STAT253/317 Lecture 12

Cong Ma

Chapter 6 Continuous-Time Markov Chains

6.2 Continuous-Time Markov Chains (CTMC)

Definitions. A stochastic process $\{X(t), t \geq 0\}$ with state space $\mathcal{X}$ is called a *continuous-time Markov chain* if for any two states i, $j \in \mathcal{X}$,

$$\begin{aligned}
&\mathrm{P}(\underbrace{X(t+s) = j}_{\text{future}} \,|\, \underbrace{X(s) = i}_{\text{present}}, \underbrace{X(u) = x(u), \text{for } 0 \leq u < s}_{\text{past}}) \\
&= \mathrm{P}(\underbrace{X(t+s) = j}_{\text{future}} \,|\, \underbrace{X(s) = i}_{\text{present}})
\end{aligned}$$

If $\mathrm{P}(X(t+s) = j | X(s) = i)$ does not depend on s for all $i, j \in \mathcal{X}$, then it is denoted as

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

and we say the CTMC is *homogeneous* in time.

In STAT253/317, we focus on homogeneous CTMC only.

Exponential Waiting Time

Let $\{X(t), t \geq 0\}$ be a homogeneous continuous-time Markov chain. For $i \in \mathcal{X}$, let T_i denote the amount of time that $X(t)$ stays in state i before making a transition into a different state.

Claim: T_i has the *memoryless property*.

$$
\begin{aligned}
&\mathrm{P}(T_i \geq t+s | T_i \geq s) \\
&= \mathrm{P}(X(u) = i, \text{ for } s \leq u \leq s+t | X(u) = i, \text{ for } 0 \leq u \leq s) \\
&= \mathrm{P}(X(u) = i, \text{ for } s \leq u \leq s+t | X(s) = i) \quad \text{(Markov property)} \\
&= \mathrm{P}(X(u) = i, \text{ for } 0 \leq u \leq t | X(0) = i) \qquad \text{(time homogeneity)} \\
&= \mathrm{P}(T_i \geq t) \quad \Rightarrow \quad \text{So } T_i \text{ is memoryless.}
\end{aligned}
$$

Recall that the exponential distribution is the only continuous distribution having the memoryless property.
Thus $T_i \sim Exp(\nu_i)$ for some rate ν_i.

An Alternative Definition of CTMC

A stochastic process $\{X(t), t \geq 0\}$ with state space $\mathcal{X}$ is a *continuous-time Markov chain* if

- (exponential waiting time) when the chain reaches a state i, the time it stays at state $i \sim Exp(\nu_i)$, where ν_i is the transition rate at state i
- (embedded with a discrete time Markov chain) when the process leaves state i, it enters anther state j with probability P_{ij}, such that

$$P_{ii} = 0, \quad \sum\nolimits_{j \in \mathcal{X}} P_{ij} = 1 \quad \text{for all } i, j \in \mathcal{X}.$$

Remark: The amount of time T_i the process spends in state i, and the next state visited, must be independent. For if the next state visited were dependent on T_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 contradicts the Markovian assumption.

6.3 Birth and Death Processes

Let $X(t) =$ the number of people in the system at time t.
Suppose that whenever there are n people in the system, then

(i) new arrivals enter the system at an exponential rate λ_n, and

(ii) people leave the system at an exponential rate μ_n.

Such an $\{X(t), t \geq 0\}$ is called a *birth and death process*.

$$0 \underset{\mu_1}{\overset{\lambda_0}{\rightleftarrows}} 1 \underset{\mu_2}{\overset{\lambda_1}{\rightleftarrows}} 2 \underset{\mu_3}{\overset{\lambda_2}{\rightleftarrows}} 3 \underset{\cdots}{\overset{\cdots}{\cdots}} n-1 \underset{\mu_n}{\overset{\lambda_{n-1}}{\rightleftarrows}} n \underset{\mu_{n+1}}{\overset{\lambda_n}{\rightleftarrows}} n+1 \underset{\cdots}{\overset{\cdots}{\cdots}}$$

