## Lecture 13. Continuous-time Markov chains. The limiting behavior of the transition probabilities.

I want to wrap up the theme of continuous-time Markov chains, so the things in this lecture will be given mostly without proofs: for some of them it will be shown how to get the proofs (in particular, modifying those in the discrete case), and some without any proof whatsoever.

We want to have for continuous-time Markov chains everything that we had for discrete chains. Start with Lecture 3.

We considered random variables $\tau_y$ defined by (3.25). In the case of continuous time we cannot, in general, consider $\min\{t > 0 \colon \xi_t = y\}$: in general we must take the $\inf$. But, of course, $\inf\{t > 0 \colon \xi_t = y\} = \min\{t \geq 0 \colon \xi_t = y\}$. So we define

$$\tau_y = \begin{cases} \min\{t \geq 0 \colon \xi_t = y\} & \text{if there are such } t, \\ \infty & \text{if there are no such } t; \end{cases} \tag{13.1}$$

only if we start at the point $y$, this $\tau_1 = 0$ and is of no use to us. We can try to introduce something more useful; in particular, we can consider the random variable

$$\tau_{1y} = \begin{cases} \min\{t \geq \tau_1 \colon \xi_t = y\} & \text{if there are such } t, \\ \infty & \text{if there are no such } t \colon \end{cases} \tag{13.2}$$

the first time of going to $y$ (returning to $y$ if $\xi_0 = y$) after the first change of state.

Let us write the analogues of formula (3.31).

For $x \neq y$ we can reach the point $y$ for the first time at time $\tau_y \leq t$, and there will remain the time $t - \tau_y$ in which we are to go from the point $y$ to the same point (not necessarily for the first time – and *never* for the first time, because by the right-continuity there exists a whole interval to the right of the time point $\tau_y$ we have $\xi_t = y$). The formula similar to (3.31):

$$p(t,\, x,\, y) = E_x[I_{\{\tau_y \leq t\}} \cdot p(t - \tau_y,\, y,\, y)]. \tag{13.3}$$

Formula (13.3) is true also for $x = y$; but since for $\xi_0 = y$ we have $\tau_y = 0$, this formula reduces to $p(t,\, y,\, y) = p(t,\, y,\, y)$: true, of course – but not very useful. Let us write for $x = y$ a different formula.

For $\xi_0 = y$ the event $\{\xi_t = y\}$ can occur in two ways: either $\tau_1 > t$, the process remaining at the state $y$ for the whole time interval $[0,\, t]$; or we leave this state at time $\tau_1 < t$, return to it for the first time at the time $\tau_{1y} \leq t$, and then we have the time $t - \tau_{1y}$ remaining to go from $y$ to the same $y$. So we right:

$$p(t,\, y,\, y) = P_y\{\tau_1 > t\} + E_y[I_{\{\tau_{1y} \leq t\}} \cdot p(t - \tau_{1y},\, y,\, y)]. \tag{13.4}$$

The arguments that we gave before writing formulas (13.3), (13.4) are not their *proofs*: we prove them writing the corresponding formulas for the discrete Markov chain $\xi_0$, $\xi_{1/2^n}$, $\xi_{2/2^n}$, $\xi_{3/2^n}$, ... and the random moments $\tau_y^n$, $\tau_1^n$, and $\tau_{1y}^n$, and taking the limit as $n \to \infty$.

Introducing the distributions $\mu_{x;\,\tau_y}$, $\mu_{y;\,\tau_{1y}}$ of the random variables $\tau_y$ with respect to the probability $P_x$ and of the random variable $\tau_{1y}$ with respect to $P_y$, and knowing the distribution of the random variable $\tau_1$ (the exponential one), we can rewrite (13.3), (13.4) as

$$p(t,\,x,\,y) = \int_{[0,\,t]} \mu_{x;\,\tau_y}(ds)\;p(t-s,\,y,\,y), \tag{13.5}$$

$$p(t,\,y,\,y) = e^{-v_y t} + \int_{[0,\,t]} \mu_{y;\,\tau_{1y}}(ds)\;p(t-s,\,y,\,y). \tag{13.6}$$

Note that the random variables $\tau_1$ and $\tau_{1y}-\tau_1$ are independent with respect to the probability $P_y$ (proved the same way as the conditional independence of $\tau_1$, $\tau_2-\tau_1$, ...: by proving independence of $\tau_1^n$ and $\tau_{1y}^n-\tau_1^n$ and taking $n\to\infty$). If $\mu$ is the distribution of the random variable $\tau_{1y}-\tau_1$ with respect to the probability measure $P_y$, we can write for $C\subseteq[0,\,\infty)$:

$$\begin{aligned}\mu_{y;\,\tau_{1y}}(C) = P_y\{\tau_1+(\tau_{1y}-\tau_1)\in C\} &= \iint\limits_{s,\,u>0,\;s+u\in C} v_y e^{-v_y s}\,ds\,\mu(du)\\ &= \int_C \Big[\int_0^t v_y e^{-v_y(t-u)}\,\mu(du)\Big]\,dt.\end{aligned} \tag{13.7}$$

We see that the distribution $\mu_{y;\,\tau_{1y}}$ has a density

$$p_{y;\,\tau_{1y}}(t) = \begin{cases} 0, & u\le 0,\\ \displaystyle\int_0^t v_y e^{-v_y(t-u)}\,\mu(du), & t>0; \end{cases} \tag{13.8}$$

this density is bounded by the constant $v_y$.

