Lecture 26. Some more examples. Martingales.

Example 26.1. $X = [-\infty, 0]$, Af is defined as f' on the domain $D_A = \bigl\{f \in \mathbf{C}^1[-\infty, 0]\colon cf'(0) = \int_{[-\infty, 0)} [f(x) - f(0)]\,\mu(dx)\bigr\}$, where $c > 0$, and μ is a finite measure on $\bigl([-\infty, 0), \mathcal{B}_{[-\infty, 0)}\bigr)$.

The linear functional $l(f) = -cf'(0) + \int_{[-\infty, 0)} [f(x) - f(0)]\,\mu(dx)$ defined on $\mathbf{C}^1[-\infty, 0]$ is unbounded, because $f'(0)$ is an unbounded functional, and $\int_{[-\infty, 0)} [f(x) - f(0)]\,\mu(dx)$ a bounded one. So D_A is dense.

The maximum principle: if f reaches its maximum at 0, we have $\int_{[-\infty, 0)} [f(x) - f(0)]\,\mu(dx) \le 0$, $f'(0) \le 0$ (if the maximum of a function $f \in \mathbf{C}^1[-\infty, 0]$ is reached at a point $x \in [-\infty, 0)$, then $f'(x) = 0$).

The existence of a solution $F \in D_A$ of the equation $\lambda F - F' = f$ is checked the same way as in Example 25.6.

If μ is a probability measure on $[-\infty, 0)$ (i. e. $\mu[-\infty, 0) = 1$), then the corresponding Markov process goes like this: if it starts at a point different from $-\infty$, it moves to the right at a unit speed until it reaches 0; it spends an exponential time at 0, after which it jumps to a random point in $[-\infty, 0)$ with distribution μ, and starts again moving to the right; on reaching 0 it remains there for an exponential time, and jumps again, independently from what happened before; etc.

Example 26.2. $X = [-\infty, 0]$, Af is defined as f' on the domain $D_A = \bigl\{f \in \mathbf{C}^1[-\infty, 0]\colon \int_{[-\infty, 0)} [f(x) - f(0)]\,\mu(dx)\bigr\} = 0$, where μ is a finite measure.

The linear functional $l(f) = \int_{[-\infty, 0)} [f(x) - f(0)]\,\mu(dx)$ is bounded, so D_A is not dense, the continuous semigroup on $\mathbf{C}$ does not exist.

The limit of exponential distributions with parameter c^{-1} as $c \to 0+$ is the distribution concentrated at 0, so the corresponding Markov process must jump *instantly* to a point with distribution μ at reaching the point 0. But such a process is *not* uniformly continuous in probability (and even not uniformly, for the initial point $\xi_0 = 0$), so it is out our scope: we are considering only uniformly continuous in probability Feller Markov processes.

Think about the following interesting problem: what are the most general boundary conditions to restrict $\mathbf{C}^1[-\infty, 0]$ so that the operator f' with these boundary conditions is the infinitesimal operator of a uniformly continuous in probability Feller process?

Can there be anything besides the boundary condition of Example 25.6?

All processes that we are considering are uniformly continuous in probability, so their trajectories can be chosen to be right-continuous with left limits. What about the *two-sided* continuity?

We had the condition

$$\sup_{x\in X} P\big(h,\, x,\, \{y\colon \operatorname{dist}(y,\, x) > \varepsilon\}\big) = o(h) \qquad (h \to 0^+) \tag{26.1}$$

for this.

We call a linear operator A *local* if the value of Af at a point x depends only on the values of the function $f \in D_A$ in an arbitrarily small neighborhood of x: for every $\varepsilon > 0$

$$f_1,\, f_2 \in D_A, \quad f_1(y) = f_2(y) \text{ if } \operatorname{dist}(y,\, x) < \varepsilon \;\Rightarrow\; Af_1(x) = Af_2(x), \tag{26.2}$$

or, which is the same,

$$f \in D_A, \quad f(y) = 0 \text{ if } \operatorname{dist}(y,\, x) < \varepsilon \;\Rightarrow\; Af(x) = 0. \tag{26.3}$$

Theorem 26.1. *If the condition* (26.1) *is satisfied, the infinitesimal operator A of the Markov process is local.*

