Lecture 27. Markov processes and martingales. Martingales and stopping times.

Example 27.1 (a class of examples, in fact). Let $\xi_0, \xi_1, \xi_2, ..., \xi_n, ...$ be a Markov chain on a countable space X with one-step transition matrix P. Let f be a column vector with components $f(x)$, $\|f\| = \sup_x |f(x)| < \infty$ (I don't denote column vectors with bold-faces letters anymore: they are just functions). Then

$$\eta_n = f(\xi_n) - \sum_{k=0}^{n-1} (Pf - f)(\xi_k) \tag{27.1}$$

is a martingale with respect to $\mathcal{F}_{\le n} = \sigma(\xi_0, ..., \xi_n)$.

Adaptedness: trivial. Now:

$$E(\eta_{n+1} - \eta_n \| \mathcal{F}_{\le n}) = E[f(\xi_{n+1}) \| \mathcal{F}_{\le n}] - (Pf)(\xi_n) = 0, \tag{27.2}$$

which follows immediately from the Markov property.

Example 27.2. Let ξ_t, $t \ge 0$, be a Markov process with respect to $(\mathcal{F}_t)$ such that for every $t \ge 0$ the function $\xi_s(\omega)$, $0 \le s \le t$, $\omega \in \Omega$, is measurable in (s, ω) with respect to the σ-algebra $\mathcal{B}_{[0,\,t]} \times \mathcal{F}_t$. Suppose $f \in D_A$.

Then the random function

$$\eta_t = f(\xi_t) - \int_0^t Af(\xi_s)\, ds \tag{27.3}$$

is a martingale with respect to $(\mathcal{F}_t)$.

First of all, Af is measurable and bounded; $Af\big(\xi_s(\omega)\big)$, $0 \le s \le t$, $\omega \in \Omega$, is $(\mathcal{B}_{[0,\,t]} \times \mathcal{F}_t)$-measurable; the integral $\int_0^t Af\big(\xi_s(\omega)\big)\, ds$ exists, and by Fubini's Theorem this random variable is $\mathcal{F}_t$-measurable: so η_t is adapted to $(\mathcal{F}_t)$. Now we have to prove that for $t_2 \ge t_1$

$$E(\eta_{t_2} - \eta_{t_1} \| \mathcal{F}_{t_1}) = E\Big(f(\xi_{t_2}) - f(\xi_{t_1}) - \int_{t_1}^{t_2} Af(\xi_s)\, ds \Big\| \mathcal{F}_{t_1}\Big) = 0. \tag{27.4}$$

Of course, $E\big(f(\xi_{t_1}) \big\| \mathcal{F}_{t_1}\big) = f(\xi_{t_1})$; by the Markov property we have almost surely

$$E\big(f(\xi_{t_2}) \big\| \mathcal{F}_{t_1}\big) = P^{t_2 - t_1} f(\xi_{t_1}). \tag{27.5}$$

It seems plausible that

$$E\Big(\int_{t_1}^{t_2} Af(\xi_s)\, ds \Big\| \mathcal{F}_{t_1}\Big) = \int_{t_1}^{t_2} P^{s-t} Af(\xi_{t_1})\, ds = \int_0^{t_2 - t_1} P^u Af(\xi_{t_1})\, du. \tag{27.6}$$

Is this so? The integral in the left-hand side is the θ_{t_1}-shift: $\displaystyle\int_{t_1}^{t_2} Af\big(\xi_s(\omega)\big)\,ds = \int_0^{t_2-t_1} Af\big(\xi_u(\theta_{t_1}\omega)\big)\,du$; so by the Markov property (the formula (15.6) rewritten with random variables and expectations) we have:

$$E\Big(\int_{t_1}^{t_2} Af(\xi_s)\,ds\Big\|\mathcal{F}_{t_1}\Big) = E_{\xi_{t_1}}\int_0^{t_2-t_1} Af(\xi_u)\,du. \tag{27.7}$$

Using Fubini's Theorem, we get:

$$E_x\int_0^{t_2-t_1} Af(\xi_u)\,du = \int_0^{t_2-t_1} E_x Af(\xi_u)\,du = \int_0^{t_2-t_1} P^u Af(x)\,du. \tag{27.8}$$

Putting here ξ_{t_1} in the place of x, we get (27.6).

