Consider now a general birth and death process $\{X(t), t \geq 0\}$ with birth rates $\lambda_0, \lambda_1, \dots$, and death rates $\mu_0, \mu_1, \dots$, where $\mu_0 = 0$. Assume that $X(0) = i$, where $i \geq 0$, and define T_i to be the time until the process enters state $i+1$ for the first time.

Last time, we verified that

$$\mathbb{E}[T_0] = \frac{1}{\lambda_0},$$

$$\mathbb{E}[T_1] = \frac{1}{\lambda_1} + \frac{\mu_1}{\lambda_1 \cdot \lambda_0}$$

$$\mathbb{E}[T_2] = \frac{1}{\lambda_2} + \frac{\mu_2}{\lambda_2}\left(\frac{1}{\lambda_1} + \frac{\mu_1}{\lambda_1} \cdot \frac{1}{\lambda_0}\right)$$

$$\cdots$$

$$\mathbb{E}[T_i] = \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i} \mathbb{E}[T_{i-1}]$$

$$\cdots$$

More generally, assuming that $X(0) = i$, we let T_{ij} be the time until the process enters state j for the first time, where $j > i$. Then we have

$$\mathbb{E}[T_{ij}] = \mathbb{E}[T_{i,i+1}] + \mathbb{E}[T_{i+1,i+2}] + \dots + \mathbb{E}[T_{j-1,j}]$$
$$= \sum_{k=i}^{j-1} \mathbb{E}[T_k].$$

6.4. The transition probability function

Let $\{X(t), t \geq 0\}$ be a continuous-time Markov chain with discrete state space and homogeneous transition probabilities. Denote

$$P_{i,j}(t) := \mathbb{P}(X(t+s) = j \mid X(s) = i).$$

In this section, we will derive some important differential equations for $P_{i,j}(t)$. But before that, let us introduce several auxiliary lemmas.

Recall that the definition of a homogeneous Markov chain can be given as follows:

A continuous-time Markov chain with homogeneous transitions and discrete state space $\mathcal{X}$ is a stochastic process such that:

(i) the times spent in different states are independent random variables (because of the Markov property);

(ii) the amount of time spent in each state $i \in \mathcal{X}$ is an exponentially distributed random variable with parameter ν_i (because of the Markov property and stationarity);

(iii) when the process leaves a state $i \in \mathcal{X}$ it enters a state j with the following transition probabilities:

$$Q_{i,i} = 0, \quad i \in \mathcal{X},$$
$$\sum_{j \in \mathcal{X}} Q_{i,j} = 1, \quad i \in \mathcal{X}.$$

(iv) Transitions follow a discrete time Markov chain.

Denote $\quad q_{i,j} := \nu_i Q_{i,j}$

Interpretation: $\quad q_{i,j}$ — rate at which the process makes a transition to state j, when in state i.

The quantities $q_{i,j}$ are called instantaneous transition rates.

Since we have that

$$\sum_{j\in\mathcal{X}} q_{i,j} = \nu_i \sum_{j\in\mathcal{X}} Q_{i,j} = \nu_i,$$

$$Q_{i,j} = \frac{q_{i,j}}{\nu_i} = \frac{q_{i,j}}{\sum_{j\in\mathcal{X}} Q_{i,j}} \Rightarrow \text{the probabilistic properties}$$

of $\{X(t), t\geq 0\}$ are completely characterised by the $q_{i,j}$'s.

Lemma 6.2 Let $P_{i,j}(t) = \mathbb{P}(X(t)=j \mid X(0)=i)$, $i,j\in\mathcal{X}$. Then

(a) $$\lim_{h\to 0} \frac{1-P_{i,i}(h)}{h} = \nu_i.$$

(b) $$\lim_{h\to 0} \frac{P_{i,j}(h)}{h} = q_{i,j} \quad \text{for all} \quad j\neq i.$$

Proof Let $X(0)=i$.

We start by noting that the amount of time T_i until a transition from i occurs is exponentially distributed. In particular, this means under some mild assumptions that

$\mathbb{P}($ there will be no transition on interval $[0,h]) =$

$$= \mathbb{P}(T_i > h) = e^{-\nu_i h} = 1-\nu_i h + o(h), \quad h\to 0$$

$\mathbb{P}($ there will be one transition on interval $[0,h]) = \nu_i h + o(h)$

$\mathbb{P}($ there will be two or more transitions on $[0,h]) = o(h)$

as $h\to 0$.