Proof. We have, for a function $f \in D_A$ that is equal to 0 in the ε-neighborhood of a point x:

$$\begin{aligned} Af(x) = \lim_{h\to 0^+} h^{-1}[(P^t f)(x) - f(x)] &= \lim_{h\to 0^+} h^{-1} \int_X f(y)\; P(h,\, x,\, dy) \\ &= \lim_{h\to 0^+} h^{-1} \int\limits_{\{y\colon \operatorname{dist}(y,\, x)>\varepsilon\}} f(y)\; P(h,\, x,\, dy), \end{aligned} \tag{26.4}$$

$$|Af(x)| \le \lim_{h\to 0^+} h^{-1}\|f\| \cdot P(h,\, x,\, \{y\colon \operatorname{dist}(y,\, x) > \varepsilon\}) = 0. \tag{26.5}$$

The infinitesimal operators in Examples 25.1 – 25.5, 26.1 were local, but the process in Example 26.1 had jumps; so the locality condition is not sufficient for (26.1), we have to add something.

Theorem 26.2. *Let X be a compact space. Let A be the infinitesimal operator of a continuous semigroup on $\mathbf{C}(X)$ corresponding to a Markov process. Let A be a local operator.*

Suppose that for every $x \in X$ and for every positive ε there exists a nonnegative function $f \in D_A$ that is equal to 0 *in the $(\varepsilon/3)$-neighborhood of x and not less than* 1 *outside its $(2\varepsilon/3)$-neighborhood.*

Then the condition (26.1) *is satisfied.*

Proof. Let x_1, ..., x_N be an $\varepsilon/3$-net in X; let f_i be the function whose existence is postulated in the theorem, for $x = x_i$. For every $\gamma > 0$ there exists a positive h_i such that for $h \le h_i$

$$\|h^{-1}(P^h f_i - f_i) - Af_i\| < \gamma. \tag{26.6}$$

Take $h_0 = \min(h_0, ..., h_N) > 0$.

Let $h \le h_0$. For an arbitrary $x \in X$, let x_i be a point in our net with $\text{dist}(x, x_i) < \gamma/3$. The function f_i is equal to 0 in an open set $\{y\colon \text{dist}(y, x_i) < \varepsilon/3\}$ containing the point x; since the operator A is local, we have $Af_i(x) = 0$ (as well as $f_i(x) = 0$). From (26.6) we get:

$$h^{-1} P^h f_i(x) = h^{-1} \int_X f_i(y)\ P(h, x, dy) < \gamma, \tag{26.7}$$

$$\begin{aligned} h^{-1} P(h, x, \{y\colon \text{dist}(y, x) > \varepsilon\}) &\le h^{-1} P(h, x, \{y\colon \text{dist}(y, x_i) > 2\varepsilon/3\}) \\ &\le h^{-1} \int_X f_i(y)\ P(h, x, dy) < \gamma, \end{aligned} \tag{26.8}$$

from which (26.1) follows.

$\boxed{\mathbf{47}}$ Let $X = [0, \infty]$, $D_A = \{f \in \mathbf{C}^2[0, \infty]\colon c \cdot \frac{1}{2} f''(0) = b \cdot f'(0)\}$, where c and b are nonnegative constants, $b + c > 0$; $Af(x) = \frac{1}{2} f''(x)$.

The operator A is the infinitesimal operator of a continuous semigroup on $\mathbf{C}[0, \infty]$ corresponding to a Markov process (you are not asked to check this).

Prove that this process can be taken with continuous trajectories.

Enough about semigroups for the time being. Let us go to the important class of random functions that I promised to talk about: martingales.

Let $\mathcal{F}_t$, $t \in T \subseteq \mathbb{R}^1$, be a non-decreasing family of σ-algebras in a probability space $(\Omega, \mathcal{F}, P)$. Let η_t, $t \in T$, be a number-valued random function.

We say that the random function η_t is a *martingale* if

- it is adapted to the family $\mathcal{F}_t$ (i. e., for every $t \in T$ the random variable η_t is $\mathcal{F}_t$-measurable);
- for every two times $t_1 < t_2$ in T almost surely

$$E(\eta_{t_2} \| \mathcal{F}_{t_1}) = \eta_{t_1}. \tag{26.9}$$

We say that a random function ζ_t is a *submartingale* if it is adapted to $\mathcal{F}_t$, and for $t_1 < t_2$ in T almost surely

$$\zeta_{t_1} \le E(\zeta_{t_2} \| \mathcal{F}_{t_1}). \tag{26.10}$$

I don't call the random functions η_t, ζ_t *stochastic processes*: why? Because the standard way to use martingales is that we have a stochastic process ξ_t; and we construct starting from it some auxiliary random fucntion η_t, which turns out to be a martingale (or a submartingale), from which we obtain some consequences about the process ξ_t.

