11.3.3 The Generator Matrix

Here, we introduce the *generator matrix*. The generator matrix, usually shown by G, gives us an alternative way of analyzing continuous-time Markov chains. Consider a continuous-time Markov chain $X(t)$. Assume $X(0) = i$. The chain will jump to the next state at time T_1, where $T_1 \sim Exponential(\lambda_i)$. In particular, for a very small $\delta > 0$, we can write

$$\begin{aligned} P(T_1 < \delta) &= 1 - e^{-\lambda_i \delta} \\ &\approx 1 - (1 - \lambda_i \delta) \\ &= \lambda_i \delta. \end{aligned}$$

Thus, in a short interval of length δ, the probability of leaving state i is approximately $\lambda_i \delta$. For this reason, λ_i is often called **the transition rate out of state** i. Formally, we can write

$$\lambda_i = \lim_{\delta \to 0^+} \left[\frac{P\big(X(\delta) \neq i | X(0) = i\big)}{\delta} \right] \qquad (11.7)$$

Since we go from state i to state j with probability p_{ij}, we call the quantity $g_{ij} = \lambda_i p_{ij}$, **the transition rate from state** i **to state** j. Here, we introduce the generator matrix, G, whose (i, j)th element is g_{ij}, when $i \neq j$. We choose the diagonal elements (g_{ii}) of G such that the rows of G sum to 0. That is, we let

$$\begin{aligned} g_{ii} &= -\sum_{j \neq i} g_{ij} \\ &= -\sum_{j \neq i} \lambda_i p_{ij} \\ &= -\lambda_i \sum_{j \neq i} p_{ij} \\ &= -\lambda_i. \end{aligned}$$

The last equality resulted as follows: If $\lambda_i = 0$, then clearly

$$\lambda_i \sum_{j \neq i} p_{ij} = \lambda_i = 0.$$

If $\lambda_i \neq 0$, then $p_{ii} = 0$ (no self-transitions), so

$$\sum_{j \neq i} p_{ij} = 1.$$

It turns out the generator matrix is useful in analyzing continuous-time Markov chains.

Example 11.20

Explain why the following approximations hold:

a. $p_{jj}(\delta) \approx 1 + g_{jj}\delta$, for all $j \in S$.
b. $p_{kj}(\delta) \approx \delta g_{kj}$, for $k \neq j$.

Solution

Let δ be small. Equation 11.7 can be written as

$$\begin{aligned} p_{jj}(\delta) &\approx 1 - \lambda_j \delta \\ &= 1 + g_{jj}\delta. \end{aligned}$$

Also, we can approximate $p_{kj}(\delta) = P(X(\delta) = j|X(0) = k)$ as follows. This probability is approximately equal to the probability that we have a single transition from state k to state j in the interval $[0, \delta]$. Note that the probability of more than one transition is negligible if δ is small (refer to the Poisson process section). Thus, we can write

$$\begin{aligned} p_{kj}(\delta) &= P(X(\delta) = j|X(0) = k) \\ &\approx P(X(\delta) \neq k|X(0) = k)p_{kj} \\ &\approx \lambda_k \delta p_{kj} \\ &= \delta g_{kj}, \text{ for } k \neq j. \end{aligned}$$

We can state the above approximations more precisely as

a. $g_{jj} = -\lim_{\delta \rightarrow 0^+} \left[\frac{1 - p_{jj}(\delta)}{\delta}\right]$, for all $j \in S$;
b. $g_{kj} = \lim_{\delta \rightarrow 0^+} \left[\frac{p_{kj}(\delta)}{\delta}\right]$, for $k \neq j$.

The Generator Matrix

For a continuous-time Markov chain, we define the **generator matrix** G. The (i,j)th entry of the transition matrix is given by

$$g_{ij} = \begin{cases} \lambda_i p_{ij} & \text{if } i \neq j \\ \\ -\lambda_i & \text{if } i = j \end{cases}$$

Example 11.21

Consider the continuous Markov chain of Example 11.17: A chain with two states $S = \{0, 1\}$ and $\lambda_0 = \lambda_1 = \lambda > 0$. In that example, we found that the transition matrix for any $t \geq 0$ is given by

$$P(t) = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{bmatrix}.$$

a. Find the generator matrix G.
b. Show that for any $t \geq 0$, we have

$$P'(t) = P(t)G = GP(t),$$

where $P'(t)$ is the derivative of $P(t)$.

