Mathematics 275C – Stochastic Processes – T. Liggett

I. The construction problem for continuous time Markov chains

Reference: Part II of
David Freedman, *Markov Chains*, Holden-Day, 1971.

Suppose that S is a countable set,
$\Omega =$ the set of all right continuous step functions $\omega : [0,\infty) \to S$ with finitely many jumps in any finite time interval, $X_t(\omega) = \omega(t)$, and
$\{P^x, x \in S\}$ are probability measures on Ω that satisfy
(i) $P^x(X_0 = x) = 1$ for all $x \in S$, and
(ii) the Markov property

$$E^x(Y \circ \theta_t \mid \mathcal{F}_t) = E^{X_t} Y \quad a.s.\ P^x$$

for all $x \in S$ and all bounded measurable Y on Ω.

Let $p_t(x,y) = P^x(X_t = y)$ be the corresponding transition probabilities. Then (i) and (ii) imply

$$p_t(x,y) \geq 0, \quad \sum_y p_t(x,y) = 1, \quad \lim_{t \downarrow 0} p_t(x,x) = p_0(x,x) = 1 \tag{1}$$

and the Chapman-Kolmogorov equation

$$p_{t+s}(x,y) = \sum_z p_s(x,z) p_t(z,y). \tag{2}$$

Theorem 1. *Suppose that $p_t(x,y)$ satisfy (1) and (2). Then*
(a) $p_t(x,x) > 0$ for all $t > 0, x \in S$.
(b) If $p_t(x,x) = 1$ for some $t > 0$ and $x \in S$, then $p_t(x,x) = 1$ for all $t > 0$ and that x.
(c) For all $x, y \in S$, $p_t(x,y)$ is uniformly continuous in t, and in fact

$$|p_t(x,y) - p_s(x,y)| \leq 1 - p_{|t-s|}(x,x).$$

(d) For every x, the following exists:

$$c(x) = -\frac{d}{dt} p_t(x,x)\Big|_{t=0} \in [0,\infty].$$

(e) $p_t(x,x) \geq e^{-c(x)t}$.
(f) If $c(x) < \infty$, then

$$q(x,y) = \frac{d}{dt} p_t(x,y)\Big|_{t=0}$$

exists and is finite, and then

$$\sum_y q(x,y) \leq 0.$$

(g) If $c(x) < \infty$ and $\sum_y q(x,y) = 0$, then $p_t(x,y)$ is continuously differentiable in t and satisfies the Kolmogorov backward equation

$$\frac{d}{dt}p_t(x,y) = \sum_z q(x,z)p_t(z,y). \tag{3}$$

As we will see shortly, $c(x)$ has the interpretation of the parameter of the exponential holding time at x. For this reason, a state x for which $c(x) = \infty$ is called instantaneous – the chain spends zero time at x before leaving x. Also $q(x,y)$ is thought of as the rate at which the chain goes from x to y.

We will not prove Theorem 1, since we are primarily interested in the converse problem: Suppose $q(x,y)$ satisfies

$$q(x,y) \geq 0 \text{ for } y \neq x \quad \text{and} \quad \sum_y q(x,y) = 0 \text{ for all } x \in S. \tag{4}$$

Under what conditions is there a unique solution $p_t(x,y)$ satisfying (1), (2) and (3)?

Before considering this problem, we look at an example due to Blackwell in which all states are instantaneous.

Example. Let $\{X_t(i), i = 0,1,...\}$ be independent Markov chains with state space $\{0,1\}$ and transitions $0 \to 1$ at rate $\beta(i)$ and $1 \to 0$ at rate $\delta(i)$. For this chain,

$$P^0(X_t(i) = 1) = \frac{\beta(i)}{\delta(i)+\beta(i)}\left(1 - e^{-t(\delta(i)+\beta(i))}\right) \leq \frac{\beta(i)}{\delta(i)+\beta(i)},$$

as can be seen by solving (3). Let's assume that

$$\sum_i \frac{\beta(i)}{\delta(i)+\beta(i)} < \infty, \tag{5}$$

let $X_t = (X_t(0), X_t(1), ...)$, and

$$S = \left\{x = (x(0), x(1), ...) \in \{0,1\}^\infty : \sum_i x(i) < \infty\right\}.$$

