Stats 150, Fall 2019
Lecture 16, Tuesday, 10/29/2019

We're moving onto chapter 4, Continuous Time Markov Chains. There are elements from discrete-time markov chains, some elements of Poisson processes, and some elements of renewal theory.

1 Continuous Time Markov Chains

The idea here is that we want to have a model for a continuous time process $(X_t, t \geq 0)$, where t is a now a continuous nonnegative variable. We say this process has the Markov property with a **transition semi-group** $(P_t, t \geq 0)$ if

$$\mathbb{P}\left(X_{s+t} = y \mid X_s = x, \text{any event determined by } (X_n, 0 \leq u \leq s)\right) = P_t(x, y),$$

for a future target y, given that at time s, we arrive at state x.
To make sense of this, we need some things. Let us start with a finite S (or countably infinite) on $\{0, 1, 2, \dots\}$. Then for $x, y \in S$, for each fixed $t \geq 0$, $P_t(x, y)$ should be a transition probability matrix.
There is a consistency issue as we vary s and t in this specification. Recall that in discrete time, we have that $P_1 = P$ is one transition matrix, then we would make $P_n = P^n$, which happens if and only if $P_n P_m = P_{n+m}$. This must be so by the addition rule of probability. Now by this exactly same discussion in continuous time, let's use $\mathbb{P}_x$ as usual to specify $\mathbb{P}_x(X_0 = x) = 1$. This simply forces the initial state to be x. More or less by definition, we have

$$\begin{aligned}
\mathbb{P}_{s+t}(x, y) &= \mathbb{P}_x(X_{s+t} = y) \\
&= \sum_{y \in S} \mathbb{P}_x\left(X_s = y, X_{s+t} = z\right) \quad \text{(conditioning on } X_s\text{)} \\
&= \vdots \\
&= \sum_{y \in S} P_s(x, y) P_t(y, z)
\end{aligned}$$

From this, we see that analogously to discrete time Markov chains, we must have the identity

$$P_{s+t} = P_s P_t = P_t P_s$$

which we call the semigroup property of the family of transition matrices driving a continuous Markov chain. The second equality follows from the first equality, under the fact that $s + t = t + s$.
We may ask, how do we create such a transition mechanism? Pitman mentions there are two answers which end up more or less the same but from different perspectives.

Derivation 1. First, we can view this via analysis and see that the only candidate for a mechanism like this is to take

$$P_t = e^{Qt},$$

for a suitable matrix Q. Now the natural question is to ask how we compute this.

$$e^{Qt} = \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!} = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n,$$

via the Taylor series expansion. Now this is well defined whenever Q is a finite matrix or an infinite matrix with bounded elements. That is, we have some finite bound λ where $|Q(i,j)| \leq \lambda < \infty$ for all i, j. Additionally, we can approach this the compound-interest way:

$$e^{Qt} = \lim_{n \to \infty} \left(I + \frac{Qt}{n} \right)^n,$$

which we can check is true easily by binomial expansion. Moreover, notice that

$$e^{Q(s+t)} = e^{Qs+Qt} = e^{Qs} \cdot e^{Qt},$$

and so if we set $P_t := e^{Qt}$, then we will see that $P_{s+t} = P_s P_t$ as desired.

Derivation 2. Alternatively, we can view this via the explicit construction of $(X_t, t \geq 0)$ from suitable building blocks. A good start to do this to consider some transition matrix U (from Durett example 4.1) for a discrete time Markov chain, say $Y_0, Y_1, Y_2, \ldots$ that are distributed like the transition matrix U. Consider a Poisson process $(N(t), t \geq 0)$ of rank 1, with iid Exponential(1) spacings. Additionally, we assume that the Poisson point process (the times of jumps) is independent of the discrete chain (where the jumps are going). The Poisson process says when the contiuous time Markov chain should (try) to move.
Now we write down and consider the process X_t is by definition:

$$X_t := Y_{N(\lambda t)},$$

where $\lambda > 0$ is a rate parameter and so $(N(\lambda t), t \geq 0)$ is a Poisson process with rate λ. We can see explicitly $N(\lambda t) \sim \text{Poisson}(\lambda t)$.
We claim that we get a continuous time Markov chain like this. In fact, we can make every chain with finite state space like this, modulo some technical assumptions.
Each of these bands are Exponential(λ). Now there is one complication in that it can happen that the chain does not move (the case where the chain has a self-transition).

