Lecture note 12. Diffusion processes and partial differential equations.

Suppose X_t is a diffusion process: $dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$. Itô's formula gives us for a smooth function $u(t,x)$ (i. e., a function having continuous partial derivatives $\dfrac{\partial u}{\partial t}$, $\dfrac{\partial u}{\partial x}$, $\dfrac{\partial^2 u}{\partial x^2}$) under some restrictions about certain expectations being finite:

$$\begin{aligned} du(t, X_t) = \frac{\partial u}{\partial t}(t, X_t)\,dt + \frac{\partial u}{\partial x}(t, X_t) \cdot b(t, X_t)\,dt \\ + \frac{\partial u}{\partial x}(t, X_t) \cdot \sigma(t, X_t)\,dW_t + \frac{1}{2}\,\frac{\partial^2 u}{\partial x^2}(t, X_t) \cdot \sigma(t, X_t)^2\,dt. \end{aligned} \tag{12.1}$$

We see that of the summands with dt (of which we take the non-stochastic – the Riemann – integral) the first is standard, being the same for every diffusion process; but the terms including the derivatives $\dfrac{\partial u}{\partial x}$, $\dfrac{\partial^2 u}{\partial x^2}$ do depend on the coefficients in the stochastic equation satisfied by our process. Let us introduce, for each t ($\geq t_0$) a second-order differential operator L_t acting on twice differentiable functions of x:

$$L_t f(x) = b(t,x) \cdot f'(x) + \frac{1}{2}\,\sigma(t,x)^2 \cdot f''(x) = \frac{a(t,x)}{2}\,f''(x) + b(t,x)\,f'(x), \tag{12.2}$$

where $a(t,x) = \sigma(t,x)^2$. The differential (12.1) can be rewritten in the form

$$du(t, X_t) = \big[\frac{\partial u}{\partial t}(t, X_t) + L_t u(t, X_t)\big]\,dt + \text{ something } dW_t, \tag{12.3}$$

where the operator L_t is applied to the function $u(t, \bullet)$, for fixed t, in its x argument (and the resulting function is taken at the point X_t).

It turns out that the function $a(t,x) = \sigma(t,x)^2$ is a more essential characteristic of the diffusion process than $\sigma(t,x)$. How can we see that?

We have spoken of uniqueness of the solution X_t of a stochastic equation; but we haven't spoken of uniqueness of the equation representing a given diffusion process. Let us speak about it. Is it possible that a diffusion process X_t is a solution of the stochastic equation (11.5), and at the same time a solution of another stochastic equation:

$$dX_t = \tilde{b}(t, X_t)\,dt + \tilde{\sigma}(t, X_t)\,d\tilde{W}_t \tag{12.4}$$

with some other coefficients $\tilde{b}$ and $\tilde{\sigma}$ and some other Wiener process $\tilde{W}_t$? It turns out that yes, possible. For example, we can take

$$\tilde{W}_t = -W_t, \qquad \tilde{\sigma}(t,x) = -\sigma(t,x), \tag{12.5}$$

and if the equation (11.5) holds, so will (12.5). (It is easy to understand that $-W_t$ is also a Wiener process.) It turns out that also if $h(t,x)$ is a function taking only values ± 1, we can take

$$\tilde{\sigma}(t,x) = h(t,x) \cdot \sigma(t,x), \qquad \tilde{W}_t = \int_{t_0}^{t} h(s, X_s)\, dW_s, \tag{12.6}$$

$\tilde{W}_t$ will be another Wiener process (I don't prove it: it wasn't in the original papers by Itô and Gikhman, but can be found in the book *Stochastic Processes* by J.L.Doob). The coefficient $b(t,x)$ remains the same for all representations of a diffusion process by means of a stochastic equation, and so does the coefficient $a(t,x) = \sigma(t,x)^{\mathbf{2}}$ (in contrast with $\sigma(t,x)$), so *these* coefficients should characterise a diffusion process rather than b and σ.

We introduce the names for the coefficients a, b: the coefficient $b(t,x)$ is called the *drift coefficient* (or just *the drift*), and $a(t,x)$ *the diffusion coefficient.* The diffusion coefficient is always nonnegative; the drift can have any sign it wishes.