By (5) and Borel-Cantelli, $P^x(X_t \in S) = 1$ for all $x \in S$ and $t \geq 0$. It is not hard to conclude that $p_t(x,y) = P^x(X_t = y)$ satisfies (1) and (2). Letting $0 = (0,0,...)$, write

$$\begin{aligned} 1 - p_t(0,0) &\geq 1 - \prod_{i=0}^n \left[\frac{\delta(i)}{\delta(i)+\beta(i)} + \frac{\beta(i)}{\delta(i)+\beta(i)} e^{-t(\delta(i)+\beta(i))}\right] \\ &\sim 1 - \prod_{i=0}^n [1 - \beta(i)t] \geq 1 - \exp\left\{-t\sum_{i=0}^n \beta(i)\right\} \end{aligned}$$

as $t \downarrow 0$. It follows that $c(0) \geq \sum_i \beta(i)$, and therefore that 0 is instantaneous if $\sum_i \beta(i) = \infty$. One can then check that all states are instantaneous in this case, and that

$$P^0(X_t \in S \ \forall\ t \geq 0) = 0.$$

From now on, we will assume (4) and set $c(x) = -q(x,x)$. Then (3), with the initial condition $p_0(x,x) = 1$, can be rewritten successively as follows, letting $\delta(x,y) = 1$ if $x = y$ and 0 otherwise:

$$\frac{d}{dt}p_t(x,y) + c(x)p_t(x,y) = \sum_{z \neq x} q(x,z)p_t(z,y),$$

$$\frac{d}{dt}\left[e^{c(x)t}p_t(x,y)\right] = e^{c(x)t}\sum_{z \neq x} q(x,z)p_t(z,y),$$

$$e^{c(x)t}p_t(x,y) = \delta(x,y) + \int_0^t e^{c(x)s}\sum_{z \neq x} q(x,z)p_s(z,y)ds,$$

and finally

$$p_t(x,y) = \delta(x,y)e^{-c(x)t} + \int_0^t e^{-c(x)(t-s)}\sum_{z \neq x} q(x,z)p_s(z,y)ds. \tag{6}$$

Next, we will try to solve (6) by successive approximations. Note in what follows that it is very important that $q(x,y),\ x \neq y$, and $c(x)$ be nonnegative. Let

$$p_t^{(0)}(x,y) = \delta(x,y)e^{-c(x)t},$$

and then recursively

$$p_t^{(n+1)}(x,y) = \delta(x,y)e^{-c(x)t} + \int_0^t e^{-c(x)(t-s)}\sum_{z \neq x} q(x,z)p_s^{(n)}(z,y)ds.$$

The following facts are then easily proved by induction:

$$\begin{aligned} p_t^{(n)}(x,y) &\geq 0 \quad \text{for all } t,x,y,n,\\ \sum_y p_t^{(n)}(x,y) &\leq 1 \quad \text{for all } t,x,n,\\ p_t^{(n+1)}(x,y) &\geq p_t^{(n)}(x,y) \quad \text{for all } t,x,y,n. \end{aligned}$$

Now define

$$p_t^*(x,y) = \lim_{n\to\infty} p_t^{(n)}(x,y).$$

It follows that

$$\begin{aligned} p_t^*(x,y) &\geq 0 \quad \text{for all } t,x,y,\\ \sum_y p_t^*(x,y) &\leq 1 \quad \text{for all } t,x,\\ p_t^*(x,y) &\text{ satisfies (6), and hence (3).} \end{aligned}$$

The problem is that p_t^* may be substochastic, i.e., it may not satisfy equality in $\sum_y p_t^*(x,y) \leq 1$.

Here is the probabilistic interpretation of $p_t^{(n)}(x,y)$ and $p_t(x,y)$:

$$p_t^{(n)}(x,y) = P^x(X_t = y, \text{ having made at most } n \text{ jumps up to time } t),$$

and

$$p_t^*(x,y) = P^x(X_t = y, \text{ having made finitely many jumps up to time } t).$$

Example. Take $S = \{0, 1, ...\}$ and let $q(x,x) = -c(x), q(x,x+1) = c(x)$, where $c(x) > 0$. The chain spends an exponential time τ_x with parameter $c(x)$ at x before jumping to $x+1$. Therefore,

$$p_t^*(x,y) = P\left(\sum_{i=x}^{y-1} \tau_i < t < \sum_{i=x}^{y} \tau_i\right) \quad \text{for } x < y.$$

So, the chain explodes in finite time iff $\sum_i \tau_i < \infty$ a.s., i.e., iff $\sum_i 1/c(i) < \infty$.

We now have the following result.