So we have almost surely:

$$E\big(\eta_{t_2}-\eta_{t_1}\|\mathcal{F}_{t_1}\big) = P^{t_2-t_1}f(\xi_{t_1}) - f(\xi_{t_1}) - \int_0^{t_2-t_1} P^u Af(\xi_{t_1})\,du. \tag{27.9}$$

But we have, by (22.14): $\displaystyle P^{t_2-t_1}f - f = \int_0^{t_2-t_1}\frac{d}{du}P^u f\,du = \int_0^{t_2-t_1} P^u Af\,du$, so the expectation (27.4) is equal to 0.

A concretization of this example: if the trajectories $\xi_t(\omega)$ are right-continuous functions with values in a metric space X, then the $(\mathcal{B}_{[0,\,t]}\times\mathcal{F}_t)$-measurability condition is satisfied.

Indeed, let us define, for every $h>0$, the random function

$$\xi_s^h(\omega) = \begin{cases} \xi_{kh}(\omega), & (k-1)h < s \le kh < t, \\ \xi_t(\omega), & (n-1)h < s \le t \le nh. \end{cases} \tag{27.10}$$

This random function is clearly $(\mathcal{B}_{[0,\,t]}\times\mathcal{F}_t)$-measurable on $[0,\,t]\times\Omega$; and $\xi_s(\omega) = \lim_{h\to 0+}\xi_s^h(\omega)$.

Let us look at what the situation is with diffusion processes.

I told you that the infinitesimal operator A of a diffusion process in $\mathbb{R}^d$ associated with the differential operator L given by

$$Lf(\boldsymbol{x}) = \tfrac{1}{2}\sum_{i,\,j=1}^d a_{ij}(\boldsymbol{x})\,\frac{\partial^2 f}{\partial x_i\partial x_j}(\boldsymbol{x}) + \sum_{i=1}^d b_i(\boldsymbol{x})\,\frac{\partial f}{\partial x_i}(\boldsymbol{x}) \tag{27.11}$$

is defined on $D_A \supseteq \mathbf{C}^2(\mathbb{R}^d)$, and for $f\in\mathbf{C}^2(\mathbb{R}^d)$ we have $Af = Lf$; but I did not prove this. Let me do it, with your help (consisting of you solving some problems). The proof will be based on the *forward* Kolmogorov equation for the transition density

$$\frac{\partial p}{\partial t}(t,\,\boldsymbol{x},\,\boldsymbol{y}) = L^*p(t,\,\boldsymbol{x},\,\bullet)(\boldsymbol{y}), \tag{27.12}$$

which holds if the coefficients a_{ij}, b_i are bounded and Hölder continuous together with the first and the second derivatives of a_{ij} and the first derivatives of b_i (a fact from the theory of parabolic equations); the estimate that I have already mentioned:

$$p(t,\, \boldsymbol{x},\, \boldsymbol{y}) \le \frac{C}{t^{d/2}}\, e^{-a|\boldsymbol{y}-\boldsymbol{x}|^2/2t} \tag{27.13}$$

(see (21.9)), and the one that I did not mentioned, but which is there in Friedman's book on the same page:

$$\Big|\frac{\partial p}{\partial y_i}(t,\, \boldsymbol{x},\, \boldsymbol{y})\Big| \le \frac{C}{t^{d/2+1}}\, e^{-a|\boldsymbol{y}-\boldsymbol{x}|^2/2t}. \tag{27.14}$$

The formula $P^t f - f = \int_0^t P^u A f \, du$ was a general formula within the Hille–Yosida theory of continuous semigroups of linear operators in a Banach space – in our case the space $\mathbf{B}$ of bounded measurable functions, – and it held for functions f that were bounded together with Af. But in concrete cases we are able to prove, stepping a little outside the framework of this theory, that the expectations $P^t f(x) = E_x f(\xi_t)$ and $P^u A f(x)$ exist for some classes of unbounded functions f.