Now we go to Lecture 4. In that lecture we considered generating functions $P_{xy}(s) = \sum_n p_{xy}^{(n)}\cdot s^n$, $F_{xy}(s) = \sum_k f_{xy}^{(k)}\cdot s^k$. In the continuous-time case we can consider the corresponding integrals $\int p(t,\,x,\,y)\cdot s^t\,dt$; but, traditionally, we take $e^{-\lambda}$ instead of $s$, and consider the *Laplace transforms.*

We consider Laplace transforms of functions $a(t)$, $t\ge 0$, and measures $\beta$ on $[0,\,\infty)$ (of course, on the Borel $\sigma$-algebra $\mathcal{B}_{[0,\,\infty)}$):

$$A(\lambda) = \int_0^\infty a(t)\,e^{-\lambda t}\,dt \tag{13.9}$$

or

$$B(\lambda) = \int_{[0,\,\infty)} e^{-\lambda t}\,\beta(dt). \tag{13.10}$$

If we denote $P_{xy}(\lambda)$ the Laplace transform of the transition probability $p(t,\,x,\,y)$, and

$$M_{xy}(\lambda) = \int_{[0,\,\infty)} e^{-\lambda t}\,\mu_{x;\,\tau_y}(dt), \qquad M_{yy}^1(\lambda) = \int_{[0,\,\infty)} e^{-\lambda t}\,\mu_{x;\,\tau_{1y}}(dt) \tag{13.11}$$

(the functions $M_{xy}(\lambda)$, $M^1_{yy}(\lambda)$ are finite at least for $\lambda \geq 0$, the function $P_{xy}(\lambda)$ for $\lambda > 0$), the formulas (13.5), (13.6) are rewritten as

$$P_{xy}(\lambda) = M_{xy}(\lambda) \cdot P_{yy}(\lambda), \tag{13.12}$$

$$P_{yy}(\lambda) = \int_0^\infty e^{-\lambda t} \cdot e^{-v_y t}\, dt + M^1_{yy}(\lambda) \cdot P_{yy}(\lambda) = \frac{1}{v_y + \lambda} + M^1_{yy}(\lambda) \cdot P_{yy}(\lambda). \tag{13.13}$$

$$P_{yy}(\lambda) = \frac{1/(v_y + \lambda)}{1 - M^1_{yy}(\lambda)}, \qquad M^1_{yy}(\lambda) = 1 - \frac{1/(v_y + \lambda)}{P_{yy}(\lambda)} \tag{13.14}$$

(I don't stop at such things as splitting $e^{-\lambda t}$ into $e^{-\lambda s} \cdot e^{-\lambda(t-s)}$, making a substitution and changing the order of integration). Taking the limit as $\lambda \to 0^+$ (as we took $\lim_{s \to 1^-}$ in Lecture 4), we get:

$$\int_0^\infty p(t,\, x,\, y)\, dt = P_x\{\tau_y < \infty\} \cdot \int_0^\infty p(t,\, y,\, y)\, dt, \tag{13.15}$$

$$\int_0^\infty p(t,\, y,\, y)\, dt = \frac{1}{P_y\{\tau_{1y} = \infty\}}, \qquad P_y\{\tau_{1y} = \infty\} = \frac{1}{\displaystyle\int_0^\infty p(t,\, y,\, y)\, dt}. \tag{13.16}$$

Now we call a state $y \in X$ recurrent if, wherever we go first from the point $y$, we return to it with probability 1; otherwise we call $y$ transient.

If a state $y$ is transient, then almost surely $\xi_t \neq y$ for sufficiently large $t$. If $y$ is recurrent, almost surely (with respect to the probability $P_y$) there exist arbitrarily large $t$ for which $\xi_t = y$.

**Theorem 13.1.** *A state $y \in X$ is recurrent if and only if* $\int_0^\infty p(t,\, y,\, y)\, dt = \infty$ (and transient if and only if this integral converges).

**Proof:** follows from formula (13.16).

From this we see at once that if $\lim_{t\to\infty} p(t,\, y,\, y)$ exists and is positive, the state $y$ is recurrent; and if $y$ is transient, then $\lim_{t\to\infty} p(t,\, y,\, y) = 0$.

This is a little more complicated than for the discrete-time case, because if a series converges, its $n$-th term necessarily goes to 0, while an integral from 0 to $\infty$ may converge, and the integral have no limit as $t \to \infty$. But we remember that the function $p(t,\, y,\, y)$ is *uniformly continuous* on $[0,\, \infty)$ (see Theorem 12.4), and for such functions if their integral converges, they go to 0 at $\infty$.

Analogue of Theorem 5.1:

**Theorem 13.2.** *If there exists* $\lim_{t\to\infty} p(t,\, y,\, y) = p(y)$, *then this limit is equal to* $\dfrac{1}{v_y \cdot E_y \tau_{1y}}$.