Theorem 2. *(a) $p_t^*(x,y)$ is the minimal solution to the Kolmogorov backward equation in the sense that if $p_t(x,y)$ is any nonnegative solution of (3), then*

$$p_t(x,y) \geq p_t^*(x,y) \quad \textit{for all } t, x, y.$$

(b) If $p_t^*(x,y)$ is stochastic, then it is the unique solution to (1) and (3).

Proof. By induction, any nonnegative solution of (3) satisfies

$$p_t(x,y) \geq p_t^{(n)}(x,y) \quad \text{for all } t, x, y,$$

and therefore

$$p_t(x,y) \geq p_t^*(x,y).$$

For the uniqueness statement, sum this inequality on y.

Our next job is to use the previous ideas to construct the Markov chain X_t. Before that, we note the fundamental role played by the exponential distribution in the theory of continuous time Markov chains.

Proposition 1. *Suppose X_t is a Markov chain on S, and let*

$$\tau = \inf\{t > 0 : X_t \neq X_0\}.$$

Then τ is exponentially distributed, i.e.,

$$P^x(\tau > t) = e^{-c(x)t} \tag{7}$$

for some $0 \leq c(x) \leq \infty$.

Proof. By the Markov property, if $s < t$,

$$P^x(X_u = x \ \forall\ u \in [s,t] \mid \mathcal{F}_s) = P^{X_s}(\tau > t - s) \quad \textit{a.s. } P^x.$$

Multiply both sides of this identity by the indicator of the event $\{\tau > s\}$ and take expected values. The result is

$$P^x(\tau > t) = E^x[P^{X_s}(\tau > t-s), \tau > s)] = P^x(\tau > t-s)P^x(\tau > s).$$

It follows that $P^x(\tau > s)$ is of the form in (7).

To construct the Markov chain corresponding to the Q-matrix in (4), let $c(x) = -q(x,x)$, and then define transition probabilities for a *discrete time* Markov chain Z_n as follows. If $c(x) = 0$, set $p(x,x) = 1$. If $c(x) > 0$, put

$$p(x,y) = \begin{cases} \frac{q(x,y)}{c(x)} & \text{if } y \neq x \\ 0 & \text{if } y = x. \end{cases}$$

Take an initial distribution $\pi(x) > 0$ for Z_n, and then given $\{Z_n, n = 0, 1, ...\}$, let τ_n be (conditionally independent) exponentially distributed random variables with parameters $c(Z_n)$. This means that the joint distributions of the Z_n's and τ_n's are given by

$$P(Z_0 = x_0, ..., Z_n = x_n, \tau_0 > t_0, ..., \tau_n > t_n) =$$
$$\pi(x_0)p(x_0,x_1)\cdots p(x_{n-1},x_n)e^{-c(x_0)t_0}\cdots e^{-c(x_n)t_n}.$$

(Note that, unconditionally, the τ_k's are neither independent nor exponentially distributed.) Now let

$$N(t) = \begin{cases} \min\{n \geq 0 : \tau_0 + \cdots + \tau_n > t\} & \text{if } \sum_k \tau_k > t \\ \infty & \text{if } \sum_k \tau_k \leq t, \end{cases}$$

and $X_t = Z_{N(t)}$ if $N(t) < \infty$.

By conditioning on the time τ_1 of the first jump and the location Z_1, we see that

$$P^x(X_t = y, N(t) \leq n+1) = \delta(x,y)e^{-c(x)t} + \int_0^t c(x)e^{-c(x)s}\sum_{z\neq x} p(x,z)P^z(X_{t-s} = y, N(t) \leq n)ds.$$

This is the same recursion satisfied by $p_t^{(n)}(x,y)$, so it follows by induction that

$$p_t^{(n)}(x,y) = P(X_t = y, N(t) \leq n \mid X_0 = x).$$

Therefore

$$p_t^*(x,y) = P(X_t = y, N(t) < \infty \mid X_0 = x) \tag{8}$$

and

$$\sum_y p_t^*(x,y) = P(N(t) < \infty \mid X_0 = x).$$

By using the Markov property of $\{Z_n\}$ and the fact that, conditionally on $\{Z_n\}$, the τ_n's are exponentially distributed, one can check that X_t satisfies the Markov property. The main point is that, conditionally on $\{Z_n\}$, if one conditions on $N(t) = n$, one knows that n jump times have occurred in $[0,t]$, but does not know when they occurred in this interval. The forgetfulness property of the exponential distribution says that no matter when the last jump before t was, the amount of time after t until the next jump has the correct exponential distribution. The representation (8) then makes it straightforward to check that $p_t^*(x,y)$ satisfies the Chapman-Kolmogorov equation (2).