48 Prove that for a continuous function $f(\boldsymbol{x})$ that doesn't grow faster than exponentially: $|f(\boldsymbol{x})| \le C_1 e^{C_2|\boldsymbol{x}|}$, $P^t f(\boldsymbol{x}) = \int_{\mathbb{R}^d} p(t,\, \boldsymbol{x},\, \boldsymbol{y}) \cdot f(\boldsymbol{y})\, d\boldsymbol{y}$ is defined, and we have $\lim_{t\to 0+} P^t f(\boldsymbol{x}) = f(\boldsymbol{x})$.

49 Prove that for a twice continuously differentiable function $f(\boldsymbol{x})$ growing, together with its first and second derivatives, not faster than exponentially: $|f(\boldsymbol{x})|$, $\big|\frac{\partial f}{\partial x_i}\big|$, $\big|\frac{\partial^2 f}{\partial x_i \partial x_j}\big| \le C_1 e^{C_2|\boldsymbol{x}|}$, we have: $\frac{\partial}{\partial t} P^t f(\boldsymbol{x}) = P^t L f(\boldsymbol{x})$ for $t > 0$ (prove this by integrating by parts), and $P^t f(\boldsymbol{x}) - f(\boldsymbol{x}) = \int_0^t P^u f(\boldsymbol{x})\, du$.

Just as in Example 27.2, we prove that for functions f growing, together with its first and second derivatives, not faster than exponentially, $\eta_t = f(\boldsymbol{\xi}_t) - \int_0^t Lf(\boldsymbol{\xi}_s)\, ds$ is a martingale (with respect to the σ-algebras $\mathcal{F}_{\le t}$, or to $(\mathcal{F}_{\le t+})$).

Concrete examples:

Example 27.3. The one-dimensional Wiener process ξ_t is a martingale with respect to the family of σ-algebras generated by it: we take $f(x) = x$ (growing definitely not faster than exponentially), and $Lf(x) = \frac{1}{2} f''(x) \equiv 0$.

Example 27.4. For the same process, $\eta_t = \xi_t^2 - t$ is a martingale: $f(x) = x^2$, $Lf(x) \equiv 1$.

Example 27.5. For the multidimensional Wiener process $Lf = \frac{1}{2}\Delta$; whereas in the one-dimensional case the only functions satisfying the equation $\Delta f = 0$ (*harmonic* functions) are linear ones, it's not so in the multidimensional case. For example, for $d = 2$ the function $f(\boldsymbol{x}) = f(x^1,\, x^2) = e^{ax^1} \cdot \cos(ax^2)$ is harmonic (you can understand why I am writing the numbers of coordinates as *super*scripts rather than subscripts), so $\eta_t = f(\boldsymbol{\xi}_t) = e^{a\xi_t^1} \cdot \cos(a\xi_t^2)$ is a martingale.

50 Check that the function $f(\boldsymbol{x}) = \dfrac{1}{|\boldsymbol{x}|}$ in $\mathbb{R}^3$ satisfies the equation $\Delta f(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \neq \mathbf{0}$. Check that for the three-dimensional Wiener process $E_{\boldsymbol{x}} f(\boldsymbol{\xi}_t) < \infty$ for $t > 0$.

Find $\lim_{t\to\infty} E_{\boldsymbol{x}} f(\boldsymbol{\xi}_t)$.

51 Is the random function $\eta_t = f(\boldsymbol{\xi}_t)$ a martingale? (This function is not covered by Example 27.5: there is no constant C_1 such that $f(\boldsymbol{x}) \le C_1 e^{C_2|\boldsymbol{x}|}$; but the expectation is finite! And f is not differentiable – but at one point only! ...)

