Lecture 38. Diffusion processes arising as solutions of stochastic equations.

Let ξ_t, $t \ge 0$, be a solution of the stochastic equation $d\xi_t = b(\xi_t)\,dt + \sigma(\xi_t)\,dW_t$ with Lipshitz coefficients and with a square-integrable initial condition ξ_0 that is measurable with respect to the σ-algebra $\mathcal{F}_0$. We are going to prove that this is a (time-homogeneous) Markov process.

But in fact, we have never just proved that such and such stochastic process is a Markov one: it was always that we proved that it is a Mrkov process *with such and such transition function.* So first of all we should write an expression for the transition function.

The value of the transition function $P(t,\,x,\,C)$ is the probability, for our process ξ_t starting at time 0 from the point x, to be in the set C at time t. In our case the solution starting at time 0 from a point x is written as $\Xi_t^x(W_\bullet)$ (see Theorem 37.1); so

$$P(t,\,x,\,C) = P\{\Xi_t^x(W_\bullet) \in C\} = E I_C\big(\Xi_t^x(W_\bullet)\big). \tag{38.1}$$

Clearly this function is a probability measure in its third argument, C.

The next thing we need is measurability of the function $P(t,\,x,\,C)$ in x.

According to Theorem 37.1, the function $\Xi_t^x(u_\bullet)$ is measurable in $(x,\,t,\,u_\bullet) \in \mathbb{R}^1 \times [0,\,\infty) \times \mathbf{C}$ with respect to the σ-algebra $\mathcal{B}^1 \times \mathcal{B}_{[0,\,\infty)} \times \mathcal{C}$. We put into this function the $\mathbf{C}$-valued random variable $W_\bullet$ instead of its third argument; so we get the function $I_C\big(\Xi_t^x\big(W_\bullet(\omega)\big)\big)$ that is $(\mathcal{B}^1 \times \mathcal{B}_{[0,\,\infty)} \times \mathcal{F})$-measurable. Of course we assume that the set C is a Borel one, so its indicator function I_C is Borel-measurable, and the function $I_C\big(\Xi_t^x(W_\bullet)\big)$ is also $(\mathcal{B}^1 \times \mathcal{B}_{[0,\,\infty)} \times \mathcal{F})$-measurable. By Fubini's Theorem, the integral of this function with respect to the probability $P(d\omega)$ is $(\mathcal{B}^1 \times \mathcal{B}_{[0,\,\infty)})$-measurable in the remaining two arguments x and t.

We don't need measurability in t right now; but from $(\mathcal{B}^1 \times \mathcal{B}_{[0,\,\infty)})$-measurability we obtain that for every t the expectation (38.1) is $\mathcal{B}^1$-measurable in x.

The next thing we need to check is that for $t_0 \ge 0$, $s \ge 0$, $C \in \mathcal{B}^1$

$$P\{\xi_{t_0+s} \in C \| \mathcal{F}_{t_0}\} = P(s,\,\xi_{t_0},\,C) \tag{38.2}$$

(almost surely, of course).

The stochastic process ξ_t was defined as the (almost-unique) solution of our stochastic equation with the initial condition ξ_0 at time 0. But this random function does take some value ξ_{t_0} at the time t_0; and at the same time ξ_t, $t \ge t_0$, is a solution of our stochastic equation with the initial condition ξ_{t_0} prescribed at the time point t_0. By Theorem 37.3 we can write (almost surely, of course) for $t \ge t_0$:

$$\xi_t = \Xi_{t-t_0}^{\xi_{t_0}}(\tilde{W}_\bullet), \qquad \xi_{t_0+s} = \Xi_s^{\xi_{t_0}}(\tilde{W}_\bullet), \tag{38.3}$$

where $\tilde{W}_t = W_{t_0+t} - W_{t_0}$.

The random variable ξ_{t_0} is measurable with respect to $\mathcal{F}_{t_0}$; and the $(\mathbf{C},\,\mathcal{C})$-valued random variable $\tilde{W}_\bullet$ is independent from this σ-algebra.