Suppose the process is at state $i > 0$ at time t. Then

$$B_i = \text{waiting time until the next birth} \sim \text{Exp}(\lambda_i)$$
$$D_i = \text{waiting time until the next death} \sim \text{Exp}(\mu_i)$$

Hence, the waiting time until the next transition out of state i is $\min(B_i, D_i) \sim \text{Exp}(\lambda_i + \mu_i)$, from which we can get

$$\nu_i = \lambda_i + \mu_i, \text{ for } i > 0$$

6.3 Birth and Death Processes (Cont'd)

Moreover, given the process is at state $i > 0$ at time t, the probability that the next transition is a birth rather than a death is

$$P_{i,i+1} = P(B_i < D_i) = \frac{\lambda_i}{\lambda_i + \mu_i},$$

which implies $P_{i,i-1} = P(D_i < B_i) = \frac{\mu_i}{\lambda_i+\mu_i}$, for $i > 0$.

As only birth is possible at state 0, we know $\nu_0 = \lambda_0$ and $P_{01} = 1$.

To sum up, a birth and death process is a CTMC with state space $\mathcal{X} = \{0, 1, 2, \ldots\}$ such that

$$\nu_i = \lambda_i + \mu_i, i > 0, \quad \nu_0 = \lambda_0,$$
$$P_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i},\ P_{i,i-1} = \frac{\mu_i}{\lambda_i + \mu_i}, i > 0$$
$$P_{01} = 1,\ P_{i,j} = 0 \quad \text{if } |i - j| > 1$$

The parameters $\{\lambda_n\}_{n=0}^{\infty}$ and $\{\mu_n\}_{n=1}^{\infty}$ are called, respectively, the arrival (or birth) and departure (or death) rates.

Examples of Birth and Death Processes

- Poisson Processes: $\mu_n = 0$, $\lambda_n = \lambda$ for all $n \geq 0$
- Pure Birth Process:

$$\mu_n = 0 \quad \Rightarrow \quad \nu_i = \lambda_i,\ P_{i,i+1} = 1,\ P_{i,i-1} = 0$$

- Yule Processes (Pure Birth Process with Linear Growth rate): If there are n people and each independently gives birth at at an exponential rate λ, then the total rate at which births occur is $n\lambda$.

$$\mu_n = 0, \quad \lambda_n = n\lambda$$

Reason: Let

$B_i =$ time until the ith individual give birth $\sim Exp(\lambda),\ i = 1, \ldots, n$

So the time until the next (first) birth when there are n individuals in the population is

$$\min(B_1, B_2, \ldots, B_n) \sim Exp(\lambda + \lambda + \cdots + \lambda) = Exp(n\lambda)$$

So the rate until the next birth is $\lambda_n = n\lambda$.

Example: Linear Growth Model with Immigration

- each individual independently gives birth at an exponential rate λ
- each individual independently die at at an exponential rate μ
- new immigrants come in at an exponential rate θ

Such a process is a birth-death process with birth and death rates

$$\mu_n = n\mu, \quad \lambda_n = n\lambda + \theta$$

Reason: Let

$B_i =$ time until the ith individual give birth $\sim Exp(\lambda),\ i = 1, \ldots, n$
$T =$ time until the next new immigrant comes in $\sim Exp(\theta)$

So the time until the population size increase from n to $n+1$ is

$$\min(B_1, \ldots, B_n, T) \sim Exp(\lambda + \cdots + \lambda + \theta) = Exp(n\lambda + \theta)$$

So the rate until the next birth is $\lambda_n = n\lambda + \theta$.

Similarly, one can show that the death rate is $\mu_n = n\mu$.

Example: $M/M/s$ Queueing Model

- s servers
- Poisson arrival of customers, rate $= \lambda$
- Exponential service time, rate $= \mu$

$\Rightarrow$ a birth and death process with constant birth rate $\lambda_n = \lambda$, and death (departure)rate $\mu_n = \min(n, s)\mu$.