A little theory of martingales. First a very little theorem:

Theorem 27.1. *If η_t is a martingale, its expectation is a constant.*

Proof. For $t < s$ we have:

$$E\eta_s = E\big(E(\eta_s \| \mathcal{F}_t)\big) = E\eta_t. \tag{27.15}$$

Theorem 27.2. *If ζ_t is a submartingale, its expectation $E\zeta_t$ is a non-decreasing function.*

The **proof** is the same with the obvious change.

The use of martingales relies on the relation between them and *stopping times*.

We start with the discrete case (the continuous case is more complicated):

Theorem 27.3. *Let η_t, $t \in T$, be a martingale with respect to the family of σ-algebras $\mathcal{F}_t$. Let τ be a stopping time taking finitely many values $t_1 < t_2 < ... < t_n$* (all in T, not $+\infty$).

Then for an arbitrary $t_* \in T$

$$E\eta_\tau = E\eta_{t_*}. \tag{27.16}$$

Proof. We have:

$$E\eta_\tau = \sum_{i=1}^{n} E\big(I_{\{\tau = t_i\}} \cdot \eta_\tau\big) = \sum_{i=1}^{n} E\big(I_{\{\tau = t_i\}} \cdot \eta_{t_i}\big). \tag{27.17}$$

For $i > 1$ the event $\{\tau = t_i\}$ whose indicator is used in this formula is equal to the difference of two:

$$\{\tau = t_i\} = \{\tau \le t_i\} \setminus \{\tau \le t_{i-1}\}; \tag{27.18}$$

and for $i = 1$ it is just

$$\{\tau = t_1\} = \{\tau \le t_1\}. \tag{27.19}$$

So the i-th summand in (27.17) can be written as

$$E\big((I_{\{\tau \le t_i\}} - I_{\{\tau \le t_{i-1}\}}) \cdot \eta_{t_i}\big) \tag{27.20}$$

for $i > 1$, and as

$$E(I_{\{\tau \le t_1\}} \cdot \eta_{t_1}) \tag{27.21}$$

for $i = 1$. So we get:

$$\begin{aligned} E\eta_\tau = & E(I_{\{\tau \le t_1\}} \cdot \eta_{t_1}) + E(I_{\{\tau \le t_2\}} \cdot \eta_{t_2}) - E(I_{\{\tau \le t_1\}} \cdot \eta_{t_2}) \\ & + E(I_{\{\tau \le t_3\}} \cdot \eta_{t_3}) - E\big(I_{\{\tau \le t_2\}} \cdot \eta_{t_3}\big) + \dots \\ & + E(I_{\{\tau \le t_{n-1}\}} \cdot \eta_{t_{n-1}}) - E(I_{\{\tau \le t_{n-2}\}} \cdot \eta_{t_{n-1}}) \\ & + E(I_{\{\tau \le t_n\}} \cdot \eta_{t_n}) - E(I_{\{\tau \le t_{n-1}\}} \cdot \eta_{t_n}) \\ = & \sum_{i=1}^{n-1} E\big(I_{\{\tau \le t_i\}} \cdot (\eta_{t_i} - \eta_{t_{i+1}})\big) + E(I_{\{\tau \le t_n\}} \cdot \eta_{t_n}). \end{aligned} \tag{27.22}$$

We represent the i-th summand ($i < n$) as the expectation of a conditional expectation:

$$E\big(I_{\{\tau \le t_i\}} \cdot (\eta_{t_{i+1}} - \eta_{t_i})\big) = E\big(E\big(I_{\{\tau \le t_i\}} \cdot (\eta_{t_i} - \eta_{t_{i+1}})\big\| \mathcal{F}_{t_i}\big)\big). \tag{27.23}$$

The event $\{\tau \le t_i\}$ belongs to $\mathcal{F}_{t_i}$, so its indicator can be taken from under the sign of the conditional expectation:

$$E\big(I_{\{\tau \le t_i\}} \cdot (\eta_{t_i} - \eta_{t_{i+1}})\big) = E\big(I_{\{\tau \le t_i\}} \cdot E(\eta_{t_i} - \eta_{t_{i+1}} \| \mathcal{F}_{t_i})\big); \tag{27.24}$$

and by the martingale property, the conditional expectation is equal to 0.