There are other names for the coefficients $b(t,x)$ and $a(t,x)$: the *local mean* and the *local variance.* Let me explain what these names mean.

It turns out, under the restriction of $a(t,x)$ and $b(t,x)$ being continuous, plus some mild supplementary restrictions, that for any t_0 and x_0, for a diffusion process X_t starting at time t_0 at the point x_0

$$E\big(X_t\big) = x_0 + b(t_0,\, x_0)\cdot(t - t_0) + o(t - t_0), \quad \mathrm{Var}(X_t) = a(t_0,\, x_0)\cdot(t - t_0) + o(t - t_0) \tag{12.7}$$

as $t \to t_0^+$ (I hope that you know the notation $o(\ \ )$ for a function that is infinitely small compared to the expression inside the parentheses).

Indeed, because

$$X_t = x_0 + \int_{t_0}^{t} b(s, X_s)\, ds + \int_{t_0}^{t} \sigma(s, X_s)\, dW_s, \tag{12.8}$$

we have:

$$E(X_t) = x_0 + \int_{t_0}^{t} E\big(b(s, X_s)\big)\, ds \tag{12.9}$$

(the expectation of the stochastic integral is equal to 0). As $t \to t_0$, the variable s, being between t_0 and t, also goes to t_0. Because the trajectory X_s is continuous, and the function $b(s,x)$ is continuous in the pair (s,x), we have almost surely: $b(s, X_s) \to b(t_0, X_{t_0}) = b(t_0, x_0)$. From this, the first equality in (12.7) follows (under some mild restrictions allowing us to use either mean-square convergence of $b(s, X_s)$ to $b(t_0, x_0)$, or Theorem 2.2).

As for (12.8), let us apply formula (12.1) to the function $u(t,x) = (x - x_0)^2$:

$$d(X_t - x_0)^2 = \big[b(t, X_t)\cdot 2(X_t - x_0) + \frac{1}{2}\,\sigma(t, X_t)^2 \cdot 2\big]\, dt + \sigma(t, X_t)\cdot 2(X_t - x_0)\, dW_t. \tag{12.10}$$

This means that

$$(X_t - x_0)^2 = (X_0 - x_0)^2 + \int_{t_0}^{t} [b(s, X_s) \cdot 2(X_s - x_0) + \sigma(s, X_s)^2] \, ds \\ + \int_{t_0}^{t} \sigma(s, X_s) \cdot 2(X_s - x_0) \, dW_s. \tag{12.11}$$

Here $X_{t_0} - x_0 = 0$, $\sigma^2 = a$. Taking the expectation, we get:

$$E\big((X_t - x_0)^2\big) = \int_{t_0}^{t} E\big(2b(s, X_s) \cdot (X_s - x_0) + a(s, X_s)\big) \, ds. \tag{12.12}$$

As t, and with it, $s \to t_0^+$, we have $2b(s, X_s) \cdot (X_s - x_0) \to 0$, $a(s, X_s) \to a(t_0, x_0)$, $E\big(2b(s, X_s) \cdot (X_s - x_0) + a(s, X_s)\big) \to a(t_0, x_0)$, and we get that

$$E\big((X_t - x_0)^2\big) = a(t_0, x_0) \cdot (t - t_0) + o(t - t_0). \tag{12.13}$$

This is not our statement about the variance yet; but

$$\begin{aligned} \operatorname{Var}(X_t) = \operatorname{Var}(X_t - x_0) &= E\big((X_t - x_0)^2\big) - \big(E(X_t - x_0)\big)^2 \\ &= a(t_0, x_0) \cdot (t - t_0) + o(t - t_0) - O\big((t - t_0)^2\big) \\ &= a(t_0, x_0) \cdot (t - t_0) + o(t - t_0). \end{aligned} \tag{12.14}$$

So the coefficients $b(t_0, x_0)$, $a(t_0, x_0)$ are determined uniquely by the diffusion process at the point (t_0, x_0): at an arbitrary point (t, x) if we can consider this diffusion process starting at an arbitrary time from an arbitrary point.

