Lecture 35. Examples of diffusion processes arising as solutions of stochastic equations. Introduction to existence and uniqueness results.

The modern theory of differential equations is not so much about "solving" the equations, i. e., writing explicit expressions for their solutions, but more about methods for finding the solutions numerically (especially so in our computer age), about establishing qualitative properties of the solutions: say, what happens with them as $t \to \infty$ (that is, on time intervals that are beyond the reach of any numerical method), etc. The same with *stochastic* equations: very seldom can we "solve" a stochastic equation, that is, write an explicit expression for the solution; we may be more interested in numerical methods, etc.; but the main thing for us is using stochastic equations to describe some more complicated stochastic processes ξ_t using simpler ones (namely, the Wiener process). However, let us consider first some examples in which the stochastic equation can be solved explicitly.

Example 35.1. Let us consider the equation

$$d\xi_t = b\,dt + \sigma\,dW_t, \tag{35.1}$$

where b and σ are constants. Let ξ_0 be a random variable that is independent with the family of random variables W_t, $t \geq 0$. If we introduce the non-decreasing family of σ-algebras $\mathcal{F}_t = \sigma(\xi_0;\, W_s, 0 \leq s \leq t)$, the Wiener process W_t will still be a Wiener process with respect to this family of σ-algebras (i. e., W_t is adapted to $(\mathcal{F}_t)$, and $P\{W_u \in C \| \mathcal{F}_t\} = P(u-t,\, W_t,\, C)$ for $0 \leq t \leq u$; the fact has been mentioned before).

The solution of equation (35.1) with the initial condition ξ_0 at time 0 can be written as

$$\xi_t = \xi_0 + \sigma \cdot (W_t - W_0) + b \cdot t \tag{35.2}$$

(indeed, (35.2) can be rewritten as $\xi_t = \xi_0 + \int_0^t \sigma\,dW_s + \int_0^t b\,ds$, and by definition this means (35.1)). The increment of the process ξ_t on a time interval from t to $s > t$ is equal to $\sigma \cdot (W_s - W_t) + b \cdot (s-t)$, and it is is a normal random variable with parameters $\big(b \cdot (s-t),\, \sigma^2 \cdot (s-t)\big)$, so $|\sigma|\sqrt{s-t}$ is its standard deviation. So σ (or rather its absolute value) is the standard deviation of the increment of our process per time unit. This is the reason why the traditional notation for this coefficient, and, in the case of a general stochastic equation (34.23), for the corresponding function is σ, $\sigma(x)$. The same for the multidimensional case (equations (30.6), (31.3) – (31.4)): the notations $\sigma_{ik}(\boldsymbol{x})$ and the matrix $\sigma(\boldsymbol{x})$ are usually used.

What about the first parameter, $b \cdot (s-t)$, of our normal distribution? Suppose we have a river with constant velocity b of flow. A particle in the water, engaged in Brownian motion, will be described by a Wiener process W_t in a coordinate system moving together with the flow, but with respect to a fixed coordinate system, the particle will be dragged by the water, so it will be described rather by the random function ξ_t (if the Brownian motion of the particle is described not by a standard Wiener process, but has standard deviation equal to σ per time unit). In the pseudo-nautical language, we can call this *drift*; and the coefficient b, be it constant, as in (35.1), (35.2), or variable: $b = b(x)$, is called *the drift coefficient*.

It is pretty easy to show that the process (35.2) is a Markov one (with transition density $p(t, x, y) = \dfrac{1}{\sqrt{2\pi\sigma^2 t}}\, e^{-(y-(x-bt))^2/2\sigma^2 t}$; I am not giving the proof here); so it is a diffusion process. The generating operator corresponding to this process is $Lf(x) = \dfrac{\sigma^2}{2} \cdot f''(x) + b \cdot f'(x)$. The coefficient $a = \sigma^2$ is called the *diffusion coefficient*. The same name is used for the variable coefficient $a(x)$ in the case of a general one-dimensional diffusion process with generator $Lf(x) = \dfrac{a(x)}{2} \cdot f''(x) + b(x) \cdot f'(x)$; the matrix $\big(a_{ij}(x)\big)$ of the coefficients in the multidimensional generating operator

$$Lf(\boldsymbol{x}) = \frac{1}{2} \sum_{i,\, j=1}^{d} a_{ij}(\boldsymbol{x}) \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}) + \sum_{i=1}^{d} b_i(\boldsymbol{x}) \cdot \frac{\partial f}{\partial x_i}(\boldsymbol{x}) \tag{35.3}$$

is called *the diffusion matrix*.

