Lecture 34. More about Markov processes.

The first thing I want to tell you – probably should have told you before: clearly, the transition probability from time t to the same time t is the unit measure concentrated at one point:

$$P(t,\, x,\, t,\, C) = \delta_x(C) = \begin{cases} 1 & \text{if } C \ni x, \\ 0 & \text{if } C \not\ni x. \end{cases} \tag{34.1}$$

`I couldn't include in this lecture note details of the proofs; take it as it is.`

According to Theorem 32.5, we can write the finite-dimensional distributions associated with a Markov process: for $0 \le t_1 \le ... \le t_n$ and $C \in \mathcal{X}^n$

$$\begin{aligned} P\{(\xi_{t_1}, ..., \xi_{t_n}) \in C\} = \int_X \nu(dx) \int_X P(0,\, x,\, t_1,\, dx_1) \int_X P(t_1,\, x_1,\, t_2,\, dx_2) \int_X ... \\ ... \int_X P(t_{n-1},\, x_{n-1},\, t_n,\, dx_n)\; I_C(x_1, ..., x_n). \end{aligned} \tag{34.2}$$

This is obtained from formula (33.1) with $t = 0$ plus the use of the generalized total probability formula with respect to the σ-algebra $\sigma(\xi_0)$: the expectation of a function of the random variable ξ_0 is equal to the integral of this function with respect to the distribution ν of this random variable.

Formula (34.2) is the analog of formula (25.1) for Markov chains. That equality was *necessary and sufficient* for a sequence of random variables to be a Markov chain with given transition matrix and initial distribution. The same turns out to be true in the case of continuous time, but the proof is a little longer.

Theorem 34.1. *Let the stochastic process ξ_t, $t \ge 0$, satisfies the equalities* (34.2). *Then it is a Markov process with transition function $P(\ ,\ ,\ ,\ )$ and the initial distribution ν.*

Proof. For all $0 \le t_1 \le ... \le t_n \le t \le u$ and every $D \in \mathcal{X}^n$ and $C \in \mathcal{X}$ we have, by formula (34.2) with $n+2$ instead of n, t and u instead of t_{n+1} and t_{n+2}, and $D \times X \times C$ instead of C:

$$\begin{aligned} P\{(\xi_{t_1}, ..., \xi_{t_n}) \in D,\, \xi_u \in C\} &= \int_X \nu(dx_0) \int_X P(0,\, x_0,\, t_1,\, dx_1) \int_X ... \\ &... \int_X P(t_n,\, x_n,\, t,\, dx) \int_X P(t,\, x,\, u,\, dy)\; I_D(x_1, ..., x_n) \cdot I_C(y) \\ &= E[I_D(\xi_{t_1}, ..., \xi_{t_n}) \cdot P(t,\, \xi_t,\, u,\, C)]. \end{aligned} \tag{34.3}$$

So for every event A of the form $A = \{(\xi_{t_1}, ..., \xi_{t_n}) \in D\}$, $D \in \mathcal{X}^n$, we have

$$P(A \cap \{\xi_u \in C\}) = E[I_A \cdot P(t,\, \xi_t,\, u,\, C)]. \tag{34.4}$$

The left-hand side and the right-hand side here are (finite) measures as functions of A; these measures coincide on the class of sets of the form $\{(\xi_{t_1}, ..., \xi_{t_n}) \in D\}$, which is a semi-algebra, so by the uniqueness theorems they coincide also on the σ-algebra generated by this semi-algebra, i. e. on $\sigma(\xi_s,\, 0 \le s \le t)$. This proves the theorem.

We had also Theorem 25.2 stating existence of a Markov chain with prescribed transition matrix and initial distribution; it's a little early for that in the case of continuous time.

Theorem 34.2. *Let ξ_t, $t \ge t$, be a Markov process with transition function $P(t, x, u, C)$. For every triple $s \le t \le u$, for every set $C \in \mathcal{X}$, for almost all $x \in X$ with respect to the one-dimensional distribution $\mu_{\xi_s}(\bullet) = P\{\xi_s \in \bullet\}$ we have:*

$$P(s, x, u, C) = \int_X P(s, x, t, dy)\; P(t, y, u, C). \tag{34.5}$$

Proof. First of all, the integral in the right-hand side makes sense; even more, it is an $\mathcal{X}$-measurable function of x. The conditional probability $P\{\xi_u \in C \| \xi_v,\, 0 \le v \le s\}$ can be written as

$$P\{\xi_u \in C \| \xi_v,\, 0 \le v \le s\} = P(s, \xi_s, u, C) \tag{34.6}$$

(almost surely, of course). But the event $\{\xi_u \in C\}$ can also be written as $\{\xi_t \in X,\, \xi_u \in C\}$, and its conditional probability is equal almost surely to

$$P\{\xi_t \in X,\, \xi_u \in C \| \xi_v,\, 0 \le v \le s\} = \int_X P(s, \xi_s, t, dy)\; P(t, y, u, C). \tag{34.7}$$

The random variables in the right-hand sides of equalities (34.6), (34.7) are equal to each other almost surely, from which we get the statement of the theorem.

