-State Chain. Consider a two-state continuous-time Markov chain that spends an exponential time with rate λ in state 0 before going to state 1, where it spends an exponential time with rate μ before returning to state 0. The forward equations yield

$$\begin{aligned} P'_{00}(t) &= \mu P_{01}(t) - \lambda P_{00}(t) \\ &= -(\lambda+\mu)P_{00}(t) + \mu, \end{aligned}$$

where the last equation follows from $P_{01}(t) = 1 - P_{00}(t)$. Hence,

$$e^{(\lambda+\mu)t}[P'_{00}(t) + (\lambda+\mu)P_{00}(t)] = \mu e^{(\lambda+\mu)t}$$

or

$$\frac{d}{dt}[e^{(\lambda+\mu)t}P_{00}(t)] = \mu e^{(\lambda+\mu)t}.$$

The Kolmogorov Differential Equations

Thus,

$$e^{(\lambda+\mu)t}P_{00}(t) = \frac{\mu}{\lambda+\mu}e^{(\lambda+\mu)t} + c.$$

Since $P_{00}(0) = 1$, we see that $c = \lambda/(\lambda+\mu)$, and thus

$$P_{00}(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}.$$

Similarly (or by symmetry),

$$P_{11}(t) = \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}e^{-(\lambda+\mu)t}.$$

Limiting Probabilities

- Since a continuous-time Markov chain is a semi-Markov process with

$$F_{ij}(t) = 1 - e^{-v_i t}$$

it follows that if the discrete-time Markov chain with transition probabilities P_{ij} is irreducible and positive recurrent, then the limiting probabilities $P_j = \lim_{t\to\infty} P_{ij}(t)$ are given by

$$P_j = \frac{\pi_j / v_j}{\sum_i \pi_i / v_i}$$

where the π_j are the unique nonnegative solution of

$$\begin{aligned} \pi_j &= \sum_i \pi_i P_{ij}, \\ \sum_i \pi_i &= 1. \end{aligned}$$

Limiting Probabilities

- From the above two equations, we see that the P_j are the unique nonnegative solution of

$$\begin{aligned} v_j P_j &= \sum_i v_i P_i P_{ij}, \\ \sum_j P_j &= 1, \end{aligned}$$

or, equivalently, using $q_{ij} = v_i P_{ij}$,

$$\begin{aligned} v_j P_j &= \sum_i P_i q_{ij}, \\ \sum_j P_j &= 1. \end{aligned} \tag{1}$$

Limiting Probabilities – Remarks

1. It follows from the results for semi-Markov processes that P_j also equals the long-run proportion of time the process is in state j.

2. If the initial state is chosen according to the limiting probabilities $\{P_j\}$, then the resultant process will be stationary. That is,

$$\sum_i P_i P_{ij}(t) = P_j \qquad \text{for all } t.$$

The above is proven as follows:

$$\begin{aligned}
\sum_i P_{ij}(t) P_i &= \sum_i P_{ij}(t) \lim_{s\to\infty} P_{ki}(s) \\
&= \lim_{s\to\infty} \sum_i P_{ij}(t) P_{ki}(s) \\
&= \lim_{s\to\infty} P_{kj}(t+s) \\
&= P_j.
\end{aligned}$$

Limiting Probabilities – Remarks

3. Another way of obtaining Equations (**??**) is by way of the forward equations

$$P'_{ij}(t) = \sum_{k \neq j} q_{kj} P_{ik}(t) - v_j P_{ij}(t).$$

If we assume that the limiting probabilities $P_j = \lim_{t \to \infty} P_{ij}(t)$ exist, then $P'_{ij}(t)$ would necessarily converge to 0 as $t \to \infty$. (Why?) Hence, assuming that we can interchange limit and summation in the above, we obtain upon letting $t \to \infty$,

$$0 = \sum_{k \neq j} P_k q_{kj} - v_j P_j.$$

It is worth noting that the above is a more formal version of the following heuristic argument — which yields an equation for P_j, the probability of being in state j at $t = \infty$ — by conditioning on the

Limiting Probabilities – Remarks

state h units prior in time:

$$\begin{aligned} P_j &= \sum_i P_{ij}(h)P_i \\ &= \sum_{i\neq j}(q_{ij}h + o(h))P_i + (1 - v_jh + o(h))P_j \end{aligned}$$

or

$$0 = \sum_{i\neq j} P_iq_{ij} - v_jP_j + \frac{o(h)}{h},$$

and the result follows by letting $h \to 0$.