The name of *diffusion process* is used because the movement of a single particle in the physical process of *diffusion* (penetration of some substance into the space occupied by another one) is similar to that of a particle in Brownian motion; only there may be *drift* involved, accounting for the summand $b(\xi_t)\, dt$, and the rate of the chaotic movements similar to Brownian motion may depend on the point of the space region where the particle is at that time.

Diffusion processes provide mathematical models for many problems in which stochastic processes appear. Let me show an application to physics:

Example 35.2. As I told you, the Wiener process provides a mathematical model for the phenomenon of Brownian motion. Of course, as every mathematical model, it works only in some approximation. In particular, in this model a Brownian particle *has no velocity*: the derivative $\dfrac{dW_t}{dt}$ does not exist.

Real-world particles do have velocities; and the next approximation is the model that takes this into account.

Let us consider a particle that moves on the line, being pushed chaotically from the right and from the left by molecules. Let the velocity of the particle at a time t be η_t. If $\eta_t > 0$, i. e., the particle moves to the right at time t, then it meets more molecules hitting it from the right than the number of molecules that hit it from the left; and a force will be applied to it directed to the left. If $\eta_t < 0$, a force will be applied to the particle directed to the right. As the first approximation, we can assume that this force is proportional to the velocity, with some coefficient of proportionality μ:

$$F = -\mu\, \eta_t. \tag{35.4}$$

If we didn't take into account the chaotic character of impulses given to the particle by the molecules hitting it, we would write:

$$\frac{d\eta_t}{dt} = -\frac{\mu}{m}\, \eta_t, \tag{35.5}$$

where m is the mass of the particle. This equation can be seen in every textbook of ordinary differential equations, and it describes the velocity of an object moving in a viscous fluid (including gasses, even though their viscosity is small) without taking into account the chaotic, random character of the molecule hits on the particle.

This random character is taken care of by adding a stochastic term:

$$d\eta_t = -c\,\eta_t\,dt + \sigma\,dW_t, \tag{35.6}$$

where $c = \dfrac{\mu}{m}$. The coefficient σ here depends on how many molecules there are per unit of length, and of the ratios of their masses to the mass m of the particle.

So the velocity of a particle in a physical Brownian motion is described, in the next approximation, by (35.6); and the solution of this equation is a diffusion process. The coefficient $b(x)$ is a linear function: $b(x) = -\,c \cdot x$, and $\sigma(x) \equiv \sigma = \text{const}$.

As for the *position* of the particle at time t, it is described by

$$\xi_t = \xi_0 + \int_0^t \eta_s\,ds. \tag{35.7}$$

The equation (35.6) is *linear*: its right-hand side depends on the unknown function linearly.

It turns out that, in a rare exception, we can "solve" this equation: that is, write an explicit formula for it. However, what does it mean: "an explicit formula"? In the case of ordinary differential equations, we are satisfied if the expression for the solution involves some integrals, even if we cannot write an "explicit formula" for the anti-derivative; so here we should be satisfied if the answer is expressed in terms of Riemann integrals and stochastic integrals.

The equation (35.6) means that

$$\eta_t = \eta_0 - c\int_0^t \eta_s\;ds + \int_0^t \sigma\;dW_s. \tag{35.8}$$

We may try to solve the equation (35.8) – which is equivalent to the stochastic differential equation (35.6) plus the initial condition at the time point 0 – just the way we would were the function W_t differentiable.

The solution η_t of the initial-value problem

$$\frac{d\eta_t}{dt} = -c\,\xi_t + \sigma \cdot \frac{dW_t}{dt}, \qquad \eta_0 = v_0 \tag{35.9}$$

where W_t (and $\dfrac{dW_t}{dt}$) is a known function, is written, according to the general rule of solving linear first-order equations, as

$$\eta_t = v_0 \cdot e^{-ct} + \sigma \int_0^t e^{-c(t-s)}\,\frac{dW_s}{ds}\;ds; \tag{35.10}$$

this can be rewritten as

$$\eta_t = v_0 \cdot e^{-ct} + \sigma \int_0^t e^{-c(t-s)}\, dW_s. \tag{35.11}$$

While formula (35.10) did not really make sense for the Wiener process, because it is not differentiable, formula (35.11) does make sense, the integral being a stochastic one. The formula is not proved yet, because in trying to write it we have made false assumptions ("*if* the Wiener process were differentiable"); but we can prove it using Itô's formula.