Hence

$$\mathbb{P}(X(h) \neq i \mid X(0)=i) = 1 - \mathbb{P}(X(h)=i \mid X(0)=i) = 1 - P_{i,i}(h)$$
$$= \lambda_i h + o(h), \quad h \to 0.$$

Hence

$$\lim_{h \to 0} \frac{1 - P_{i,i}(h)}{h} = \lambda_i .$$

Similarly, if $i \neq j$, we get

$$P_{i,j}(h) = \mathbb{P}(X(h)=j \mid X(0)=i) = \lambda_i Q_{i,j} h + o(h), \quad \text{whence}$$

$$\lim_{h \to 0} \frac{P_{i,j}(h)}{h} = \lambda_i Q_{i,j} = q_{i,j}$$

◢

Lemma 6.3 (Chapman-Kolmogorov equations) For all $s, t \geq 0$ and $i,j \in \mathcal{X}$, we have

$$P_{i,j}(t+s) = \sum_{k \in \mathcal{X}} P_{i,k}(t) P_{k,j}(s).$$

Proof We have:

$$P_{i,j}(t+s) = \mathbb{P}(X(t+s)=j \mid X(0)=i)$$

$$= \sum_{k \in \mathcal{X}} \mathbb{P}(X(t+s)=j, X(t)=k \mid X(0)=i) =$$
$$= \sum_{k \in \mathcal{X}} \mathbb{P}(X(t+s)=j \mid X(t)=k, X(0)=i)\, \mathbb{P}(X(t)=k \mid X(0)=i)$$
$$= \sum_{k \in \mathcal{X}} \mathbb{P}(X(t+s)=j \mid X(t)=k)\, \mathbb{P}(X(t)=k \mid X(0)=i)$$
$$= \sum_{k \in \mathcal{X}} P_{i,k}(t) P_{k,j}(s)$$

◢

Theorem 6.1 For all $t \geq 0$ and states $i,j \in \mathcal{X}$, we have

$$P'_{i,j}(t) = \sum_{k \in \mathcal{X} \setminus \{i\}} q_{i,k} P_{k,j}(t) - \lambda_i P_{i,j}(t). \qquad \text{(KBE)}$$

Proof By Lemma 6.3, we have:

$$P_{i,j}(t+h) - P_{i,j}(t) = \sum_{k\in\mathcal{X}} P_{i,k}(h)\, P_{k,j}(t) - P_{i,j}(t)$$

$$= \sum_{k\in\mathcal{X}\setminus\{i\}} P_{i,k}(h)\, P_{k,j}(t) + P_{i,i}(h)\, P_{i,j}(t) - P_{i,j}(t)$$

$$= \sum_{k\in\mathcal{X}\setminus\{i\}} P_{i,k}(h)\, P_{k,j}(t) - P_{i,j}(t)\,(1 - P_{i,i}(h))$$

Hence

$$\frac{P_{i,j}(t+h) - P_{i,j}(t)}{h} = \sum_{k\in\mathcal{X}\setminus\{i\}} \frac{P_{i,k}(h)}{h}\, P_{k,j}(t) - P_{i,j}(t)\, \frac{1 - P_{i,i}(h)}{h}$$

By $h \to 0$ and Lemma 6.2, we obtain the result

Kolmogorov backward equations:

$$P'_{i,j}(t) = \sum_{k\in\mathcal{X}\setminus\{i\}} q_{i,k}\, P_{k,j}(t) - \lambda_i P_{i,j}(t), \quad i,j \in \mathcal{X}.$$

Remark Recall that

$$\lambda_i = \sum_{j\in\mathcal{X}} q_{i,j} = \sum_{j\in\mathcal{X}} \lambda_i Q_{i,j} = [Q_{i,i}=0] = \sum_{j\in\mathcal{X}\setminus\{i\}} q_{i,j}$$