As an aside, if $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are iid Exponential(λ) with mean $\frac{1}{\lambda}$ and G(p) is geometric (p) on $\{1, 2, \ldots\}$ with mean $\frac{1}{p}$, then $\mathcal{E}_1 + \mathcal{E}_2 + \cdots + \mathcal{E}_{G(p)}$ is Exponential($p\lambda$).
To see this, condition on $G(p)$ and notice

$$\begin{aligned}\mathbb{E}\left(\mathcal{E}_1 + \mathcal{E}_2 + \cdots + \mathcal{E}_{G(p)}\right) &= (\mathbb{E}\mathcal{E}_1)\left(\mathbb{E}(G(p))\right) \\ &= \frac{1}{\lambda}\frac{1}{p} = \frac{1}{p\lambda}\end{aligned}$$

Returning to our probability argument, let us find Q for the Poisson construction:

$$\begin{aligned}
P_t(x,y) &= \mathbb{P}_x(Y_{N(\lambda t)} = y) \\
&= \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} U^n(x,y) \\
&= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t U)^n}{n!}(x,y) \\
&= \left[e^{-\lambda t} \cdot e^{\lambda t U}\right](x,y) \\
&= e^{Qt},
\end{aligned}$$

where

$$Q := \lambda(U - I),$$

and $I = P^0$ is the identity matrix. Here, $\mathbb{P}_x$ has $Y_0 = x$ and N in a Poisson point process is independent of Y, and the (x,y) at the right is the evaluation of a matrix at row x and column y.

2 Simple Examples

1. Take $Y_n := n$ for $n = 0, 1, 2, \ldots$, which is the (very boring and simple) deterministic Markov chain that increase by 1 at each time. Then

$$U(x,y) = \mathbb{1}(Y = x+1).$$

This gives

$$X_t = Y_{N(\lambda t)} = N(\lambda t),$$

which is simply a Poisson process with rate λ.

2. Consider $Y_n = \Delta_1 + \Delta_2 + \cdots + \Delta_n$ for some iid Δ_i. Then add a Poisson number of these terms:

$$X_t = \Delta_1 + \Delta_2 + \cdots + \Delta_{N(\lambda t)},$$

which is called a compound Poisson process, or in other words a Negative Binomial process as per Homework 7.
Note that probability generating functions are very handy for this example, as we have seen in our homework.

3 Interpretation of Rate Matrix Q

We agree that we are creating our semigroup by the matrix formalism in that

$$P_t = e^{Qt}.$$

Now this implies that something nice happens when we differentiate:

$$\implies \frac{d}{dt} P_t = Q e^{Qt} = e^{Qt} Q = Q P_t = P_t Q, \tag{1}$$

where via Taylor's,

$$\begin{aligned}\frac{d}{dt}\left(I + Q + \frac{Q^2t^2}{2!} + \frac{Q^3t^3}{3!} + \cdots\right) &= 0 + Q + \frac{Q^2}{1!}t + \frac{Q^3t^2}{2!} + \cdots \\ &= Q\left(I + \frac{Qt}{1!} + \frac{Q^2t^2}{2!} + \cdots\right) \\ &= Qe^{Qt}.\end{aligned}$$

Notice that Q commutes with every power of itself, so we could very well have (1) above, which are the Kolmogorov, backwards, and forwards equations respectively.
Formally, $(P_t, t \geq 0)$ is defined by a matrix system of differential equations, given by the set of Kolmogorov's differential equations $P_0 = I$ and $\frac{d}{dt}P_t = QP_t = P_tQ$. Pitman warns this is fairly abstract, and we will focus on the interpretation. From the formalism and equations introduced above, we have:

$$Q = \frac{d}{dt}P_t\Big|_{t=0^+} = \lim_{t\downarrow 0}\left(\frac{P_t - I}{t}\right),$$

which we can perform for each entry one at a time. Now, notice that for a column vector $\mathbb{1}$ of all entries 1,

$$P_t\mathbb{1} = \mathbb{1}.$$

Similarly,

$$\frac{d}{dt}P_t\mathbb{1} = \frac{d}{dt}\mathbb{1} = 0,$$

the zero vector. This implies that

$$Q\mathbb{1} = \frac{d}{dt}(P_t) = 0.$$

Remark: Now this tells us that all row sums of a rate matrix Q must be identically zero.
Additionally, the off-diagonal entries (for $i \neq j$) are

$$Q(i,j) = \lim_{t\to 0}\frac{P_t(i,j)}{t} \geq 0,$$

because this is the limit of nonnegative things. Hence the main diagonal entries $Q(i,i)$ MUST all be negative to compensate so that the row sums are all 0.

Now let us define

$$\lambda_i := -Q(i,i) \geq 0$$

which has the interpretation that the holding time of the chain in state i is Exponential(λi). Then for $i \neq j$, the

$$\frac{Q(i,j)}{\lambda i}$$

are transition probabilities when the chain jumps.
We keep moving around in the state space, and every time we hold, we hold for an exponential time. This can be read from the rate matrix Q. This is a very high level view of the continuous time Markov chain theory.

Lecture ends here.