Lecture 11. The Poisson process. Markov processes in queueing theory. Kolmogorov equations.

We are considering an example started in Lecture 10, with $v_x = a_{x,x+1} = \lambda$.

We know that the distribution of the sum of n independent random variables having the same exponential distribution with parameter λ is the Γ-distribution with parameters n, λ. This means that the distribution of the random variable τ_n has the probability density

$$p_{\tau_n}(t) = \begin{cases} 0, & t \le 0, \\ \dfrac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t}, & t > 0 \end{cases} \tag{11.1}$$

(the distribution is the same independently with respect to which probability P_x we take it). So we have for $n > 0$:

$$p(t, x, x+n) = \int_0^t \frac{\lambda^n s^{n-1}}{(n-1)!} e^{-\lambda s}\, ds - \int_0^t \frac{\lambda^{n+1} s^n}{n!} e^{-\lambda s}\, ds = \frac{\lambda^n}{n!} \int_0^t (n s^{n-1} - \lambda s^n) e^{-\lambda s}\, ds. \tag{11.2}$$

The integrand is equal to $\dfrac{d}{ds}(s^n e^{-\lambda s})$, so

$$p(t, x, x+n) = \frac{\lambda^n t^n e^{-\lambda t}}{n!}. \tag{11.3}$$

For $n = 0$ (taking into account that $\tau_0 = 0$, $P_x\{\tau_0 \le t\} = 1$ for $t \ge 0$:

$$p(t, x, x) = 1 - P_x\{\tau_1 \le t\} = 1 - \int_0^t \lambda e^{-\lambda s}\, ds = e^{-\lambda t}; \tag{11.4}$$

i. e., formula (11.3) is true also for $n = 0$.

The entries of the transition matrix P^t are given by formula (11.3) for $n \ge 0$, and $p(t, x, y) = 0$ for $y < x$. It is easy to check that for every $t \ge 0$ this is a stochastic matrix, $P^0 = 0$, and $P^s \cdot P^t = P^{s+t}$ (the last equality is equivalent to the statement that the sum of two independent random variables having Poisson distributions with parameters λs and λt has the Poisson distribution with parameter $\lambda(s+t)$).

It is easy to check that the matrix $A = \dfrac{d^+}{dt} P^t\Big|_{t=0}$ is the one that we started with: $-\lambda$'s on the main diagonal, and λ's on the first diagonal above it.

Equality (11.3) means that the difference $\xi_t - \xi_0$ (and also every difference $\xi_{s+t} - \xi_s$) has the Poisson distribution with parameter λt. This is why the Markov process ξ_t is called *the Poisson process.*

At this time I would like to introduce the terms for the quantities a_{xy} and $v_x = -a_{xx}$. They are (right-hand) *derivatives*, i. e. rates of change; so I'll call v_x the *exit rate* (from the state x), and a_{xy}, $y \ne x$, the *transition rates.*

One more example.

Continuous-time Markov chains have applications in what is called *Queueing Theory* (*queue* is the French-British word for a *line*, used traditionally in the Queueing Theory).

Let us consider a serving system (say, a gas station). The state of the system at a time t can be characterized by the number ξ_t of customers in the system at this time.

To be concrete, let us suppose that it is a gas station with two pumps `(in the lecture it was` *one* `pump)`, and the cars that arrive while both pumps are busy form a queue of length not more than 2 (if a car drives up and finds both pumps busy and both places in the queue occupied, it drives away). The state space (the set of all possible states) here is $X = \{0, 1, 2, 3, 4\}$.

Suppose the process starts at time $t = 0$. Let $U_1 > 0$ be the time before the first new customer comes (*new* means that it is possible that there was some number of customers in the system at time $t = 0$; so we are speaking of the next customer to come); U_2 the time between the coming of the first new customer and the next one; U_3, between the second customer and the third; etc. (these U_k can be described as *interarrival times*; I seem to run out of Greek letters for denoting random variables, so I am using capital Roman ones). Of course, if we don't suppose that the distribution of these random variables is exponential, we'll never get a Markov process. Suppose that the random variables U_k have an exponential distribution with parameter λ.