And $\sigma(t, x)$, let me repeat it once more, is not determined uniquely by the diffusion process.

Now let us go to partial differential equations.

The differential equality (12.3) means that for $t \geq t_0$, the initial value X_0 being a non-random point $x_0 \in \mathbb{R}$,

$$u(t, X_t) - u(t_0, x_0) = \int_{t_0}^{t} \Big[\frac{\partial u}{\partial t}(s, X_s) + L_s u(s, X_s)\Big] \, ds + \int_{t_0}^{t} \frac{\partial u}{\partial x}(s, X_s) \cdot \sigma(s, X_s) \, dW_s, \tag{12.15}$$

provided

$$\int_{t_0}^{t} E\Big(\Big[\frac{\partial u}{\partial x}(s, X_s) \cdot \sigma(s, X_s)\Big]^2\Big) \, ds < \infty \tag{12.16}$$

(see Theorem 10.1). Let us take $t = T > t_0$, denote the variable of integration with t instead of s (not that this will make any difference), and take the expectation of both sides of (12.15). We know that the expectation of a stochastic integral is always equal to 0, so we have:

$$E\big(u(T, X_T)\big) = u(t_0, x_0) + E\Big(\int_{t_0}^{T} \Big[\frac{\partial u}{\partial t}(t, X_t) + L_t u(t, X_t)\Big] \, dt\Big) \tag{12.17}$$

(if the expectation $E\big(u(T, X_T)\big)$ exists). If we have

$$\frac{\partial u}{\partial t}(t, x) + L_t u(t, x) = g(t, x), \tag{12.18}$$

we can rewrite (12.17) as

$$u(t_0, x_0) = E\Big(u(T, X_T) - \int_{t_0}^{T} g(t, X_t)\; dt\Big). \tag{12.19}$$

The equation (12.18) is a partial differential equation; and the equality (12.19) establishes a link between diffusion processes and partial differential equations. This equality can be used in two opposite directions: on one hand, we can use what is known about partial differential equations to study diffusion processes; on the other, we can go from what we know about diffusion processes to get some new information about solutions of PDEs. For example, if we are able to *simulate* a diffusion process on a computer, we can find the solution of the equation at the point (t_0, x_0) as the expectation of a certain functional of our diffusion process (of the functional within the large parentheses in (12.19)): to get an approximate value of the expectation, we repeat our simulation independently very many times, and take the arithmetic mean of the values obtained.

An important class of partial differential equations is that of *parabolic equations.* These are equations for functions of two or more variables, one of which is interpreted as time (and denoted usually with t); the left-hand side of such an equation is the time derivative of the unknown function, and in the right-hand side stand a second-order differential operator applied in the rest of the variables, plus some known function:

$$\frac{\partial u}{\partial t}(t, x) = \mathcal{L}_t u(t, \bullet)(x) + f(t, x) \tag{12.20}$$

(I am describing, and we will be using, only *linear* parabolic equations). The operators $\mathcal{L}_t$ for every t are supposed to satisfy some conditions.

Let us restrict ourselves to the case of the "space" variable x being one-dimensional.

A linear second-order operator has the form:

$$\mathcal{L}_t f(x) = \frac{\alpha(t, x)}{2} f''(x) + \beta(t, x)\, f'(x) + \gamma(t, x) f(x). \tag{12.21}$$

The great controversy between specialists in stochastic processes and those in differential equations is whether to write the coefficient by the highest (second) derivative *with* a factor $1/2$, or without. You can see that I fancy this $1/2$ (we saw that it appears quite naturally in our formulas), and you can guess that the specialists in differential equations don't use this factor. Another controversy is whether to write the independent variables in order (t, x), or (x, t). But this does not preclude mutual understanding and cooperation.

An equation of the form (12.20) is called a *parabolic* equation if the coefficient $\alpha(t, x)$ by the highest (second) derivative is either everywhere positive, or everywhere negative.