Reason: Suppose, there are n customer in the system at time t. At most $\min(n, s)$ of them are being served. Let S_i be remaining service time of the ith server $\sim$ Exp(μ). Then, the waiting time until the next departure is

$$\min(S_1, \ldots, S_{\min(s,n)}) \sim \text{Exp}(\min(s, n)\mu).$$

6.4 The Transition Probability Function $P_{ij}(t)$

Recall the transition probability function $P_{ij}(t)$ of a CTMC $\{X(t), t \geq 0\}$ is

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

Example. (Poisson Processes with rate λ)

$$\begin{aligned} P_{ij}(t) &= \mathrm{P}(N(t+s) = j | N(s) = i) \\ &= \mathrm{P}(N(t+s) - N(s) = j - i) = \begin{cases} e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!} & \text{if } j \geq i \\ 0 & \text{if } j < i \end{cases} \end{aligned}$$

Properties of Transition Probability Functions

- $P_{ij}(t) \geq 0$ for all $i, j \in \mathcal{X}$ and $t \geq 0$
- (Row sums are 1) $\sum_j P_{ij}(t) = 1$ for all $i \in \mathcal{X}$ and $t \geq 0$

Lemma 6.3 Chapman-Kolmogorov Equation

For all $i, j \in \mathcal{X}$ and $t \geq 0$,

$$P_{ij}(t+s) = \sum_{k \in \mathcal{X}} P_{ik}(t) P_{kj}(s)$$

Proof.

$$\begin{aligned}
&P_{ij}(t+s) \\
&= \mathrm{P}(X(t+s) = j | X(0) = i) \\
&= \sum_{k \in \mathcal{X}} \mathrm{P}(X(t+s) = j, X(t) = k | X(0) = i) \\
&= \sum_{k \in \mathcal{X}} \mathrm{P}(X(t+s) = j | X(t) = k, X(0) = i) \mathrm{P}(X(t) = k | X(0) = i) \\
&= \sum_{k \in \mathcal{X}} \mathrm{P}(X(t+s) = j | X(t) = k) \mathrm{P}(X(t) = k | X(0) = i) \text{ (Markov Property)} \\
&= \sum_{k \in \mathcal{X}} P_{kj}(s) P_{ik}(t)
\end{aligned}$$

Lemma 6.2a

For any $i, j \in \mathcal{X}$, we have

$$\lim_{h \to 0} \frac{1 - P_{ii}(h)}{h} = \nu_i$$

Proof. Let T_i be the amount of time the process stays in state i before moving to other states.

$$\begin{aligned}
P_{ii}(h) &= \mathrm{P}(X(h) = i | X(0) = i) \\
&= \mathrm{P}(X(h) = i, \text{no transition in (0,h]} | X(0) = i) \\
&\quad + \mathrm{P}(X(h) = i, \text{2 or more transition in (0,h]} | X(0) = i) \\
&= \mathrm{P}(T_i > h) + o(h) \\
&= e^{-\nu_i h} + o(h) \\
&= 1 - \nu_i h + o(h)
\end{aligned}$$

Lemma 6.2b

For any $i \neq j \in \mathcal{X}$, we have

$$\lim_{h \to 0} \frac{P_{ij}(h)}{h} = \nu_i P_{ij} \overset{\text{defined as}}{=} q_{ij}$$

where $q_{ij} = \nu_i P_{ij}$ is called the *instantaneous transition rates*.

