Lecture 6.5

In this lecture Poisson processes are derived from first principles and the notion of Markov chains is extended to include continuous time Markov chains. This is necessary for our original aim of modelling the M/M/1 queue with Markov chains (remember this stands for a queue with Poisson arrivals, Poisson service time and one server).

Poisson processes

There are many ways to derive a Poisson process. Let $A(t)$ (for $t \geq 0$) be the number of customers arriving in the interval $[0, t]$ (with $A(t) = 0$). Consider the following requirements,

1. The number of arrivals in disjoint time periods are independent.
2. For a small time period δt the probability of a single arrival in the period is given by

$$\mathbb{P}\left[A(t+\delta t) - A(t) = 1\right] = \lambda \delta t + o(\delta t),$$

   where λ is known as the *rate* of the process and $o(\delta t)$ is some function (which maybe negative or positive) which tends to zero in the limit as δt tends to zero.
3. The probability that two or more arrivals occur in the same small time period is negligible. More formally,

$$\mathbb{P}\left[A(t+\delta t) - A(t) \geq 2\right] = o(\delta t).$$

From the above we can derive

$$\mathbb{P}\left[A(t+\delta t) - A(t) = 0\right] = 1 - \lambda \delta t + o(\delta t).$$

Define $P_n(t) = \mathbb{P}\left[A(t) = n\right]$. Now, consider how $P_n(t)$ evolves in some short time period δt.

$$\begin{aligned} P_n(t+\delta t) &= P_n(t)(1-\lambda\delta t) + P_{n-1}(t)\lambda\delta t + o(\delta t) \qquad n > 0, \\ P_0(t+\delta t) &= P_0(t)(1-\lambda\delta t) + o(\delta t). \end{aligned}$$

These can be rewritten as

$$\begin{aligned} \frac{P_n(t+\delta t) - P_n(t)}{\delta t} &= \frac{-P_n(t)(\lambda\delta t) + P_{n-1}(t)\lambda\delta t + o(\delta t)}{\delta t} \qquad n > 0, \\ \frac{P_0(t+\delta t) - P_0(t)}{\delta t} &= \frac{P_0(t)(-\lambda\delta t) + o(\delta t)}{\delta t}. \end{aligned}$$

These equations are sometimes called differential difference equations and are often used in queuing theory. Now take the limit as $\delta t \to 0$ which causes the terms $o(\delta t)$ to vanish and gives

$$\begin{aligned} \frac{dP_n(t)}{dt} &= -\lambda P_n(t) + \lambda P_{n-1}(t) \qquad n > 0 \\ \frac{dP_0(t)}{dt} &= -\lambda P_0(t). \end{aligned}$$

Now, the equation for $P_0(t)$ can be trivially solved and this gives

$$P_0(t) = Ce^{-\lambda t},$$

where C is a constant which is shown to be 1 from the boundary condition $P_0(0) = 1$ (since $A(0) = 0$).

Substituting $n = 1$ gives,

$$\frac{dP_1(t)}{dt} = -\lambda P_1(t) + \lambda P_0(t) = -\lambda P_1(t) + \lambda e^{-\lambda t}.$$

Solving gives

$$P_1(t) = \lambda t e^{-\lambda t},$$

which can be checked by substitution into the previous equation.

Continuing this analysis will give

$$P_2(t) = \frac{(\lambda t)^2 e^{-\lambda t}}{2!}$$
$$P_3(t) = \frac{(\lambda t)^3 e^{-\lambda t}}{3!},$$

and so on.

In general the relationship

$$P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!},$$

is suspected and this can be shown trivially by induction.

Now, this implies,

$$\mathbb{P}\left[A(t) = n\right] = P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!},$$

and since one of our other conditions was that the arrivals in disjoint time periods are independent then we could repeat this analysis to find the number of packets arriving since some time t and get,

$$\mathbb{P}\left[A(t+\tau) - A(t) = n\right] = \mathbb{P}\left[A(\tau) = n\right] = \frac{(\lambda \tau)^n e^{-\lambda \tau}}{n!},$$

which is the equation for a Poisson process. It can therefore be seen that the three simple conditions defined in the beginning are sufficient that a Poisson process is the only process which meets those conditions.

