Lecture 13

Scribe: Hanxin Zhang

February 14, 2017

Note: *These lecture notes are still rough, and have only have been mildly proofread.*

Continuous-Time Markov Chains II

As we discussed in the last lecture, for a continuous-time Markov chain $\{X(t), t \geq 0\}$, we can define the *transition probabilities* $P_{ij}(t)$ as

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

Let q_{ij} be the rate at which the Markov chain makes a transition into state j from state i. We have the Kolmogorov's Backward Equations:

$$\frac{dP_{ij}(t)}{dt} = -v_i P_{ij}(t) + \sum_{k \neq i} q_{ik} P_{kj}(t)$$

where $v_i = \sum_{j \neq i} q_{ij}$.

Under certain conditions, we also have the Kolmogorov's Forward Equations

$$\frac{dP_{ij}(t)}{dt} = -v_j P_{ij}(t) + \sum_{k \neq j} q_{kj} P_{ik}(t)$$

Please refer to Ross 11th edition section 6.4 for the proof of Kolmogorov's Equations.

Define the *rate matrix* Q as

$$Q = \begin{bmatrix} -v_1 & q_{12} & q_{13} & \cdots \\ q_{21} & -v_2 & q_{23} & \cdots \\ q_{31} & q_{32} & -v_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Then, we can rewrite the Kolmogorov equations as follows

$$\frac{dP(t)}{d} = QP(t) \qquad \text{(Backward)}$$

$$\frac{dP(t)}{d} = P(t)Q \qquad \text{(Forward)}$$

The solution for these two differential equations is

$$P(t) = e^{Qt}$$

If we factorize the rate matrix Q by eigendecomposition, i.e.

$$Q = VDV^{-1}$$

where D is a diagonal matrix, we can compute $P(t)$ as

$$P(t) = e^{Qt} = Ve^{Dt}V^{-1} = V\text{diag}\{e^{D_{11}t}, e^{D_{22}t}, ...\}V^{-1}$$

Stationary distribution

Usually, the probability that the continuous-time Markov process will be in state j is independent of its initial state i when time t goes to infinity. Let P_j denotes this limiting probabilities or stationary probabilities

$$P_j = \lim_{t\to\infty} P_{ij}(t)$$

An example

Let's consider nucleotide substitution on a site of a DNA sequence. It may change as follows

$$\begin{array}{ll} \text{State} & \text{T} \to \text{A} \to \text{C} \to \text{G} \to \text{C} \to \cdots \\ \text{Time} & 0 \to t_1 \to t_2 \to t_3 \to t_4 \to \cdots \end{array}$$

We can use a continuous-time Markov chain $\{X(t), t \geq 0\}$ to describe this substitution process, which is given by

$$\begin{gathered} X(t) = \text{T}, \quad 0 \leq t < t_1 \\ X(t) = \text{A}, \quad t_1 \leq t < t_2 \\ \vdots \end{gathered}$$

To compute the transition probabilities $P_{ij}(t)$ and the stationary probabilities P_j of this substitution process, we have to determine the rate matrix Q first. One simple choice is the Jukes-Cantor model, in which the rate matrix is

$$Q = \begin{bmatrix} -3\mu & \mu & \mu & \mu \\ \mu & -3\mu & \mu & \mu \\ \mu & \mu & -3\mu & \mu \\ \mu & \mu & \mu & -3\mu \end{bmatrix}$$

where μ is the mutation rate or substitution rate. The transition probabilities are

$$P(t) = \begin{bmatrix} \frac{1}{4} + \frac{3}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} \\ \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} + \frac{3}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} \\ \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} + \frac{3}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} \\ \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} - \frac{1}{4}e^{-4\mu t} & \frac{1}{4} + \frac{3}{4}e^{-4\mu t} \end{bmatrix}$$

The stationary probabilities are

$$P_j = \lim_{t \to \infty} P_{ij}(t) = \frac{1}{4}$$

This means the long-run probability that A, C, G, or T occurring in this process is all equal to 1/4.