Equality (34.5) is called *the Chapman – Kolmogorov equation.* It is the continuous-time analogue of equality (25.24) for Markov chains.

In the discrete case we did not need to say anything about that equality being satisfied almost everywhere: the proof of the equality being satisfied for all values of the variables in it was there: in particular, we could use the equality $P^{nm} = P_{n+1} \cdot P_{n+2} \cdot ... \cdot P_m$ and obtain $P^{nr} = P^{nm} \cdot P^{mr}$ from it. Not so in the case of continuous time: there is nothing that corresponds to the *one-step* transition probabilities (or transition matrix).

The question arises: can we always replace the transition function that does not satisfy the Chapman – Kolmogorov equation everywhere by one that satisfies this equation *everywhere*, and is equal to our original transition function almost everywhere (and so is another version of transition function for the same Markov process)? The question turns out to be too deep and difficult for us; so we decide that we are going to consider only transition functions that satisfy the Chapman – Kolmogorov equation (34.5) *everywhere.*

If the space X is discrete (countable, finite or infinite), the Chapman – Kolmogorov equation can be written as

$$p(s, x, u, z) = \sum_{y \in X} p(s, x, t, y) \cdot p(t, y, u, z), \tag{34.8}$$

where $p(s, x, t, y) = P(s, x, t, \{y\})$; in this case we can introduce the transition matrices $P^{st} = \big(p(s, x, t, y)\big)_{x,\, y \in X}$, and rewrite the Chapman – Kolmogorov equation in the matrix form as

$$P^{su} = P^{st} \cdot P^{tu}. \tag{34.9}$$

Problem **54** Invent a family of stochastic matrices P^{st}, $0 \le s \le t$, such that $P^{tt} = I$ (the identity matrix), the Chapman – Kolmogorov equation (34.9) is satisfied for all $0 \le s \le t \le u$, $P^{st} \ne \text{const}$, and P^{st} depends on (s, t) continuously.

If the space X is the Euclidean space $\mathbb{R}^r$ or a Borel set in a Euclidean space, we can consider transition probabilities $P(s, x, t, C)$ having *densities* (we'll call them *transition densities*):

$$P(s, x, t, C) = \int_C p(s, x, t, y)\, dy, \tag{34.10}$$

$p(s, x, t, y)$ being the density (with respect to the Lebesgue measure) as a function of its last argument. Of course, in the case of $X = \mathbb{R}^r$ there are plenty of measures that have no density with respect to the Lebesgue measure; but there are some that have densities; let us consider such.

Only there is no hope that transition densities exist for all $0 \le s \le t$: for $s = t$ the distribution $P(t, x, t, \bullet) = \delta_x(\bullet)$ is a discrete one, and cannot have a density with respect to the Lebesgue measure. So let us consider the class of transition functions (satisfying the Chapman – Kolmogorov equation everywhere) that are described by densities for all $0 \le s < t$.

What conditions should we impose on a function $p(s, x, t, y)$ for the function $P(s, x, t, C)$ defined through it by formula (34.10) to satisfy the conditions that we require from a transition function (in particular, for the Chapman – Kolmogorov equation to be satisfied)?

Clearly we have to require its Borel measurability in its y argument – otherwise the integral (34.10) makes no sense. It must be nonnegative, and $\displaystyle\int_X p(s, x, t, y)\, dy$ must be equal to 1 (otherwise $P(s, x, t, \bullet)$ is not a probability measure). But how to make sure that $P(s, x, t, C)$ is Borel measurable in x? Measurability of $p(s, x, t, y)$ in x does not imply the measurability of the integral. What is enough is requiring that $p(s, x, t, y)$ should be measurable in the *couple* (x, y): by Fubini's theorem then the integral in y is measurable as a function of x. We are going to have some examples of such functions $p(s, x, t, y)$ that might be transition densities of Markov processes.