Our equation (12.18) can be written as

$$\frac{\partial u}{\partial t}(t,x) = -\frac{a(t,x)}{2}\,\frac{\partial^2 u}{\partial x^2}(t,x) - b(t,x)\,\frac{\partial u}{\partial x}(t,x) + g(t,x), \tag{12.22}$$

with $\alpha = -a$, $\beta = -b$, $\gamma \equiv 0$, $f = g$, and it is a parabolic equation (with an everywhere negative α) if $a(t,x)$ is everywhere positive (this means that $\sigma(t,x) \neq 0$). But $a(t,x)$ is always nonnegative; if it is equal to 0 at some points, we can call such an equation *degenerate parabolic.*

The most popular example of a parabolic equation is the *heat equation*:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\,\frac{\partial^2 u}{\partial x^2}. \tag{12.23}$$

Here the coefficient $a \equiv 1$ is everywhere *positive.* This equation describes heat transfer in a uniform unbounded one-dimensional medium (a rod going from $-\infty$ to ∞).

We are accustomed to considering for ordinary differential equations *initial-value problems.* (For stochastic differential equations we also considered initial-value problems, prescribing the value of the unknown stochastic process X_t at a time point t_0; but in contrast with the situation of ordinary differential equations, we cannot solve the equation "backwards": only from the initial point t_0 to values of $t > t_0$).

The solution of a parabolic equation (12.21) is, of course, not unique, just as solutions of ordinary or stochastic differential equations without any initial conditions prescribed. So what *problems* can be solved for parabolic equations? Can we prescribe *initial conditions* at some $t = t_0$?

It turns out that for equations (12.21) with $\alpha(t,x) > 0$ one can consider an initial-value problem:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = \dfrac{\alpha(t,x)}{2}\,\dfrac{\partial^2 u}{\partial^2 x}(t,x) + \beta(t,x)\,\dfrac{\partial u}{\partial x}(t,x) + \gamma(t,x)u(t,x) + f(t,x), \\ u(t_0,\,x) = \varphi(x), \qquad x \in \mathbb{R}, \end{cases} \tag{12.24}$$

the solution being sought for $t \geq t_0$.

In contrast with this, if $\alpha(t,x) < 0$, a solution that we are looking for is defined for the times t that are *smaller* than the time at which the value of the unknown function is given. We could call it, instead of *initial*, a *final* condition. Let us write a final-value problem with the final value prescribed at the time point T:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = \mathcal{L}_t u(t,x) + f(t,x), \\ u(T,x) = \varphi(x). \end{cases} \tag{12.25}$$

Problems (12.24), (12.25) are called *Cauchy problems* for parabolic equations.

In particular, a Cauchy problem of the type of (12.25) should be set for the equation (12.18):

$$\begin{cases} \dfrac{\partial u}{\partial t} = -L_t u(t,x) + g(t,x), \quad t \leq T,\ x \in \mathbb{R}, \\ u(T,x) = \varphi(x), \qquad x \in \mathbb{R}; \end{cases} \tag{12.26}$$

and (12.19) provides a formula for the value of the solution of this problem at a point (t_0, x_0), $t_0 \le T$:

$$u(t_0, x_0) = E\Big(\varphi(X_T) - \int_{t_0}^{T} g(t, X_t)\ dt\Big), \tag{12.27}$$

where X_t is the diffusion process governed by the stochastic equation $dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$, starting at time t_0 at the point x_0.

For the heat equation (12.23), we should solve the Cauchy problem:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t, x) = \dfrac{1}{2}\,\dfrac{\partial^2 u}{\partial x^2}(t, x), & t \ge t_0,\ \ x \in (-\infty, \infty), \\ u(t_0,\, x) = \varphi(x), & x \in (-\infty, \infty). \end{cases} \tag{12.28}$$

Of course, equations with $\alpha(t, x)$ being positive can be made into one with the corresponding coefficient negative (and the opposite can be done) by simple time substitution: taking $-t$, or $C - t$ instead of t. The sign of the time derivative changes to the opposite, the "space" derivatives remain as they were, an initial-value problem turns into a final-value one, and vice versa.

Other problems also are considered for parabolic differential equations: boundary problems, initial-boundary-value problems, etc.; but the Cauchy problem is enough for us right now.

More on parabolic equations and diffusion processes in the next lecture note.