Now, there are *service times*: for the first pump, the first service time S_{11} (from the time point 0 if the system was busy at this time, or from the time when it starts to be busy if ξ_0 was equal to 0, to the time at which the service of the first customer is finished); the second service time S_{12}, etc.; and S_{21}, S_{22}, S_{23}, ... for the second pump. Again, we have no chance of obtaining a continuous-time Markov chain unless we suppose that the service times S_{ik} are exponential. Let them be exponential with parameter μ.

And let all random variables U_1, U_2, ..., U_n, ..., S_{11}, S_{12}, ..., S_{1n}, ..., S_{21}, S_{22}, ..., S_{2n}, ..., be independent.

What are the exit rates v_x and the transition rates a_{xy} for the Markov chain X_t describing the number of customers in this system?

Note that the problem of finding the *transition probabilities* $p(t, x, y), 0 \le t < \infty$, $x, y = 0, 1, 2, 3, 4$, is much more complicated than that of finding the infinitesimal (differential) characteristics a_{xy}.

First of all, transitions from one state to another occur at the times at which either a new customer comes (and then we move from the state x, which is equal to 0, 1, 2, or 3, to $x + 1$); or at times at which serving a customer is finished (and at such times we go from a state x to $x - 1$; note that there is no transition to the left from the point $x = 0$ – because no customer is being served if there are no customers in the system). The possible transitions can be shown as follows:

$$0 \rightleftarrows 1 \rightleftarrows 2 \rightleftarrows 3 \rightleftarrows 4 \,. \tag{11.5}$$

So we have $a_{xy} = 0$ if $|y - x| > 1$.

Transitions from the point $x = 0$ are possible only to the state $y = 1$; the time before this happens is equal to the random variable U_1, whose distribution is exponential with parameter λ: so the first state change time $\tau_1 = U_1$ if $\xi_0 = 0$; so

$$v_0 = a_{01} = \lambda. \tag{11.6}$$

If we are at a state $x > 0$ at time $t = 0$, the situation is more complicated: if a new customer comes – at the time U_1 – *before* the service of the customer(s) currently being served is finished – at time S_{11} or S_{21} – then the system jumps to the state $x+1$; and if S_{11} or S_{21} is less than U_1, the customer that has been served leaves the system, and we go to the state $x-1$.

If $\xi_0 = x = 1$, only one customer is being served at time $t = 0$; suppose it (the customer = a car) is being served at pump number 1. Then we have:

$$\begin{cases} \tau_1 = U_1, \quad \xi_{\tau_1} = x+1 & \text{if } U_1 < S_{11}, \\ \tau_1 = S_{11}, \quad \xi_{\tau_1} = x-1 & \text{if } S_{11} < U_1. \end{cases} \tag{11.7}$$

We may think also of the case of $U_1 = S_{11}$, when the first service is ended *exactly* at the time that the first new customer comes (so the system remains in the same state at this time); but we have a feeling that it is practically impossible, and we'll see that it can be *proved* that $P\{U_1 = S_{11}\} = 0$.

So if $U_1 \neq S_{11}$, we have:

$$\tau_1 = \min(U_1,\, S_1). \tag{11.8}$$

We want to find the distribution of this random variable (if we discover that it is an exponential distribution, its parameter will be the value of v_x); and find also the probabilities of the two alternatives in (11.7) (and these will be $a_{x,\,x+1}/v_x$ and $a_{x,\,x-1}/v_x$). To do this, let us write the joint density of the random variables U_1, S_{11}: since they are independent, we have

$$f_{U_1,\,S_{11}}(u,\,s) = f_{U_1}(u)\cdot f_{S_{11}}(s) = \begin{cases} 0 & \text{if } u \le 0 \text{ or } s \le 0, \\ \lambda\mu\, e^{-\lambda u - \mu s} & \text{if } u,\, s > 0. \end{cases} \tag{11.9}$$

First of all, the distribution of the random variable (11.8) is clearly continuous. To find its density, the standard way is to find the cumulative distribution function $F_{\min(U_1,\,S_{11})}$ first:

$$F_{\min(U_1,\,S_{11})}(b) = P\{\min(U_1,\,S_1) \le b\} = \begin{cases} 0 & \text{for } b < 0, \\ \displaystyle\iint_{\min(u,\,s)\le b} f_{U_1,\,S_{11}}(u,\,s)\;du\,ds & \text{for } b \ge 0. \end{cases} \tag{11.10}$$

Since the integrand here is equal to 0 except for u and s positive, and given by the second formula (11.9) if they are positive, the integral in (11.10) is equal to

$$\iint_{0<\min(u,\,s)\le b} \lambda\mu\, e^{-\lambda u - \mu s}\;du\,ds. \tag{11.11}$$

Make a picture and show in it the angular-shaped range of integration $\{(u,\,s) :\ 0 < \min(u,\,s) \le b\}$.