Of course, the requirement of the random function being adapted is much simpler, belonging to the set-theoretic introduction to the theory of stochastic processes. The requirement (26.9) means two things: one is the $\mathcal{F}_{t_1}$-measurability of η_{t_1} once again (but this requirement is not repeated in (26.10)); and the other is that for every event $A \in \mathcal{F}_{t_1}$

$$E(I_A \cdot \eta_{t_1}) = E(I_A \cdot \eta_{t_2}); \tag{26.11}$$

or, equivalently,

$$E\big(I_A \cdot (\eta_{t_2} - \eta_{t_1})\big) = 0. \tag{26.12}$$

The requirement (26.10) is equivalent to

$$E(I_A \cdot \zeta_{t_1}) \le E(I_A \cdot \zeta_{t_2}), \qquad E\big(I_A \cdot (\zeta_{t_2} - \zeta_{t_1})\big) \ge 0. \tag{26.13}$$

Before we consider examples, let me formulate a result that facilitates working with martingales in the discrete-time case.

Theorem 26.3. *Let* η_t, ζ_t *be random functions defined for some* (finite or infinite) *interval of integers* t: $t \in T = \{(...,)\ m,\ m+1,\ m+2,\ ...,\ n-1,\ n\ (, ...)\}$. *Then* η_t, $t \in T$, *is a martingale* (ζ_t *is a submartingale*) *with respect to* $\mathcal{F}_t$ *if and only if for every* $t \in T$ *such that* $t+1 \in T$

$$\eta_t = E(\eta_{t+1} \| \mathcal{F}_t) \qquad (\text{or } \zeta_t \le E(\zeta_{t+1} \| \mathcal{F}_t)). \tag{26.14}$$

Proof. We have to prove that not only (26.14) is satisfied, but also

$$\eta_t = E(\eta_{t+2} \| \mathcal{F}_t) \qquad (\text{or } \zeta_t \le E(\zeta_{t+2} \| \mathcal{F}_t)), \tag{26.15}$$

$$\eta_t = E(\eta_{t+3} \| \mathcal{F}_t) \qquad (\text{or } \zeta_t \le E(\zeta_{t+3} \| \mathcal{F}_t)), \tag{26.16}$$

etc. By the properties of the conditional expectation, we have:

$$E(\eta_{t+2} \| \mathcal{F}_t) = E\big(E(\eta_{t+2} \| \mathcal{F}_{t+1}) \big\| \mathcal{F}_t\big). \tag{26.17}$$

By (26.14), we have:

$$E(\eta_{t+2} \| \mathcal{F}_{t+1}) = \eta_{t+1} \qquad (\text{or } E(\zeta_{t+2} \| \mathcal{F}_{t+1}) \ge \zeta_{t+1}), \tag{26.18}$$

so the expectation (26.17) is equal to (or $\ge$)

$$E(\eta_{t+1} \| \mathcal{F}_t) = \eta_t \qquad (\text{or } \ge \zeta_t) \tag{26.19}$$

by (26.14). To prove (26.16), we insert *two* conditional expectations: with respect to $\mathcal{F}_{t+2}$, and with respect to $\mathcal{F}_{t+1}$; etc.

Formulas (26.14) can be rewritten as

$$E(\eta_{t+1} - \eta_t \| \mathcal{F}_t) = 0 \qquad (E(\zeta_{t+1} - \zeta_t \| \mathcal{F}_t) \ge 0), \tag{26.20}$$

$$E\big(I_A \cdot (\eta_{t+1} - \eta_t)\big) = 0 \qquad \big(E\big(I_A \cdot (\zeta_{t+1} - \zeta_t)\big) \ge 0\big) \tag{26.21}$$

for every event $A \in \mathcal{F}_t$.

Example 26.3. Let ξ_1, ξ_2, ..., ξ_n, ... be a sequence of independent identically distributed random variables. Is this sequence a martingale with respect to the family of σ-algebras $F_{\le n} = \sigma(\xi_1, ..., \xi_n)$?