The last summand in (27.22) is equal to $E\eta_{t_n}$, because the event $\{\tau \le t_n\}$ is the whole Ω, and its indicator is identically 1. And of course, $E\eta_{t_n}$ is equal to $E\eta_{t_*}$ for any other $t_* \in T$.

The same way we prove

Theorem 27.4. *Let η_t, $t \in T$, be a submartingale with respect to $(\mathcal{F}_t)$. Let τ be a stopping time taking finitely many values $t_1 < t_2 < \dots < t_n = t_{\max}$.*

Then

$$E\eta_\tau \le E\eta_{t_{\max}}. \tag{27.25}$$

Proof. Formulas (27.22), (27.23), (27.24) still hold; but instead of saying "by the martingale property the conditional expectation is equal to 0" we say "by the submartingale property the conditional expectation is ≤ 0".

For random times τ that are *not* stopping times, the statements of Theorems 27.3, 27.4 are not necessarily true. E. g., if $\xi_0 \equiv 0$, and ξ_1 is a random variables with density $p(x)$ with $E\xi_1 = 0$, the sequence of two random variables ξ_0, ξ_1 is a martingale (with respect to the σ-algebras generated by them). Let us define the random time τ by

$$\tau = I_{[0,\infty)}(\xi_1). \tag{27.26}$$

This is *not* a stopping time: to know whether $\tau \le 0$ (which means $= 0$ in this case), it is not enough to watch ξ_0 (as if there were anything to watch here!), but we need to know ξ_1. We have:

$$E\xi_\tau = E(I_{\{\tau=0\}} \cdot \xi_0 + I_{\{\tau=1\}} \cdot \xi_1) = E\big(I_{[0,\infty)}(\xi_1) \cdot \xi_1\big) = \int_0^\infty x \cdot p(x)\, dx > 0, \tag{27.27}$$

while $E\xi_1 = \displaystyle\int_{-\infty}^{\infty} x \cdot p(x)\, dx = 0.$

Let η_t, $t \in T$, be a random function; τ a random variable taking values in $T \cup \{\infty\}$. The random function $\hat{\eta}_t$, $t \in T$, obtained by stopping η_t at the time τ is defined as

$$\hat{\eta}_t = \begin{cases} \eta_t, & t < \tau, \\ \eta_\tau, & t \geq \tau. \end{cases} \tag{27.28}$$

Theorem 27.5. *If* η_t, $t \in T$, *is a martingale* (*a submartingale*), *and* τ *a stopping time taking finitely many values, the stopped random function* $\hat{\eta}_t$, $t \in T$, *is also a martingale* (*a submartingale*).

The **proof** is similar to that of Theorems 27.3, 27.4; I omit it: we'll not be using this theorem. I gave it only in order to explain why stopping times are called so: you stop the random function at such a time, and still it is a martingale (a submartingale).

Theorem 27.6 (the Kolmogorov-type inequality). *If* η_t, $t \in T$, *is a nonnegative submartingale, and* $t_1 < t_2 < ... < t_n = t_{\max}$ *are time points in* T, *then for every* $a > 0$

$$P\{\max(\eta_{t_1}, \eta_{t_2}, ..., \eta_{t_n}) \geq a\} \leq \frac{E\eta_{t_{\max}}}{a}. \tag{27.29}$$

This is a strengthening of the Chebyshev-type inequality:

$$P\{\eta_{t_i} \geq a\} \leq \frac{E\eta_{t_i}}{a} \leq \frac{E\eta_{t_{\max}}}{a}: \tag{27.30}$$

the Chebyshev-type inequality is about *one* random variable, while (27.29) is about the maximum of an arbitrarily large number of random variables.