Stat205B: Probability Theory (Spring 2003) **Lecture: 6**

Finite State Continuous Time Markov Chain

Lecturer: James W. Pitman *Scribe: Bo Li* `boli@stat.berkeley.edu`

See Durrett Sec 5.6 for the theory of discrete time recurrent Markov Chains with uncountable state space, as developed following Harris. The general idea is to recognize a suitable regenerative structure, like what happens to a discrete time, discrete space Markov chain each time it comes back to a point. Then decompose the path into blocks which are i.i.d. This idea can also be applied to continuous time, discrete space chains.

We now discuss a continuous time, discrete space Markov Chain, with time-homogeneous transition probabilities. Let S = state space. The theory is easiest if S is finite. Some aspects can be extended to S countable. The book of Karlin and Taylor [5], provides details for most of the following discussion. See also [3, 4] for further developments. General theory of countable state space, continuous time chains is very tricky: see Chung [1] and Freedman [2].

Suppose $(X_t, t \geq 0)$ is a process defined on $(\Omega, \mathcal{F}, \mathbb{P})$, with values in S which is finite. For each t, $\omega \to X_t(\omega)$ is a measurable map from $(\Omega, \mathcal{F}) \to S$. Look at the path: $t \to X_t(\omega)$ for fixed ω. In the finite state space case, we expect this path to be almost surely a step function, with only a finite number of jumps in any finite interval of time.
Make the convention that the path is right-continuous with left limits. Such a process has the time-homogeneous Markov property if

- conditionally given X_t, the processs $(X_s, 0 \leq s \leq t)$ and $(X_u, t \leq u < \infty)$ are independent;
- $(X_{t+v}, v \geq 0 | X_t = i)$ is distributed like $(X_v, v \geq 0)|X_0 = i)$.

Introduce the transition matrices $\mathbf{P}_t = \|P_t(i,j)\|_{i,j}$, where

$$P_t(i,j) = P(X_{s+t} = j | X_s = i), s \geq 0, i, j \in S$$

The definition of $\mathbf{P}_t$ and the time-homogeneous Markov property yield:

- $P_t(i,j) \geq 0$
- $\sum_{j \in S} P_t(i,j) = 1$
- the semi-group property (Chapman-Kolmogorov equation): $\mathbf{P}_s \mathbf{P}_t = \mathbf{P}_{s+t}$

Right-continuous paths make $X_t \to X_0, a.s.$ as $t \to 0^+$, which implies

$$\lim_{t \to 0^+} \mathbf{P}_t = \mathbf{I}$$

(the identity matrix). Combining with the semigroup property, we know

$$\lim_{r \to t^+} \mathbf{P}_r = \lim_{s \to 0^+} \mathbf{P}_{t+s} = \lim_{s \to 0^+} \mathbf{P}_s \mathbf{P}_t = \mathbf{I}\mathbf{P}_t = \mathbf{P}_t.$$

Thus $\mathbf{P}_t$ is a right continuous function of t. In fact, $\mathbf{P}_t$ is not only right continuous but also continuous and even differentiable. Accepting this, let

$$\mathbf{Q} = \frac{d}{dt}\mathbf{P}_t|_{t=0}$$

The semi-group property easily implies the following backwards equations and forwards equations:

$$\frac{d}{dt}\mathbf{P}_t = \mathbf{Q}\mathbf{P}_t = \mathbf{P}_t\mathbf{Q}$$

Hence there is representation:

$$\mathbf{P}_t = \exp(\mathbf{Q}t) = \mathbf{I} + \mathbf{Q}t + \mathbf{Q}^2t^2/2! + \ldots$$

In particular,

$$P_t(i,j) = 1_{(i=j)} + Q(i,j)t + o(t) \quad as \quad t \to 0^+$$

Note that $P_t(i,j) \geq 0$, so

$$Q(i,j) \geq 0 \quad for \quad j \neq i$$

And $\sum_{j\in S} P_t(i,j) = 1$ implies

$$\sum_{j\in S} Q(i,j) = 0.$$

Let

$$q_i := -Q(i,i) = \sum_{j\neq i} Q(i,j) \geq 0.$$

Let J_r denote time of the rth jump. By the Markov property, J_1 has the memoryless property

$$\mathbb{P}_i(J_1 > s+t|J_1 > s) = \mathbb{P}_i(J_1 > t)$$

Notice

$$\begin{array}{rcl}
\frac{d}{dt}\mathbb{P}_i(J_1 > t)|_{t=0} & = & \frac{\mathbb{P}_i(t<J_1\leq t+dt)}{dt}|_{t=0} \\
& = & \sum_{j\neq i} \frac{\mathbb{P}_i(X_t=i, X_{t+dt}=j)}{dt}|_{t=0} \\
& = & \sum_{j\neq i} P_t(i,i)\frac{P_{dt}(i,j)}{dt}|_{t=0} \\
& = & \sum_{j\neq i} Q(i,j) \\
& = & -q_i
\end{array}$$

Hence

$$\mathbb{P}_i(J_1 > t) = e^{-q_i t} \qquad (t \geq 0).$$

That is, the $\mathbb{P}_i$ distribution of J_1 is exponential(q_i). Note that $q_i = 0$ means i is absorbing: $P_t(i,i) = 1$ for all t. .
Now assume $q_i > 0$. Let

$$\widehat{p}(i,j) := \{ {}^{Q(i,j)/q_i \quad j\neq i}_{0 \qquad\quad otherwise}$$