Since the σ-algebra $\mathcal{C}$ is generated by sets having to do with finitely many values of a function u_t, $t \geq 0$, the last statement is, in essence, the statement that the events

$$\{(\tilde{W}_{s_1}, ..., \tilde{W}_{s_n}) \in D\} = \{(W_{t_0+s_1} - W_{t_0}, ..., W_{t_0+s_n} - W_{t_0}) \in D\} \tag{38.4}$$

are independent from $\mathcal{F}_{t_0}$ for every $0 \leq s_1 < ... < s_n$ and $D \in \mathcal{B}^n$. This means that

$$P\{(W_{t_0+s_1} - W_{t_0}, ..., W_{t_0+s_1} - W_{t_0}) \in D \| \mathcal{F}_{t_0}\} \tag{38.5}$$

does not depend on $\omega \in \Omega$ (to be precise, that *one of the versions* of this conditional probability is just a constant), and so it is equal to the unconditional probability.

By the Markov property for the Wiener process we have:

$$P\{(W_{t_0+s_1} - W_{t_0}, ..., W_{t_0+s_1} - W_{t_0}) \in D \| \mathcal{F}_{t_0}\} = f(\xi_{t_0}), \tag{38.6}$$

where

$$\begin{aligned} f(x) = \int P(s_1, x, dy_1) \int P(s_2 - s_1, y_1, dy_2) \int ... \\ ... \int P(s_n - s_{n-1}, y_{n-1}, dy_n)\; I_D(y_1 - x, y_2 - x, ..., y_n - x) \end{aligned} \tag{38.7}$$

(all integrals are taken over the real line).

Introduce here new variables of integration $z_i = y_i - x$. Using the fact that $P(t, x, C) = P(t, 0, C - x)$ (that the transition probabilities of the Wiener process don't change if we shift both the point x and the set C by the same amount), we get:

$$\begin{aligned} f(x) = \int P(s_1, 0, dz_1) \int P(s_2 - s_1, z_1, dz_2) \int ... \\ ... \int P(s_n - s_{n-1}, z_{n-1}, dz_n)\; I_D(z_1, z_2, ..., z_n), \end{aligned} \tag{38.8}$$

which is obviously a constant (does not depend on x). So the conditional probability (38.4) is equal to the unconditional one: independence.

Now we use the result of Problem $\boxed{\mathbf{65}}$: almost surely

$$P\{\xi_{t_0+s} \in C \| \mathcal{F}_{t_0}\} = E\big(I_C\big(\Xi_s^{\xi_{t_0}}(\tilde{W}_\bullet)\big) \big\| \mathcal{F}_{t_0}\big) = G(\xi_{t_0}), \tag{38.9}$$

where

$$G(x) = EI_C\big(\Xi_s^x(\tilde{W}_\bullet)\big) = EI_C\big(\Xi_s^x(W_\bullet)\big) = P(s, x, C) \tag{38.10}$$

(according to (38.1)). This proves (38.2).

The remaining properties of a transition function ($P(0, x, C) = \delta_x(C)$, and the Chapman – Kolmogorov equation) are automatical if we can consider a Markov process for which (38.2) is satisfied starting from an arbitrary point x.

So we have a Markov process (in fact, a family of Markov processes with a common transition function, starting at time 0 from all possible points x) with trajectories that are almost surely continuous. The solution ξ_t^x with the initial condition $\xi_0^x = x$ depends on the initial condition x continuously (in the mean squares: $E(\xi_t^x - \xi_t^y)^2 \le 2(x-y)^2 \cdot e^{2C^2(1+\sqrt{T})^2 t}$ – see formula (32.6)). It follows from this that the limit holds in probability: $\lim_{y \to x}(P)\, \xi_t^y = \xi_t^y$, and from this that weak continuity holds for the corresponding distributions: $P(t,\, y,\, \bullet) \to^w P(t,\, x,\, \bullet)$ as $y \to x$. This means that the Markov process ξ_t satisfies the Feller property: $P^t\mathbf{C} \subseteq \mathbf{C}$ (where $\mathbf{C}$, in contrast with our recent use of the same letter, is the space of *bounded* continuous functions). It follows from this that the process ξ_t is a strong Markov one (we had a theorem about a Feller Markov process with right-continuous trajectories being strong Markov; here only *almost* all trajectories are continuous, but the result remains true. At the time of formulating and proving that result I did not foresee the need to generalize it allowing right-continuity of only *almost* all trajectories). In the one-dimensional case we get that the process arising as the solution of the stochastic equation $d\xi_t = b(\xi_t)\,dt + \sigma(\xi_t)\,dW_t$ is a diffusion process with generating differential operator $Lf(x) = \dfrac{\sigma(x)^2}{2} f''(x) + b(x)\,f'(x)$.