Solution

a. First, we have

$$\begin{aligned} g_{00} &= -\lambda_0 \\ &= -\lambda, \\ g_{11} &= -\lambda_1 \\ &= -\lambda. \end{aligned}$$

The transition matrix for the corresponding jump chain is given by

$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Therefore, we have

$$\begin{aligned} g_{01} &= \lambda_0 p_{01} \\ &= \lambda, \\ g_{10} &= \lambda_1 p_{10} \\ &= \lambda. \end{aligned}$$

Thus, the generator matrix is given by

$$G = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}.$$

b. We have

$$P'(t) = \begin{bmatrix} -\lambda e^{-2\lambda t} & \lambda e^{-2\lambda t} \\ \lambda e^{-2\lambda t} & -\lambda e^{-2\lambda t} \end{bmatrix},$$

where $P'(t)$ is the derivative of $P(t)$. We also have

$$P(t)G = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{bmatrix} \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix} = \begin{bmatrix} -\lambda e^{-2\lambda t} & \lambda e^{-2\lambda t} \\ \lambda e^{-2\lambda t} & -\lambda e^{-2\lambda t} \end{bmatrix},$$

$$GP(t) = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix} \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{bmatrix} = \begin{bmatrix} -\lambda e^{-2\lambda t} & \lambda e^{-2\lambda t} \\ \lambda e^{-2\lambda t} & -\lambda e^{-2\lambda t} \end{bmatrix}.$$

We conclude

$$P'(t) = P(t)G = GP(t).$$

The equation $P'(t) = P(t)G = GP(t)$ (in the above example) is in fact true in general. To see the proof idea, we can argue as follows. Let δ be small. By Example 11.20, we have

$$\begin{aligned} p_{jj}(\delta) &\approx 1 + g_{jj}\delta, \\ p_{kj}(\delta) &\approx \delta g_{kj}, \quad \text{for } k \neq j. \end{aligned}$$

Using the Chapman-Kolmogorov equation, we can write

$$\begin{aligned}
P_{ij}(t+\delta) &= \sum_{k\in S} P_{ik}(t)p_{kj}(\delta) \\
&= p_{ij}(t)p_{jj}(\delta) + \sum_{k\neq j} P_{ik}(t)p_{kj}(\delta) \\
&\approx p_{ij}(t)(1+g_{jj}\delta) + \sum_{k\neq j} P_{ik}(t)\delta g_{kj} \\
&= p_{ij}(t) + \delta p_{ij}(t)g_{jj} + \delta \sum_{k\neq j} P_{ik}(t)g_{kj} \\
&= p_{ij}(t) + \delta \sum_{k\in S} P_{ik}(t)g_{kj}.
\end{aligned}$$

Thus,

$$\frac{P_{ij}(t+\delta)-p_{ij}(t)}{\delta} \approx \sum_{k\in S} P_{ik}(t)g_{kj},$$

which is the (i,j)th element of $P(t)G$. The above argument can be made rigorous.

Forward and Backward Equations

The **forward equations** state that

$$P'(t) = P(t)G,$$

which is equivalent to

$$p'_{ij}(t) = \sum_{k\in S} p_{ik}(t)g_{kj}, \text{ for all } i,j \in S.$$

The **backward equations** state that

$$P'(t) = GP(t),$$

which is equivalent to

$$p'_{ij}(t) = \sum_{k\in S} g_{ik}p_{kj}(t), \text{ for all } i,j \in S.$$

One of the main uses of the generator matrix is finding the stationary distribution. So far, we have seen how to find the stationary distribution using the jump chain. The following result tells us how to find the stationary matrix using the generator matrix.

> Consider a continuous Markov chain $X(t)$ with the state space S and the generator Matrix G. The probability distribution π on S is a stationary distribution for $X(t)$ if and only if it satisfies
>
> $$\pi G = 0.$$

Proof:

For simplicity, let's assume that S is finite, i.e., $\pi = [\pi_0, \pi_1, \cdots, \pi_r]$, for some $r \in \mathbb{N}$. If π is a stationary distribution, then $\pi = \pi P(t)$. Differentiating both sides, we obtain

$$\begin{aligned} 0 &= \frac{d}{dt}[\pi P(t)] \\ &= \pi P'(t) \\ &= \pi G P(t) \quad \text{(backward equations)} \end{aligned}$$

Now, let $t = 0$ and remember that $P(0) = I$, the identity matrix. We obtain

$$0 = \pi G P(0) = \pi G.$$

Next, let π be a probability distribution on S that satisfies $\pi G = 0$. Then, by backward equations,

$$P'(t) = G P(t).$$

Multiplying both sides by π, we obtain

$$\pi P'(t) = \pi G P(t) = 0.$$

Note that $\pi P'(t)$ is the derivative of $\pi P(t)$. Thus, we conclude $\pi P(t)$ does not depend on t. In particular, for any $t \geq 0$, we have

$$\pi P(t) = \pi P(0) = \pi.$$

Therefore, π is a stationary distribution.

Example 11.22