Embedded discrete-time Markov chain

Interestingly, there is an embedded discrete-time Markov chain for every continuous Markov chain. The transition probability P_{ij} (do not be confused with $P_{ij}(t)$) is the conditional probability that the process will enter state j when it is in state i, which is given by

$$P_{ij} = \begin{cases} \dfrac{q_{ij}}{v_i} & \text{for } i \neq j \\ \\ 0 & \text{for } i = j \end{cases}$$

Besides, the amount of time it stays in state i before it makes a transition follows an exponential distribution with parameter v_i, i.e.,

$$T_i \sim \text{Exp}(v_i)$$

Global balance and local balance

For a discrete-time Markov chain, we have the global balance condition

$$\pi = \pi P$$

where π is the stationary probability vector.

For a time-reversible process, we have the local balance condition

$$\pi_i P_{ij} = \pi_j P_{ji}$$

Similarly, for a continuous-time Markov chain, we still have the global balance condition

$$\text{rate out } j = \text{rate into } j$$

$$\pi_j v_j = \sum_{k \neq j} \pi_k q_{kj}$$

where $\pi_j = P_j$ is the stationary probability.

If it is time reversible, the local balance condition will be

$$\text{realized rate into } j \text{ from } i = \text{realized rate into } i \text{ from } j$$

$$\pi_i q_{ij} = \pi_j q_{ji}$$

Now, let's consider a birth-death process for which the transition rates are

$$q_{0,1} = \lambda_0$$

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

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

$$q_{i,i} = -(\lambda_i + \mu_i)$$

$$\text{otherwise} \quad q_{ij} = 0$$

The state of this process is represented by the population size i. We can use the global balance condition to solve the stationary probabilities P_i.

State	Rate at which leave $=$ Rate at which enter
0	$\lambda_0 P_0 = \lambda_1 P_1$
1	$(\lambda_1 + \mu_1)P_1 = \lambda_0 P_0 + \mu_2 P_2$
2	$(\lambda_2 + \mu_2)P_2 = \lambda_1 P_1 + \mu_3 P_3$
$j \geq 1$	$(\lambda_i + \mu_i)P_i = \lambda_{i-1} P_{i-1} + \mu_{i+1} P_{i+1}$

Solving these equations recursively, we obtain

$$P_i = \frac{\lambda_{i-1}\lambda_{i-2}...\lambda_1\lambda_0}{\mu_i\mu_{i-1}...\mu_2\mu_1} P_0$$

Using the fact $\sum_{i=0}^{\infty} P_i = 1$, we have

$$P_0 = \frac{1}{1 + \sum_{i=1}^{\infty} \frac{\lambda_{i-1}\lambda_{i-2}...\lambda_1\lambda_0}{\mu_i\mu_{i-1}...\mu_2\mu_1}}$$

and

$$P_i = \frac{\lambda_{i-1}\lambda_{i-2}...\lambda_1\lambda_0}{\mu_i\mu_{i-1}...\mu_2\mu_1 \left(1 + \sum_{i=1}^{\infty} \frac{\lambda_{i-1}\lambda_{i-2}...\lambda_1\lambda_0}{\mu_i\mu_{i-1}...\mu_2\mu_1}\right)}, \qquad i \geq 1$$

It indicates the stationary probabilities exist only if

$$\sum_{i=1}^{\infty} \frac{\lambda_{i-1}\lambda_{i-2}...\lambda_1\lambda_0}{\mu_i\mu_{i-1}...\mu_2\mu_1} < \infty$$

This condition is actually sufficient.

A special case is when $\lambda_i = \lambda$ and $\mu_i = \mu$, it is easy to show

$$P_i = \left(\frac{\lambda}{\mu}\right)^i \left(1 - \frac{\lambda}{\mu}\right), \qquad i \geq 0$$

It is necessary that $\lambda/\mu \leq 1$. We can see that P_i follows a geometric distribution.