Of course, the results of this lecture are true for multidimensional processes as well: if $\boldsymbol{\xi}_t = (\xi_t^1,\, ...,\, \xi_t^d)$ is the process arising as the solution of the system

$$d\xi_t^i = b_i(\boldsymbol{\xi}_t)\,dt + \sum_{k=1}^{r} \sigma_{ik}(\boldsymbol{\xi}_t)\,dW_t^k, \qquad i = 1,\, ...,\, d, \tag{38.11}$$

or the vector equation

$$d\boldsymbol{\xi}_t = \boldsymbol{b}(\boldsymbol{\xi}_t)\,dt + \sigma(\boldsymbol{\xi}_t)\,d\boldsymbol{W}_t, \tag{38.12}$$

it is a Markov process in $\mathbb{R}^d$ with trajectories that are almost surely continuous.

We cannot state that this is a diffusion process, and write its generating operator before writing the multidimensional Itô formula. Let me do it now.

If $\boldsymbol{\xi}_t = (\xi_t^1,\, ...,\, \xi_t^d)$, $t \ge t_0$,

$$d\xi_t^i = f_i(t,\, \omega)\,dt + \sum_{k=1}^{r} \sigma_{ik}(t,\, \omega)\,dW_t^k, \qquad i = 1,\, ...,\, d, \tag{38.13}$$

and $F(t,\, \boldsymbol{x}) = F(t,\, x_1,\, ...,\, x_d)$ is a function that is continuously differentiable once in t and twice in the spatial variables, we have:

$$\begin{aligned} dF(t,\, \boldsymbol{\xi}_t) &= \\ \Big[\frac{\partial F}{\partial t}(t,\, \boldsymbol{\xi}_t) &+ \sum_{i=1}^{d} \frac{\partial F}{\partial x_i}(t,\, \boldsymbol{\xi}_t) \cdot f_i(t,\, \omega) + \frac{1}{2} \sum_{i,\, j=1}^{d} \sum_{k=1}^{r} \frac{\partial^2 F}{\partial x_i \partial x_j}(t,\, \boldsymbol{\xi}_t) \cdot g_{ik}(t,\, \omega)\, g_{jk}(t,\, \omega)\Big]\,dt \\ &+ \sum_{i=1}^{d} \sum_{k=1}^{r} \frac{\partial F}{\partial x_i}(t,\, \boldsymbol{\xi}_t) \cdot g_{ik}(t,\, \omega)\,dW_t^k - \end{aligned} \tag{38.14}$$

provided that

$$\int_{t_0}^{t} E\big[\frac{\partial F}{\partial x_i}(s,\, \boldsymbol{\xi}_t) \cdot g_{ik}(s,\, \omega)\big]^2\, ds < \infty, \qquad 1 \le i \le d, \quad 1 \le k \le r, \tag{38.15}$$

and that the functions $\dfrac{\partial F}{\partial t}(t,\, \boldsymbol{\xi}_t)$, $\dfrac{\partial F}{\partial x_i}(t,\, \boldsymbol{\xi}_t) \cdot f_i(t,\, \omega)$, $\dfrac{\partial^2 F}{\partial x_i \partial x_j}(t,\, \boldsymbol{\xi}_t) \cdot g_{ik}(t,\, \omega)\, g_{jk}(t,\, \omega)$ are almost surely Lebesgue integrable.