Aside, the mean and variance of a Poisson distribution

Assume the variable X has a Poisson distribution with parameter λ. Its expectation is given by

$$\begin{aligned}
\mathrm{E}\left[X\right] = \sum_{n=0}^{\infty} n \mathbb{P}\left[X = n\right] &= \sum_{n=1}^{\infty} n \frac{\lambda^n e^{-\lambda}}{n!} \\
&= \sum_{n=1}^{\infty} \lambda \frac{\lambda^{n-1} e^{-\lambda}}{(n-1)!} \\
&= \lambda \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \\
&= \lambda.
\end{aligned}$$

A similar calculation will show that $\overline{X}$ (the sample mean) is an unbiased estimator of λ. Proceeding in the same way,

$$\begin{aligned}
\mathrm{E}\left[X^2\right] = \sum_{n=0}^{\infty} n^2 \mathbb{P}\left[X = n\right] &= \sum_{n=1}^{\infty} n^2 \frac{\lambda^n e^{-\lambda}}{n!} \\
&= \lambda \sum_{n=1}^{\infty} n \frac{\lambda^{n-1} e^{-\lambda}}{(n-1)!} \\
&= \lambda \sum_{n=0}^{\infty} (n+1) \frac{\lambda^n e^{-\lambda}}{n!} \\
&= \lambda \sum_{n=0}^{\infty} \left[n \frac{\lambda^n e^{-\lambda}}{n!} + \frac{\lambda^n e^{-\lambda}}{n!} \right] \\
&= \lambda^2 + \lambda.
\end{aligned}$$

The variance is therefore given by

$$\sigma^2 = \mathrm{E}\left[X^2\right] - \mathrm{E}\left[X\right]^2 = \lambda.$$

Continuous time Markov chains

The Markov chain considered previously was a homogeneous discrete time Markov chain where homogeneous in this case implies that the transition probabilities remain constant and discrete time implies that the process has a certain step length (in the original hitch-hiking hippy example it was assumed that the hippy travelled once and once only every day). The continuous time Markov chain weakens this assumption by allowing a transition between states of the chain to occur at any time. Instead of having discrete steps and transition probabilities, instead the states act like Poisson processes with given flow rates connecting them. Define λ_{ij} where $i \neq j$ as the flow rate between state i and j.

Consider first the discrete time Markov chain which has a time step δt and which has its transition matrix defined by

$$\mathbf{P}(\delta t) = \begin{bmatrix} 1 - \lambda_{00}\delta t & \lambda_{01}\delta t & \lambda_{02}\delta t & \dots \\ \lambda_{10}\delta t & 1 - \lambda_{11}\delta t & \lambda_{12}\delta t & \dots \\ \lambda_{20}\delta t & \lambda_{21}\delta t & 1 - \lambda_{22}\delta t & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

In order that this is a Markov chain then it must be the case that $\lambda_{ii} = \sum_{j \neq i} \lambda_{ij}$ in order that the rows add up to one. It must also be the case that δt is sufficiently small that $\forall i, j : \lambda_{ij}\delta t < 1$. The continuous time Markov chain could be thought of as the limit of this process as $\delta t \to 0$. However, the matrix defined above would simply become the identity matrix. A new approach is needed.

Define the following (assuming that the states of the chain are numbered $(0, 1, 2, \dots)$.

- $X(t)$ is the state of the chain at some time $t \geq 0$.
- $\mathbf{f}(t) = (f_0(t), f_1(t), \dots)$ is the vector of probabilities at time t, formally $f_i = \mathbb{P}\left[X(t) = i\right]$.
- $q_{ij}(t_1, t_2)$ where $t_1 < t_2$ is $\mathbb{P}\left[X(t_2) = j | X(t_1) = i\right]$.