Hence $\sum_{k\in\mathcal{X}\setminus\{i\}} q_{i,k} - \lambda_i = 0.$

We now assume that $\mathcal{X} = \{1, 2, \dots, n\}$ and introduce the following matrices:

$$R := \begin{pmatrix} -\nu_1 & q_{1,2} & q_{1,3} & \cdots & q_{1,n} \\ q_{2,1} & -\nu_2 & q_{2,3} & \cdots & q_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{n,1} & q_{n,2} & q_{n,3} & \cdots & -\nu_n \end{pmatrix}$$

$$P(t) = \begin{pmatrix} P_{1,1}(t) & P_{1,2}(t) & \cdots & P_{1,n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ P_{n,1}(t) & P_{n,2}(t) & \cdots & P_{n,n}(t) \end{pmatrix}$$

Then the Kolmogorov backward equations can be written as

$$P'(t) = R\,P(t).$$

Example 6.10 Let $\{X(t),\ t \geq 0\}$ be a birth and death process with rates $\{\lambda_n\}_{n\geq 0}$, $\{\mu_n\}_{n\geq 0}$. We have:

$$q_{n,n+1} = (\lambda_n + \mu_n) \cdot \frac{\lambda_n}{\lambda_n + \mu_n} = \lambda_n, \quad n \geq 0$$

$$q_{n,n-1} = (\lambda_n + \mu_n) \cdot \frac{\mu_n}{\lambda_n + \mu_n} = \mu_n, \quad n \geq 0$$

$$q_{n,k} = 0 \quad \text{otherwise.}$$

The matrix R then has the form

$$R = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$P'(t) = R\,P(t)$$

$$P'_{0,j}(t) = -\lambda_0 P_{0,j}(t) + \lambda_0 P_{1,j}(t) = \lambda_0 P_{1,j}(t) - \lambda_0 P_{0,j}(t)$$

$$P'_{i,j}(t) = \mu_i P_{i-1,j}(t) + \lambda_i P_{i+1,j}(t) - (\lambda_i + \mu_i) P_{i,j}(t), \quad i \neq 0.$$

◢

Theorem 6.2 For all $t \geq 0$ and states $i,j \in \mathcal{X}$ we have:

$$P'_{i,j}(t) = \sum_{k \in \mathcal{X} \setminus \{j\}} P_{i,k}(t)\, q_{k,j} - P_{i,j}(t)\, \nu_j \qquad \text{(KFE)}$$

Proof By Lemma 6.2,

$$P_{i,j}(t+h) - P_{i,j}(t) = \sum_{k \in \mathcal{X}} P_{i,k}(t) \cdot P_{k,j}(h) - P_{i,j}(t) =$$

$$= \sum_{k \in \mathcal{X} \setminus \{j\}} P_{i,k}(t) P_{k,j}(h) - \big(P_{i,j}(t) - P_{i,j}(t) P_{j,j}(h) \big) =$$

$$= \sum_{k \in \mathcal{X} \setminus \{j\}} P_{i,k}(t) P_{k,j}(h) - P_{i,j}(t) \big(1 - P_{j,j}(h) \big).$$

Now divide both sides by h and apply Lemma 6.3 ◢

Equations (KFE) are called Kolmogorov forward equations.

Matrix form:

$$P'(t) = P(t)\,R$$

Remark (Warning) In the "proofs" of Lemma 6.2, Theorem 6.1 and Theorem 6.2 we performed unjustified interchanges of limiting operations, interchanges of sums and limits etc

For "reasonable" Markov chains they hold, but in general one should be very careful: one can come up with an example of a Markov chain such that for some its state i

$$\lim_{h \to 0} \frac{1 - P_{i,i}(h)}{h} = -\infty,$$

which completely spoils all the arguments we had. However, these technicalities are out of scope of our course.

Example 6.12 Let $\{X(t), t \geq 0\}$ be again a birth and death process.

$$R = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Then the Kolmogorov forward equations look as follows:

there was a typo here in the previous version of the lecture

$$P'_{i,0}(t) = P_{i,1}(t)\mu_1 - P_{i,0}(t)\lambda_0$$

$$P'_{i,j}(t) = P_{i,j-1}(t)\lambda_{j-1} + P_{i,j+1}(t)\mu_{j+1} - (\lambda_j+\mu_j)P_{i,j}(t)$$