When we had our "preliminary investigation" for the Itô formula, we wrote tentative stochastic integrals with respect to $(dW_t)^2$ in the one-dimensional case, and with respect to $dW_t^k\, dW_t^l$ in the multidimensional case. It turned out that $(dW_t)^2$ (and also, of course, $(dW_t^k)^2$) should be replaced with ds. It turns out (and the proof of it relies on Problem $\boxed{\mathbf{61}}$) that we should replace $dW_t^k\, dW_t^l$ for $k \neq l$ with 0.

Using formula (38.14), we can write the generating differential operator L for the multidimensional process arising as the solution of (38.11), (38.12): for a smooth function $F(t,\, \boldsymbol{x})$

$$dF(t,\, \boldsymbol{\xi}_t) = \big[\frac{\partial F}{\partial t}(t,\, \boldsymbol{\xi}_t) + LF(t,\, \bullet)(\boldsymbol{\xi}_t)\big]\, dt + \ \ldots \ d\boldsymbol{W}_t, \tag{38.16}$$

where

$$Lf(\boldsymbol{x}) = \frac{1}{2} \sum_{i,\, j=1}^{d} a_{ij}(\boldsymbol{x}) \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}) + \sum_{i=1}^{d} b_i(\boldsymbol{x}) \frac{\partial f}{\partial x_i}(\boldsymbol{x}), \tag{38.17}$$

$$a_{ij}(\boldsymbol{x}) = \sum_{k=1}^{r} \sigma_{ik}(\boldsymbol{x}) \cdot \sigma_{jk}(\boldsymbol{x}). \tag{38.18}$$

In the matrix form the last equality is written as

$$\big(a_{ij}(\boldsymbol{x})\big) = \sigma(\boldsymbol{x}) \cdot \sigma(\boldsymbol{x})^T \tag{38.19}$$

(^{T} meaning the transposed matrix).

The diffusion matrix $\big(a_{ij}(\boldsymbol{x})\big)$ is symmetric, and not necessarily positive definite, but always non-negative definite: for a row vector $\boldsymbol{\lambda} = (\lambda_1,\, ...,\, \lambda_d)$

$$\sum_{i,\, j=1}^{d} a_{ij}(\boldsymbol{x}) \cdot \lambda_i \lambda_j = \boldsymbol{\lambda} \cdot \big(a_{ij}(\boldsymbol{x})\big) \cdot \boldsymbol{\lambda}^T = \boldsymbol{\lambda}\sigma(\boldsymbol{x}) \cdot \sigma(\boldsymbol{x})^T \boldsymbol{\lambda}^T = \big(\boldsymbol{\lambda}\sigma(\boldsymbol{x})\big) \cdot \big(\boldsymbol{\lambda}\sigma(\boldsymbol{x})\big)^T, \tag{38.20}$$

which is equal to the square of the length of the r-dimensional vector $\boldsymbol{\lambda}\sigma(\boldsymbol{x})$, and is always nonnegative.

We have two approaches to diffusion processes: one based on the results about partial differential equations of the parabolic type; and another based on stochastic equations. The second approach is more direct, and it allows us to consider a wider class of processes – the main difference being that this approach is not sensitive to degeneration of the diffusion matrix $\big(a_{ij}(\boldsymbol{x})\big)$. But there are some points in favor of the PDE approach: e. g., we obtain by it the existence of the transition density $p(t,\, \boldsymbol{x},\, \boldsymbol{y})$, with good estimates for this density and its derivatives.

In the next lecture I'll show how stochastic equations are used to handle elliptic equations not only of the form (28.39), but also

$$\begin{gathered}\frac{1}{2}\sum_{i,\,j=1}^{d} a_{ij}(\boldsymbol{x})\cdot\frac{\partial^2 u}{\partial x_i \partial x_j}(\boldsymbol{x})+\sum_{i=1}^{d} b_i(\boldsymbol{x})\cdot\frac{\partial u}{\partial x_i}(\boldsymbol{x})+c(\boldsymbol{x})\cdot u(\boldsymbol{x})=g(\boldsymbol{x}),\quad \boldsymbol{x}\in G,\\ u(\boldsymbol{x})=\varphi(\boldsymbol{x}),\quad \boldsymbol{x}\in\partial G.\end{gathered}\tag{38.21}$$

Not that such equations couldn't be handled using our semigroup approach; but with stochastic equations this seems to be easier.