Proof.

$$\begin{aligned}
P_{ij}(h) &= \mathrm{P}(X(h) = j | X(0) = i) \\
&= \mathrm{P}(X(h) = j, 1 \text{ transition in } (0,\text{h}] | X(0) = i) \\
&\quad + \mathrm{P}(X(h) = j, 2 \text{ or more transition in } (0,\text{h}] | X(0) = i) \\
&= \mathrm{P}(T_i < h) P_{ij} + o(h) \\
&= (1 - e^{-\nu_i h}) P_{ij} + o(h) \\
&= \nu_i P_{ij} h + o(h)
\end{aligned}$$

Theorem 6.1 Kolmogorov's Backward Equations

From Lemma 6.3 (Chapman-Kolmogorov equations), we obtain

$$\begin{aligned} P_{ij}(h+t) - P_{ij}(t) &= \sum_{k\in\mathcal{X}} P_{ik}(h)P_{kj}(t) - P_{ij}(t) \\ &= \sum_{k\in\mathcal{X},k\neq i} P_{ik}(h)P_{kj}(t) - (1-P_{ii}(h))P_{ij}(t) \end{aligned}$$

and thus

$$\lim_{h\to 0} \frac{P_{ij}(t+h) - Pij(t)}{h} = \lim_{h\to 0}\left\{\sum_{k\neq i}\frac{P_{ik}(h)}{h}P_{kj}(t) - \frac{1-P_{ii}(h)}{h}P_{ij}(t)\right\}$$

Now assuming that we can interchange the limit and the summation in the preceding and applying Lemma 6.2, we obtain

$$P'_{ij}(t) = \sum_{k\in\mathcal{X},k\neq i} q_{ik}P_{kj}(t) - \nu_i P_{ij}(t)$$

It turns out that this interchange can indeed be justified.

Theorem 6.2 Kolmogorov's Forward Equations

From Lemma 6.3 (Chapman-Kolmogorov equations), we obtain

$$\begin{aligned} P_{ij}(t+h) - P_{ij}(t) &= \sum_{k \in \mathcal{X}} P_{ik}(t) P_{kj}(h) - P_{ij}(t) \\ &= \sum_{k \in \mathcal{X}, k \neq j} P_{ik}(t) P_{kj}(h) - (1 - P_{jj}(h)) P_{ij}(t) \end{aligned}$$

and thus

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

Now assuming that we can interchange the limit and the summation in the preceding and applying Lemma 6.2, we obtain

$$P'_{ij}(t) = \sum\nolimits_{k \neq j} P_{ik}(t) q_{kj} - \nu_j P_{ij}(t)$$

Unfortunately, this interchange is not always justifiable. However, the forward equations do hold in most models, including all birth and death processes and all finite state models.

Recall that we define the instantaneous transition rates

$$q_{ij} = \nu_i P_{ij}, \quad \text{for} i, j \in \mathcal{X}, i \neq j$$

If we define q_{ii} as $-\nu_i$. For finite state space case $\mathcal{X} = \{1, 2, \ldots, m\}$, define the matrices

$$\mathbf{P}(t) = \begin{bmatrix} P_{11}(t) & \cdots & P_{1m}(t) \\ \vdots & & \vdots \\ P_{m1}(t) & \cdots & P_{mm}(t) \end{bmatrix}, \quad \mathbf{P}'(t) = \begin{bmatrix} P'_{11}(t) & \cdots & P'_{1m}(t) \\ \vdots & & \vdots \\ P'_{m1}(t) & \cdots & P'_{mm}(t) \end{bmatrix},$$

$$\mathbf{Q} = \begin{bmatrix} q_{11} & \cdots & q_{1m} \\ \vdots & & \vdots \\ q_{m1} & \cdots & q_{mm} \end{bmatrix} = \begin{bmatrix} -\nu_1 & \nu_1 P_{12} & \cdots & \nu_1 P_{1m} \\ \nu_2 P_{21} & -\nu_2 & \cdots & \nu_2 P_{2m} \\ \vdots & \vdots & & \vdots \\ \nu_m P_{m1} & \nu_m P_{m2} & \cdots & -\nu_m \end{bmatrix}$$

In matrix notation,
Forward equation: $\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}$
Backward equation: $\mathbf{P}'(t) = \mathbf{Q}\mathbf{P}(t)$