It is easier to find this probability through the probability of the opposite event: the integral (11.11) is equal to

$$1 - \iint\limits_{\min(u,\, s) > b} \lambda\mu\, e^{-\lambda u - \mu s}\; du\, ds = 1 - \int_b^\infty \Big[\int_b^\infty \lambda\mu\, e^{-\lambda u - \mu s}\; ds\Big]\; du$$
$$= 1 - \int_b^\infty \lambda\, e^{-\lambda u} \cdot e^{-\mu b}\; du = 1 - e^{-(\lambda+\mu)b}. \tag{11.12}$$

Differentiating the cumulative distribution function given by formulas (11.10),(11.12), we obtain:

$$f_{\min(U_1,\, S_{11})}(t) = \begin{cases} 0, & t \le 0, \\ (\lambda+\mu)\, e^{-(\lambda+\mu)t}, & t > 0; \end{cases} \tag{11.13}$$

so the distribution is exponential with parameter $\lambda + \mu$, and we have:

$$v_1 = \lambda + \mu, \qquad a_{11} = -\lambda - \mu. \tag{11.14}$$

As for the probabilities of the two alternatives in (11.7), we have:

$$P\{U_1 < S_{11}\} = \iint\limits_{u<s} f_{U_1,\, S_{11}}(u,\, s)\; du\, ds = \iint\limits_{0<u<s} \lambda\mu\, e^{-\lambda u - \mu s}\, du\, ds$$
$$= \int_0^\infty \Big[\int_u^\infty \lambda\mu\, e^{-\lambda u - \mu s}\; ds\Big]\; du = \int_0^\infty \lambda\, e^{-(\lambda+\mu)u}\, du = \frac{\lambda}{\lambda+\mu}, \tag{11.15}$$

$$P\{S_{11} < U_1\} = \iint\limits_{0<s<u} \lambda\mu\, e^{-\lambda u - \mu s}\, du\, ds$$
$$= \int_0^\infty \Big[\int_s^\infty \lambda\mu\, e^{-\lambda u - \mu s}\; du\Big]\; ds = \int_0^\infty \mu\, e^{-(\lambda+\mu)s}\, ds = \frac{\mu}{\lambda+\mu}; \tag{11.16}$$

so

$$\frac{a_{12}}{v_1} = \frac{\lambda}{\lambda+\mu}, \qquad \frac{a_{10}}{v_1} = \frac{\mu}{\lambda+\mu}, \tag{11.17}$$

and

$$a_{12} = \lambda, \qquad a_{10} = \mu. \tag{11.18}$$

As for the promised equality $P\{U_1 = S_{11}\} = 0$, here is its proof:

$$P\{U_1 = S_{11}\} = \iint\limits_{u=s} f_{U_1,\, S_{11}}(u, s)\; du\, ds = 0, \tag{11.19}$$

because this is an two-dimensional integral over a *line*, which has two-dimensional area equal to 0.

If the initial point $\xi_0 = x > 2$ or 3, then both pumps are serving their cars at time $t = 0$, and we have to consider the minimum

$$\tau_1 = \min(U_1,\, S_{11},\, S_{12}); \tag{11.20}$$

the same kind of calculations, with a triple integral instead of double, leads us to

$$v_x = \lambda + 2\mu, \qquad \pi_{x,\,x+1} = \frac{\lambda}{\lambda + 2\mu}, \quad \pi_{x,\,x-1} = \frac{2\mu}{\lambda + 2\mu}, \tag{11.21}$$

$$a_{x,\,x+1} = \lambda, \qquad a_{x,\,x-1} = 2\mu. \tag{11.22}$$

Finally, if $\xi_0 = x = 4$, the number of customers in the system cannot increase, $\tau_1 = \min(S_{11},\, S_{21})$, and

$$v_4 = -\,a_{44} = a_{43} = 3\mu. \tag{11.23}$$

So we can write the matrix A of rates:

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 \\ \mu & -\lambda-\mu & \lambda & 0 & 0 \\ 0 & 2\mu & -\lambda-2\mu & \lambda & 0 \\ 0 & 0 & 2\mu & -\lambda-2\mu & \lambda \\ 0 & 0 & 0 & 2\mu & -2\mu \end{pmatrix}. \tag{11.24}$$

This can be shown in the following scheme:

$$0 \underset{\mu}{\overset{\lambda}{\rightleftarrows}} 1 \underset{2\mu}{\overset{\lambda}{\rightleftarrows}} 2 \underset{2\mu}{\overset{\lambda}{\rightleftarrows}} 3 \underset{2\mu}{\overset{\lambda}{\rightleftarrows}} 4\,. \tag{11.25}$$

Let me start (only start) a new thing.

In the discrete-time case the whole family of n-step transition matrices P^n can be produced if we know only the one-step transition matrix P. Can all information about the family of transition matrices P^t in the continuous-time case be obtained from one matrix only? We suspect that this should be the rate matrix A.

It turns out that we can write some differential equations for the transition probabilities $p(t,\, x,\, y)$ (or, which is the same: for the transition matrices P^t). These equations were derived by A.N.Kolmogorov, and they are called *Kolmogorov equations.*

Let us start with the equalities

$$P^{t+h} = P^h \cdot P^t = P^t \cdot P^h; \tag{11.26}$$

or, for the entries of these matrices:

$$p(t+h,\, x,\, z) = \sum_y p(h,\, x,\, y) \cdot p(t,\, y,\, z) = \sum_y p(t,\, x,\, y) \cdot p(h,\, y,\, z). \tag{11.27}$$

Let us write the expression for the derivative $\dfrac{d}{dt}\, p(t,\, x,\, z)$:

$$\begin{aligned} \frac{d}{dt}\, p(t,\, x,\, z) = \lim_{h\to 0} \frac{p(t+h,\, x,\, z) - p(t,\, x,\, z)}{h} &= \lim_{h\to 0} \frac{\sum_y p(h,\, x,\, y)\cdot p(t,\, y,\, z) - p(t,\, x,\, z)}{h} \\ &= \lim_{h\to 0} \frac{\sum_y p(t,\, x,\, y)\cdot p(h,\, y,\, z) - p(t,\, x,\, z)}{h}. \end{aligned} \tag{11.28}$$

In the first representation, let us replace $p(t,\, x,\, z)$ with $\sum_y \delta_{xy} \cdot p(t,\, y,\, z)$, in the second, with $\sum_y p(h,\, x,\, y) \cdot \delta_{yz}$. The limit (11.28) is equal to each of the two limits:

$$\lim_{h \to 0} \frac{\sum_y [p(h,\, x,\, y) - \delta_{xy}] \cdot p(t,\, y,\, z)}{h}, \quad \lim_{h \to 0} \frac{\sum_y p(t,\, x,\, y) \cdot [p(h,\, y,\, z) - \delta_{yz}]}{h}. \tag{11.29}$$

We know that $\lim_{h \to 0+} \dfrac{p(h,\, x,\, y) - \delta_{xy}}{h} = a_{xy}$ (same in the second limit, with a_{yz} instead of a_{xy}); so it would seem that

$$\frac{d}{dt}\, p(t,\, x,\, z) = \sum_y a_{xy} \cdot p(t,\, y,\, z) = \sum_y p(t,\, x,\, y) \cdot a_{yz}. \tag{11.30}$$

In the matrix form it is

$$\frac{d}{dt}\, P^t = A\, P^t = P^t\, A. \tag{11.31}$$

However, I have to admit that it only *seems* so to us: we see only the direction in which we have to move. But there are awful gaps in our reasoning (which was no real reasoning, but only writing formulas). First of all, a_{xy} are only *right-hand* derivatives of $p(t,\, x,\, y)$ at $t = 0$; so what we have now is, at most, the statement about the right-hand derivative: $\dfrac{d^+}{dt} P^t = A\, P^t = P^t\, A$. And secondly, we can replace the limit of the sums by the sum of the limits in the case of *finite* sums (i. e., for the space X, over which y changes, being finite); for infinite sums we need some extra conditions.

We see that there is something to think about.

But with this to the next lecture.