**Proof:** It follows from (13.14): $P(\lambda) = \dfrac{p(y)}{\lambda} + o\big(\dfrac{1}{\lambda}\big)$ as $\lambda \to 0^+$, and $E_y \tau_{1y} = \lim_{\lambda\to 0^+} \dfrac{1 - M_{y;\,\tau_{1y}}(\lambda)}{\lambda} = \lim_{\lambda\to 0^+} \dfrac{1/(v_y+\lambda)}{\lambda \cdot [p(y)/\lambda + o(1/\lambda)]} = \dfrac{1}{v_y \cdot p(y)}$.

Now the analogue of Theorem 5.2:

**Theorem 13.3.** *The limit* $\lim_{t\to\infty} p(t,\, x,\, y)$ *exists for every* $y \in X$.

You see, *periodicity* is impossible for continuous time and discrete space, so this theorem is formulated simpler than Theorem 5.2.

The **proof** is essentially the same as that of Theorem 5.2, with integrals instead of sums, and use of formula (13.6) instead of (4.5). The fact is used that the distribution $\mu_{y;\,\tau_{1y}}$ has the density (13.8) that is strictly positive for sufficiently large $u$ (for $u \ge u_0$, where $u_0$ is such that $\mu(0,\, u_0] > 0$, and $\mu$ is the distribution with respect to $P_y$ of the random variable $\tau_{1y} - \tau_1$).

Theorem 6.1 is reformulated for continuous-time Markov chains with changing $n$ to $t$ and replacing sums with integrals; Theorem 6.2 is preserved with the same formulation. We can introduce equivalence classes; we don't need any analogue of Theorem 7.1 (no periodicity is possible), and we easily prove the reformulated Theorems 7.2 – 7.5. Problems $\boxed{\mathbf{9}}$–$\boxed{\mathbf{12}}$, $\boxed{\mathbf{14}}$ are reformulated too, introducing the requirement of almost surely only finite number of jumps in any finite time interval: $P_x\{\lim_{n\to\infty} \tau_n = \infty\} = 1$. In particular,

*For a continuous-time Markov chain with one equivalence class of states satisfying the condition* $P_x\{\lim_{n\to\infty} \tau_n = \infty\} = 1$, *and* $q(y)$, $y \in X$, *is a nonnegative solution of the system* $\sum_x q(x) \cdot a_{xy} = 0$, $y \in X$, *and* $0 < \sum_y q(y) < \infty$, *then* $\lim_{t\to\infty} p(t,\, x,\, y) = q(y)/\sum_y q(y)$;

*If* $\sum_y q(y) = \infty$, *then* $\lim_{t\to\infty} p(t,\, x,\, y) = 0$.

Let us consider our example of a two-pump gas station with a queue not longer than 2, with exit densities $v_0 = \lambda$, $v_1 = \lambda + \mu$, $v_2 = v_3 = \lambda + 2\mu$, $v_4 = 2\mu$ and transition densities $a_{01} = \lambda$, $a_{12} = a_{23} = a_{34} = \lambda$, $a_{10} = \mu$, $a_{21} = a_{32} = a_{43} = 2\mu$. The equations $\sum_x q(x) \cdot a_{xy} = 0$ are:

$$\begin{cases} 0 = -\,\lambda q(0) + \mu q(1), \\ 0 = \lambda q(0) - (\mu + \lambda) q(1) + 2\mu q(2), \\ 0 = \lambda q(1) - (2\mu + \lambda) q(2) + 2\mu q(3), \\ 0 = \lambda q(2) - (2\mu + \lambda) q(3) + 2\mu q(4), \\ 0 = \lambda q(3) - 2\mu q(4). \end{cases} \tag{13.17}$$

From the 0-th equation we get $q(1) = \dfrac{\lambda}{\mu}\, q(0)$; in the next equation the terms $-\,\lambda q(0)$ and $\mu q(1)$ cancel, and $-\,\lambda q(1) + 2\mu q(2)$ remain, so $q(2) = \dfrac{\lambda}{2\mu}\, q(1) = \dfrac{\lambda^2}{2\mu^2}\, q(0)$. Proceeding lime this, we express all $q(y)$ in terms of $q(0)$:

$$q(1) = \frac{\lambda}{\mu}\, q(0), \quad q(2) = \frac{\lambda^2}{2\mu^2}\, q(0), \quad q(3) = \frac{\lambda^3}{4\mu^3}\, q(0), \quad q(4) = \frac{\lambda^4}{8\mu^4}\, q(0), \tag{13.18}$$

from which we get:

$$p(0) = \lim_{t\to\infty} p(t,\, x,\, 0) = \frac{1}{1 + \lambda/\mu + \lambda^2/2\mu^2 + \lambda^3/4\mu^3 + \lambda^4/8\mu^4}, \tag{13.19}$$

etc.

In particular, for $\lambda = 3$, $\mu = 2$ we have (up to 0.001:

$$p(0) = 0.196, \quad p(1) = 0.294, \quad p(2) = 0.221, \quad p(3) = 0.165, \quad p(4) = 0.124. \tag{13.20}$$