Then

$$\sum_{j\neq i} \widehat{p}(i,j) = 1$$

so $\widehat{p}$ is a transition probability matrix. From the exponential(q_i) distribution of J_1,

$$\mathbb{P}_i(X \text{ first leaves } i \text{ in } (t,t+dt)|X_s = i, 0 \leq s \leq t) = q_i dt$$

Hence, for $j \neq i$

$$\mathbb{P}_i(X \text{ jumps to } j \text{ in } (t,t+dt)|X_s = i, 0 \leq s \leq t) = P_i(X_{J_1} = j)q_i dt$$

On the other hand

$$\mathbb{P}_i(X \text{ jumps to } j \text{ in } (t, t+dt) | X_s = i, 0 \le s \le t) = P_{dt}(i,j) = Q(i,j)dt$$

Comparing the last two facts

$$\mathbb{P}_i(X_{J_1} = j) = Q(i,j)/q_i \quad for\ j \neq i.$$

Set $J_0 = 0$, similar arguments enable us to get for $r = 0, 1, 2, \ldots$

$$P(X_{J_{r+1}} = j | X_{J_r} = i) = \widehat{p}(i,j)$$

Starting from i, $(X_0, X_{J_1}, X_{J_2}, \ldots)$ is a discrete time Markov chain with transition matrix $\widehat{p}$, which is called the *embedded jump chain.* Moreover, conditionally given $X_0 = i_0, X_{J_1} = i_1, X_{J_2} = i_2, \ldots$, the holding times $J_1, J_2 - J_1, J_3 - J_2, \ldots$ are independent exponential variables with parameters $q_{i_0}, q_{i_1}, q_{i_2}, \ldots$. And given any matrix Q with non-negative off-diagonal elements and row sums identically zero, we can construct a Markov chain with semigroup $\mathbf{P}_t = \exp(\mathbf{Q}t)$ as such a hold-jump process. Say, the chain starts from i_0, it stays at i_0 for a period of time with exponential(q_{i_0}) distribution. Then it jumps to another point i_1 with probability $\widehat{p}(i_0, i_1)$. And stays at i_1 for a period of time with exponential(q_{i_1}) distribution, then jumps to i_2 with probability $\widehat{p}(i_1, i_2)$. And so on.

Provided Q is bounded or not too badly unbounded, this construction also makes sense for infinite S. Here are some examples:

Example 1: Poisson process. $S = 0, 1, 2, \ldots$.

$$\mathbf{Q} = \begin{pmatrix} -\lambda & \lambda & 0 & \ldots \\ 0 & -\lambda & \lambda & \ldots \\ 0 & 0 & -\lambda & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

$X =$ Poisson process with rate λ, and $P_t(i,j) = 1_{(j \ge i)} e^{-\lambda t} (\lambda t)^{j-i}/(j-i)!, \quad i, j \ge 0.$

Example 2: Birth and death process

$$\mathbf{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \ldots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & \ldots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Here, λ_is are birth rates, μ_is are death rates. And $\lambda_i + \mu_i$ must not increase too rapidly.

Example 3: Pure birth process
Now $\lambda_i > 0$, $\mu_i = 0$. The mean holding time of state i is $1/\lambda_i$. What if $\sum_i 1/\lambda_i < \infty$? Note $E_0 J_n = \frac{1}{\lambda_1} + \ldots + \frac{1}{\lambda_n}$ increases and is bounded. So $J_n \uparrow J_\infty < \infty$ *a.s.*, with $E_0 J_\infty = \sum_i 1/\lambda_i < \infty$. This phenomenon is called explosion. There is not a unique way to bring the process back from ∞: it could jump back to $0, 1, 2, \ldots$ according to any probability distribution over these states, then explode again,come back again independently, and so on. Note that the paths of the process can now have infinitely many jumps in finite time. And the way the process comes back from ∞ is not determined by the Q matrix. Much more subtle behaviour is possible with countable state space (Markov chains with *instantaneous states*).

Returning to the case with S finite. Consider limit behavior of the chain as $t \to \infty$. First observe the following are equivalent:

- $\mathbf{P}_t$ is irreducible for some $t > 0$
- $\widehat{p}$, transition matrix of the embedded jumping chain, is irreducible
- $P_t(i, j) > 0$ for all $t > 0$, $i, j \in S$

These conditions imply that $\mathbf{P}_t$ is aperiodic. Moreover, if $\mathbf{P}_t$ is positive recurrent, there exists a unique stationary distribution π so that

$$\lim_{t\to\infty} P_t(i, j) = \pi_i$$

The semi-group property implies that for each t

$$\pi \mathbf{P}_t = \pi$$

Or

$$\pi\left(\frac{\mathbf{P}_t - \mathbf{I}}{t}\right) = 0.$$

Let $t \to 0^+$ to see that

$$\pi \mathbf{Q} = 0.$$

Thus π is determined by the system of linear equations $\pi\mathbf{Q} = 0$ and $\pi\mathbf{1} = 1$.

References

[1] K.L. Chung. *Markov Chains with Stationary Transition Probabilities.* Springer–Verlag, second edition, 1967.

[2] D. Freedman. *Markov Chains.* Springer–Verlag, 1983. Reprint of 1971 Holden-Day edition.

[3] S. Karlin and H.M. Taylor. *A First Course in Stochastic Processes.* Academic Press, second edition, 1975.

[4] S. Karlin and H.M. Taylor. *A Second Course in Stochastic Processes.* Academic Press, 1981.

[5] H.M. Taylor and S. Karlin. *An Introduction to Stochastic Modeling (revised ed.).* Academic Press, Boston, 1994.