Theorem 3. *The minimal solution to the backward equation* $p_t^*(x,y)$ *is stochastic if and only if any of the following equivalent conditions is satisfied:*

(a) $P(N(t) < \infty) = 1$ *for all* t.

(b) $\sum_n \tau_n = \infty$ *a.s.*

(c) $\sum \frac{1}{c(Z_n)} = \infty$ *a.s.*

Proof. The equivalence of (a) and (b) and to the fact that $p_t^*(x,y)$ is stochastic is clear. For the equivalence of (b) and (c), use the following fact: If τ_k are independent exponentials with parameters λ_k, then $\sum \tau_k < \infty$ a.s. if $\sum_k 1/\lambda_k < \infty$ and $\sum \tau_k = \infty$ a.s. if $\sum_k 1/\lambda_k = \infty$. To see this, take $\sigma > 0$ and write

$$Ee^{-\sigma(\tau_1+\cdots+\tau_n)} = \prod_{k=1}^{n} \frac{\lambda_k}{\lambda_k + \sigma},$$

and note that this expression tends to zero as $n \to \infty$ if and only if $\sum_k 1/\lambda_k = \infty$. This implies that

$$P\left(\sum_n \tau_n = \infty \,\middle|\, \{Z_k, k \geq 0\}\right) = P\left(\sum_n \frac{1}{c(Z_n)} = \infty \,\middle|\, \{Z_k, k \geq 0\}\right) \quad a.s.$$

Now take expected values.

Remark. Note that the events in (b) and (c) of Theorem 3 may not be 0-1 events, since the summands are not independent. One can easily make examples in which Z_n has two different limiting directions with positive probability, and the series in (c) converges in one direction and diverges in the other.

Corollary 1. $p_t^*(x,y)$ *is stochastic if (a)* $\sup_x c(x) < \infty$ *or if (b)* Z_n *is recurrent.*

Example. The linear birth and death chain on $S = \{0, 1, ...\}$. Take $\beta, \delta > 0$ and put

$$q(x,y) = \begin{cases} \delta x & \text{if } y = x-1, \\ \beta x & \text{if } y = x+1, \\ -(\beta+\delta)x & \text{if } y = x. \end{cases}$$

Then $c(x) = (\beta + \delta)x$ and for $x \geq 1$,

$$p(x,y) = \begin{cases} \frac{\delta}{\beta+\delta} & \text{if } y = x-1, \\ \frac{\beta}{\beta+\delta} & \text{if } y = x+1, \\ 0 & \text{otherwise.} \end{cases}$$

In this case, the minimal solution to the backward equation is stochastic, since

$$\frac{Z_n}{n} \to \beta - \delta \quad a.s. \text{ on } \{Z_n \geq 1 \ \forall\ n\}.$$

Note that Corollary 1 does not apply if $\beta > \delta$, but still (c) of Theorem 3 holds since the harmonic series diverges. For the quadratic birth and death chain, in which x is replaced by x^2, the minimal solution is stochastic if and only if $\delta \geq \beta$.

To make the transition from Markov chains to Markov processes (with general state space), we define

$$T(t)f(x) = E^x f(X_t) = \sum_y p_t(x,y)f(y) \tag{9}$$

for bounded f. This has the following properties:

(a) $T(t)1 = 1$.
(b) $T(t)f \geq 0$ if $f \geq 0$.
(c) $||T(t)f|| \leq ||f||$, where $|| \cdot ||$ is the sup norm.
(d) The semigroup property $T(t+s) = T(t)T(s)$, which follows from the Chapman-Kolmogorov equation:

$$\begin{aligned} T(t)[T(s)f](x) &= \sum_y p_t(x,y)T(s)f(y) = \sum_y p_t(x,y) \sum_z p_s(y,z)f(z) \\ &= \sum_z p_{s+t}(x,z)f(z) = T(t+s)f(x). \end{aligned}$$

Formally differentiating (9) and using the backward equation (3) gives

$$\frac{d}{dt}T(t)f(x) = \mathcal{L}T(t)f(x),$$

where

$$\mathcal{L}f(x) = \sum_y q(x,y)f(y).$$

$\mathcal{L}$ is called the infinitesimal generator of the semigroup $T(t)$. Note that

$$|\mathcal{L}f(x)| \leq 2c(x)||f||,$$

so there is no reason to expect $\mathcal{L}$ to be a bounded operator.