4. Equation (**??**) has a nice interpretation, which is as follows:
   - In any interval $(0, t)$, the number of transitions into state j must equal to within 1 the number of transitions out of state j. (Why?) Hence, in the long run the rate at which transitions into state j occur must equal the rate at which transitions out of state j occur.

Limiting Probabilities – Remarks

- Now when the process is in state j it leaves at rate v_j, and, since P_j is the proportion of time it is in state j, it thus follows that

$$v_j P_j = \text{rate at which the process leaves state } j.$$

- Similarly, when the process is in state i it departs to j at rate q_{ij}, and, since P_i is the proportion of time in state i, we see that the rate at which transitions from i to j occur is equal to $q_{ij} P_i$. Hence,

$$\sum_i P_i q_{ij} = \text{rate at which the process enters state } j.$$

- Therefore, (**??**) is just a statement of the equality of the rate at which the process enters and leaves state j. Because it balances (that is, equates) these rates, Equations (**??**) are sometimes referred to as *balance equations*.

Limiting Probabilities – Remarks

5. When the continuous-time Markov chain is irreducible and $P_j > 0$ for all j, we say that the chain is *ergodic*.

Limiting Probabilities – An Example

- Let us now determine the limiting probabilities for a birth and death process.
- From Equations (**??**), or, equivalently, by equating the rate at which the process leaves a state with the rate at which it enters that state, we obtain

State	Rate Process Leaves		Rate Process Enters
0	$\lambda_0 P_0$	$=$	$\mu_1 P_1$
$n, n > 0$	$(\lambda_n + \mu_n) P_n$	$=$	$\mu_{n+1} P_{n+1} + \lambda_{n-1} P_{n-1}$

Rewriting these equations gives

$$\begin{aligned} \lambda_0 P_0 &= \mu_1 P_1, \\ \lambda_n P_n &= \mu_{n+1} P_{n+1} + (\lambda_{n-1} P_{n-1} - \mu_n P_n), \qquad n \geq 1, \end{aligned}$$

Limiting Probabilities – An Example

or, equivalently,

$$
\begin{aligned}
\lambda_0 P_0 &= \mu_1 P_1, \\
\lambda_1 P_1 &= \mu_2 P_2 + (\lambda_0 P_0 - \mu_1 P_1) = \mu_2 P_2, \\
\lambda_2 P_2 &= \mu_3 P_3 + (\lambda_1 P_1 - \mu_2 P_2) = \mu_3 P_3, \\
\lambda_n P_n &= \mu_{n+1} P_{n+1} + (\lambda_{n-1} P_{n-1} - \mu_n P_n) = \mu_{n+1} P_{n+1}.
\end{aligned}
$$

Solving in terms of P_0 yields

$$
\begin{aligned}
P_1 &= \frac{\lambda_0}{\mu_1} P_0, \\
P_2 &= \frac{\lambda_1}{\mu_2} P_1 = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} P_0, \\
P_3 &= \frac{\lambda_2}{\mu_3} P_2 = \frac{\lambda_2 \lambda_1 \lambda_0}{\mu_3 \mu_2 \mu_1} P_0.
\end{aligned}
$$

Limiting Probabilities – An Example

$$P_n \quad = \quad \frac{\lambda_{n-1}}{\mu_n} P_{n-1} = \frac{\lambda_{n-1}\lambda_{n-2}\cdots\lambda_1\lambda_0}{\mu_n\mu_{n-1}\cdots\mu_2\mu_1} P_0.$$

Using $\sum_{n=0}^{\infty} P_n = 1$ we obtain

$$1 = P_0 + P_0 \sum_{n=1}^{\infty} \frac{\lambda_{n-1}\cdots\lambda_1\lambda_0}{\mu_n\cdots\mu_2\mu_1}$$

or

$$P_0 = [1 + \sum_{n=1}^{\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n}]^{-1},$$

and hence

$$P_n = \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n(1 + \sum_{n=1}^{\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n})}, \qquad n \geq 1.$$

Limiting Probabilities – An Example

- The above equations also show us what condition is needed for the limiting probabilities to exist. Namely,

$$\sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} < \infty.$$