We consider the stochastic process

$$\zeta_t = \int_{t_0}^t e^{cs}\, dW_s; \tag{35.12}$$

its stochastic differential is

$$d\zeta_t = e^{ct}\, dW_t. \tag{35.13}$$

Taking the function

$$F(t,\, x) = v_0 \cdot e^{-ct} + \sigma e^{-ct} \cdot x, \tag{35.14}$$

we can write that

$$\eta_t = F(t,\, \zeta_t). \tag{35.15}$$

The rest is left to you as a problem:

64 Applying Itô's formula to the random function (35.15), prove that for

$$\eta_t = v_0 e^{-ct} + \sigma \int_0^t e^{-c(t-s)}\, dW_s, \qquad t \geq 0, \tag{35.16}$$

we have:

$$d\eta_t = -c\eta_t\, dt + \sigma\, dW_t. \tag{35.17}$$

The process η_t turns out to be a Markov one: a diffusion process (I won't speak about it now, waiting until we have a general statement about it). Its generating differential operator is

$$Lf(x) = \frac{\sigma^2}{2} f''(x) - cx\, f'(x). \tag{35.18}$$

We couldn't have handled this diffusion process using the PDE results given in Lecture note 21, because the drift coefficient $b(x) = -\,cx$ is not bounded – but perhaps we could look for PDE results for parabolic equations that allow the coefficients to grow linearly. We'll see that we can handle it using stochastic equations with ease.

Example 35.3. It turns out that the stochastic process given by formula (35.7) is *not* a diffusion process; but the *pair* $(\xi_t,\, \eta_t)$ is one.

What is the system of stochastic equations that generates this process? It is

$$\begin{aligned} d\xi_t &= \eta_t\, dt, \\ d\eta_t &= -\frac{\mu}{m}\, \eta_t\, dt + \sigma\, dW_t, \end{aligned} \tag{35.19}$$

or, in the vector-matrix form,

$$d\begin{pmatrix}\xi_t\\ \eta_t\end{pmatrix} = \begin{pmatrix}\eta_t\\ -\dfrac{\mu}{m}\,\eta_t\end{pmatrix} dt + \begin{pmatrix}0\\ \sigma\end{pmatrix} dW_t. \tag{35.20}$$

Applying the Itô formula (34.22), we can write the generating differential operator for the two-dimensional process (ξ_t, η_t): it is given by formula (35.3) with $d = 2$, the drift $\boldsymbol{b}(\boldsymbol{x}) = \boldsymbol{b}(x, v) = \begin{pmatrix}v\\ -\dfrac{\mu}{m}\,v\end{pmatrix}$ and the diffusion matrix $\big(a_{ij}(\boldsymbol{x})\big) \equiv \begin{pmatrix}0 & 0\\ 0 & \sigma^2\end{pmatrix}$:

$$Lf(x, v) = v\cdot\frac{\partial f}{\partial x}(x, v) - cv\cdot\frac{\partial f}{\partial v}(x, v) + \frac{\sigma^2}{2}\cdot\frac{\partial^2 f}{\partial v^2}(x, v). \tag{35.21}$$

The matrix $\big(a_{ij}(\boldsymbol{x})\big)$ is *not* positive-definite, so we couldn't apply here any results about parabolic differential equations; again, handling this process using stochastic equations is quite easy.

Now we go to the results on existence and uniqueness of solutions for stochastic equations. But first let me remind you the situation in the ordinary differential equations.

One of approaches to constructing solutions of differential equations is to replace the initial-value problem

$$\frac{dx_t}{dt} = b(x_t), \qquad x_{t_0} = x \tag{35.22}$$

with the integral equation equivalent to it:

$$x_t = x + \int_{t_0}^{t} b(x_s)\, ds. \tag{35.23}$$

Suppose the right-hand side function $b(x)$ satisfies a Lipschitz condition:

$$|b(x) - b(y)| \le C\cdot|x - y|, \tag{35.24}$$

where C is a constant.