Since the context is still homogeneous chains then these probabilities are just a function of $\tau = t_2 - t_1$. Hence, define for $i \neq j$

$$q_{ij}(\tau) = q_{ij}(t_2 - t_1) = q_{ij}(t_1, t_2) = \mathbb{P}\left[X(\tau) = j | X(0) = i\right].$$

Define the limit

$$q_{ij} = \lim_{\tau \to 0} \frac{q_{ij}(\tau)}{\tau}.$$

Consider the transition rates between states i and j which were defined as λ_{ij} (where $i \neq j$). As the process is like a Poisson then for a given state i and a small time period δt

$$f_i(t + \delta t) = f_i(t) - \sum_{j \neq i} f_i(t)\lambda_{ij}\delta t + \sum_{j \neq i} f_j(t)\lambda_{ji}\delta t + o(\delta t).$$

Taking the difference, dividing through by δt and taking the limit as $\delta t \to 0$ (a differential difference equation again) then

$$\frac{df_i(t)}{dt} = -\sum_{j \neq i} f_i(t)\lambda_{ij} + \sum_{j \neq i} f_j(t)\lambda_{ji}.$$

Similarly, this could be derived in terms of $q_{ij}(t)$ so

$$f_i(t + \delta t) = f_i(t) - \sum_{j \neq i} f_i(t)q_{ij}(\delta t) + \sum_{j \neq i} f_j(t)q_{ji}(\delta t).$$

As before this becomes

$$\frac{df_i(t)}{dt} = -\sum_{j \neq i} f_i(t)q_{ij} + \sum_{j \neq i} f_j(t)q_{ji}.$$

Therefore, for $i \neq j$ it is the case that $q_{ij} = \lambda_{ij}$. Now, q_{ii} still needs definition. It is handy to define

$$q_{ii} = -\sum_{i \neq j} q_{ij}.$$

Now, define the matrix

$$\mathbf{Q} = \begin{bmatrix} q_{00} & q_{01} & q_{02} & \cdots \\ q_{10} & q_{11} & q_{12} & \cdots \\ q_{20} & q_{21} & q_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

It can now be seen that

$$\frac{d\mathbf{f}(t)}{dt} = \mathbf{f}(t)\mathbf{Q}.$$

This is similar to the equation for the discrete time homogenous Markov chain but the rows of $\mathbf{Q}$ add up to zero not one. Notice that

$$\mathbf{Q} = \mathbf{P}(1) - \mathbf{I},$$

where $\mathbf{I}$ is the identity matrix.

Consider the ergodicity conditions as before. In this case irreducible is as before. The continuous time chain can never be periodic because there is always a transition from a state i to itself. The recurrent non-null requirement is similar to that for discrete time chains.

Now, assuming the chain in question is ergodic then the equilibrium probilities can be calculated. In this case

$$\boldsymbol{\pi} = \lim_{t \to \infty} \mathbf{f}(t).$$

Therefore

$$\boldsymbol{\pi}\mathbf{Q} = \frac{d\lim_{t\to\infty}\mathbf{f}(t)}{dt} = \frac{d}{dt}\lim_{t\to\infty}\mathbf{f}(t) = \frac{d\boldsymbol{\pi}}{dt} = 0,$$

where the equality to zero comes from the fact that by definition $\boldsymbol{\pi}$ is constant. The equation $\boldsymbol{\pi}\mathbf{Q} = 0$ is the continuous Markov chain equivalent of $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$ for discrete chains.

This gives a new version of our balance equations. For all i then

$$\sum_j \pi_j q_{ji} = 0$$

Expanding q_{jj} from its definition and multiplying by -1 gives

$$\sum_{j\neq i} \pi_i q_{ij} - \sum_{j\neq i} \pi_j q_{ji} = 0.$$

This can also be seen as a balance of flows into and out of the state. Also as usual $\sum_i \pi_i = 1$.