By the properties of the conditional expectation, we have:

$$E(\xi_{n+1} \| \mathcal{F}_{\le n}) = E(\xi_{n+1} \| \xi_1, ..., \xi_n) = E\xi_{n+1} = \text{const}, \tag{26.22}$$

which is *not* almost surely equal to the random variable ξ_n (unless all our random variables take with probability one one and the same constant value). So the sequence of ξ_n is not a martingale. It is not a submartingale either, because ξ_n may be greater or smaller than the expectation (which is also the conditional expectation) $E\xi_{n+1}$.

Example 26.4 (the first example of a random function that *is* a martingale). Let ξ_1, ξ_2, ..., ξ_n, ... be independent random variables taking values ± 1 with probabilities 1/2. For every $n \ge 1$, let $f_n(x_1, ..., x_{n-1})$ be a function of the arguments x_i taking values ± 1 (for $n = 1$, f_1 is a function of zero arguments, which is taken to mean a constant). Let us take a constant y_0 and form the sums

$$\begin{gathered} \eta_0 = y_0, \quad \eta_1 = y_0 + f_1 \cdot \xi_1, \quad \eta_2 = y_0 + f_1 \cdot \xi_1 + f_2(\xi_1) \cdot \xi_2, \\ \eta_3 = y_0 + f_1 \cdot \xi_1 + f_2(\xi_1) \cdot \xi_2 + f_3(\xi_1, \xi_2) \cdot \xi_3, \quad ..., \\ \eta_n = y_0 + \sum_{i=1}^{n} f_i(\xi_1, ..., \xi_{i-1}) \cdot \xi_i, \quad .... \end{gathered} \tag{26.23}$$

Let us check that the sequence η_0, η_1, η_2, ..., η_n, ... is a martingale with respect to the σ-algebras $\mathcal{F}_{\le n} = \sigma(\xi_1, ..., \xi_n)$ (for $n = 0$ we take the σ-algebra $\mathcal{F}_{\le 0}$ generated by an empty set of random variables to be the smallest possible σ-algebra $\mathcal{O} = \{\Omega, \emptyset\}$).

The sequence $\{\eta_n\}$ can be interpreted as follows: We toss a coin time after time, and bet a quantity f_i on the outcome of the i-th tossing. The size of our bet depends, in general, on the outcomes of the previous tosses (we would be too lucky if we knew ξ_n before assigning the size f_n of our bet; but we take it to be impossible). The quantity y_0 represents our initial capital, and η_n our capital after n bets.

We have: $\eta_0 \equiv y_0$ is a constant, and it is measurable with respect to the σ-algebra $\mathcal{F}_{\le 0} = \mathcal{O}$; and clearly every η_n, $n \ge 1$, is measurable with respect to $\mathcal{F}_{\le n}$.

Now we have to check that

$$E(\eta_{n+1} - \eta_n \| \mathcal{F}_{\le n}) = 0. \tag{26.24}$$

By the properties of conditional expectations we have:

$$E(\eta_{n+1} - \eta_n) \| \mathcal{F}_{\le n}) = f_{n+1}(\xi_1, ..., \xi_n) \cdot E(\xi_{n+1} \| \mathcal{F}_{\le n}) = f_{n+1}(\xi_1, ..., \xi_n) \cdot E\xi_{n+1} = 0. \tag{26.25}$$

Generalization: if ξ_1, ξ_2, ..., ξ_n, ... are independent random variables with *arbitrary* distributions with $E\xi_i = 0$, and $f_n(x_1, ..., x_{n-1})$ arbitrary functions such that $E|f_n(\xi_1, ..., \xi_{n-1})| < \infty$, the sequence (26.23) is a martingale: our capitals after participating n times in a *fair game* ($E\xi_i = 0$).

If $E\xi_i \ge 0$ (the game is favorable to us) and $f_n(x_1, ..., x_{n-1}) \ge 0$ (we don't bet against our luck), the sequence (26.23) is a *submartingale*.

When J.L.Doob, the American mathematician who wrote the book *Stochastic Processes*, where much attention was paid to martingales, visited Moscow, we asked him why this class of random functions is called *martingales*: what does this word mean? He explained that a martingale was a device checking upward movement of a horse's head: a strap fastened to the bit. This seemed to have nothing to do with the mathematical concept: or did it contain some depths that we couldn't grasp? But in fact, there is another meaning of the word "martingale": "any of several systems of betting in which a player increases his stake usually by doubling each time he loses a bet". The above example shows that, in a fair game, no system of betting can increase (or decrease) your expected winnings.