Then the existence of a solution can be proved by using *successive approximations*:

$$x_t^{(0)} \equiv x, \tag{35.25}$$

and for $n \ge 1$

$$x_t^{(n)} = x + \int_{t_0}^{t} b(x_s^{(n-1)})\, ds. \tag{35.26}$$

Both for $t \ge t_0$ and for $t < t_0$ we have:

$$|x_t^{(1)} - x_t^{(0)}| = \left|\int_{t_0}^{t} b(x)\, ds\right| = |b(x)|\cdot|t - t_0|; \tag{35.27}$$

for $n > 1$ we have (for a little shorter notations, we consider only $t \geq t_0$):

$$|x_t^{(n)} - x_t^{(n-1)}| = \Big|\int_{t_0}^{t} \big(b(x_s^{(n-1)}) - b(x_s^{(n-2)})\big)\ ds\Big| \leq \int_{t_0}^{t} C \cdot |x_s^{(n-1)} - x_s^{(n-2)}|\ ds, \quad (35.28)$$

and by induction we prove that

$$|x_t^{(n)} - x_t^{(n-1)}| \leq K \cdot C^{n-1} \cdot \frac{|t - t_0|^n}{n!}, \quad (35.29)$$

where K is a constant. Since the series

$$\sum_{n=1}^{\infty} K \cdot C^{n-1} \cdot \frac{|t - t_0|^n}{n!} = \frac{K}{C} \cdot e^{C|t-t_0|} \quad (35.30)$$

converges, we see that the series

$$x + (x_t^{(1)} - x_t^{(0)}) + (x_t^{(2)} - x_t^{(1)}) + ... + (x_t^{(n)} - x_t^{(n-1)}) + ... \quad (35.31)$$

converges, uniformly in t changing in every finite interval, that is, there exists the limit

$$x_t = \lim_{n \to \infty} x_t^{(n)}. \quad (35.32)$$

Limit passage in (35.26) yields (35.23).

As for proving the uniqueness of the solution, we are able to get at the same price also its continuous dependence on the initial condition x: if x_t is a solution of the initial-value problem (35.22), or, which is the same, of the integral equation (35.23), and y_t is the solution of

$$\frac{dy_t}{dt} = b(y_t), \qquad y_{t_0} = y, \quad (35.33)$$

then we have:

$$|x_t - y_t| \leq |x - y| \cdot e^{C|t-t_0|}. \quad (35.34)$$

This is obtained by subtracting (35.23) and the corresponding integral equation for y_t (again I'll write this only for $t \geq t_0$ even if the same is true for $t < t_0$):

$$x_t - y_t = (x - y) + \int_{t_0}^{t} \big(b(x_s) - b(y_s)\big)\ ds, \quad (35.35)$$

$$|x_t - y_t| \leq |x - y| + \int_{t_0}^{t} C \cdot |x_s - y_s|\ ds \quad (35.36)$$

This is an integral *inequality*; it turns out that its every solution is less or equal that the solution of the corresponding integral *equation*:

$$|x_t - y_t| \leq U_t, \quad (35.37)$$

where

$$U_t = |x - y| + \int_{t_0}^{t} C \cdot U_s \, ds. \tag{35.38}$$

The equation (35.38) is equivalent to the initial-value problem

$$\frac{dU_t}{dt} = C \cdot U_t, \qquad U_{t_0} = |x - y|, \tag{35.39}$$

so

$$U_t = |x - y| \cdot e^{C(t - t_0)} \tag{35.40}$$

(I'll remind you that we consider now only the case of $t \geq t_0$).

To prove (35.37) it is enough to prove that for every $\varepsilon > 0$, $t \geq t_0$

$$|x_t - y_t| < (|x - y| + \varepsilon) \cdot e^{C(t - t_0)}. \tag{35.41}$$

At $t = t_0$ this is true. Suppose at some $t_1 > t_0$ it is not. The left-hand side of (35.41) and the right-hand side are continuous functions, so there is the smallest t_* such that for $t = t_*$ (35.41) does not hold; at this point the left-hand side is *equal* to the right-hand side. We have:

$$\begin{aligned} 0 &= |x_{t_*} - y_{t_*}| - (|x - y| + \varepsilon) \cdot e^{C(t_* - t_0)} \\ &\leq |x - y| + \int_{t_0}^{t_*} C \cdot |x_s - y_s| \, ds \\ &\qquad - (|x - y| + \varepsilon) - \int_{t_0}^{t_*} C \cdot (|x - y| + \varepsilon) \cdot e^{C(s - t_0)} \, ds \\ &< -\varepsilon < 0, \end{aligned} \tag{35.42}$$

because $|x_s - y_s| < (|x - y| + \varepsilon) \cdot e^{C(s - t_0)}$ for all $s \in [t_0, t^*)$. The contradiction obtained proves that it is impossible that inequality (35.37) shouldn't hold; then since a positive ε can be chosen arbitrarily small, we get (35.37).

The uniqueness is a particular case with $|x - y| = 0$.

We know that there are other theorems of existence (and uniqueness) for ordinary differential equations and that the Lipschitz condition (35.24) is not *necessary* for the existence and uniqueness to take place.