MATH 657 STOCHASTIC DIFFERENTIAL EQUATIONS SPRING 2007

Lecture note 11. Diffusion processes.

From now on, I am dropping my attempts to number my lecture notes so that it corresponds as precisely as possible to the numbers of the lectures. So Lecture note 10 is probably corresponds to something like Lecture 12.

In Theorem 10.1, I gave a more general formulation of conditions under which Itô's formula holds; and I told you, in particular, that a stochastic integral $\int_a^b g(t,\omega)\,dW_t$ is defined if the integrand $g(t,\omega)$ is determined by the past of the Wiener process and

$$\int_a^b E\big(g(t,\omega)^2\big)\,dt < \infty \tag{11.1}$$

(this integral is, by the way, equal to $E\Big(\int_a^b g(t,\omega)^2\,dt\Big)$). In serious books on stochastic integrals, the existence of the stochastic integral is proved under a weaker condition: only that

$$\int_a^b g(t,\omega)^2\,dt < \infty \qquad \text{almost surely.} \tag{11.2}$$

We are not going to consider stochastic integrals of random functions satisfying only the condition (11.2), because, while under the condition (11.1) the stochastic integral has zero expectation:

$$E\Big(\int_a^b g(t,\omega)\,dW_t\Big) = 0 \tag{11.3}$$

(because the expectation of the limit is equal to the limit of the expectation if the limit of random variables is understood as a mean-square limit), under the condition (11.2) it is no longer necessarily so.

So in all stochastic integrals to come up in this course, the condition (11.1) will be always assumed.

In the previous lectures, I supposed that a function $F(t,x)$, $t \ge t_0$, has a continuous partial derivative $\dfrac{\partial F}{\partial t}(t,x)$, $t \ge t_0$, $x \in \mathbb{R}$. But a function defined only for $t \ge t_0$ cannot have a (partial) derivative at $t = t_0$: it can have only a one-sided derivative. So what I meant was that the partial derivative exists for $t > t_0$, at $t = t_0$ a one-sided derivative exists, and the function defined as the derivative for $t > t_0$, and as the right-hand derivative for $t = t_0$ is continuous for $t \ge t_0$, $x \in \mathbb{R}$. It can be expressed also thus: for every $x \in \mathbb{R}$ there exist finite limits

$$\lim_{t \to t_0^+,\, y \to x} \frac{\partial F}{\partial t}(t,y) \tag{11.4}$$

(and we can use the notation $\dfrac{\partial F}{\partial t}(t_0,x)$ for this limit). This is enough for what we used about the partial derivative (the formula $F(t',y) - F(t,y) = \dfrac{\partial F}{\partial t}(t^*,y)\cdot(t'-t)$, where t^* is between t and t').

For the time being, I am postponing writing the Itô formula for multidimensional processes with stochastic differentials.

If a stochastic process is a solution of a stochastic differential equation

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t, \qquad t \geq t_0, \tag{11.5}$$

we call it a *diffusion process*. This name 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 a *drift* involved, accounting for the summand $b(t, X_t)\, dt$, and the rate of the chaotic movements similar to Brownian motion may depend on time t and the point X_t of the space region where the particle is at that time.

We have proved the existence of a diffusion process starting at time t_0 from a nonrandom point x_0 under some conditions imposed on the coefficients $b(t, x)$, $\sigma(t, x)$ (continuity in the pair (t, x) and a Lipschitz condition in the second argument). This does not mean that a solution does not exist or is not unique if the coefficients don't satisfy, say, the Lipschitz condition. Even if the solution exists, but is *not* unique, we'll still call X_t a diffusion process.

Diffusion processes provide mathematical models for many problems in which stochastic processes are applied. Let me show two such applications.

One is to physics.

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 V_t. If $V_t > 0$, 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 $V_t < 0$, 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:

$$F = -\mu\, V_t. \tag{11.6}$$

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

$$\frac{dV_t}{dt} = -\frac{\mu}{m}\, V_t, \tag{11.7}$$

where m is the mass of the particle (in the lecture, I took this mass to be 1 in order not to have to write $-\dfrac{\mu}{m}\, V_t$). 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 introducing a stochastic term:

$$dV_t = -\frac{\mu}{m}\, V_t\, dt + \sigma\, dW_t. \tag{11.7}$$

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 (11.7); and it is a diffusion process. The coefficients $b(t, x)$, $\sigma(t, x)$ here don't depend on t: $b(t, x) = -\dfrac{\mu}{m} \cdot x$, $\sigma(t, x) \equiv \sigma = \text{const}$.

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

$$X_t = X_{t_0} + \int_{t_0}^{t} V_s\, ds. \tag{11.8}$$

We'll return to this model later.

A second example is an application to finance. Stock prices are subject to irregular oscillations. If we invest some amount of money in some stock, how this amount changes in time will be described by a stochastic process X_t. A reasonable, and pretty simple, mathematical model is a diffusion process described by a stochastic equation having the form (11.5). As for the coefficients $b(t, x)$, $\sigma(t, x)$, we can say this: if we have two dollars invested in a stock at some time, this amount grows or decreases at twice the rate it would had we invested only one dollar, and the random oscillations of the amount that is \$2 at some time are twice those of the amount that is just one dollar at this time. So the coefficients $b(t, x)$, $\sigma(t, x)$ should be taken proportional to the variable x. The simplest model, not taking into account seasonal or multi-year changes of these rates, is just that

$$b(t, x) = \mu \cdot x, \qquad \sigma(t, x) = \sigma \cdot x, \tag{11.9}$$

where μ and σ are some constants; so the stochastic equation for X_t is:

$$dX_t = \mu X_t\, dt + \sigma X_t\, dW_t. \tag{11.10}$$

Of course, this is just the simplest model. One of its weak points is that it is one-dimensional, taking into account the oscillations of one kind of stock only. A richer model can consider several different kinds of stock – but for this we need multidimensional diffusion processes that we don't consider at this time. Another weak point of this model is that it presupposes that the increments of the Wiener process W_t, driving the changes in X_t, over non-overlapping time intervals are independent. But it is believed that the oscillations in stock prices in non-overlapping time intervals are dependent; some believe that this dependence is "positive", some that it is "negative"; some think that there is some *long-time dependence*. So mathematicians hired by financial companies try to push forward each its own mathematical technique; the question remains for some time whether their approaches are profitable. Let us stick to the simples model (11.10).

Of course, this one-dimensional model hardly is what is really needed: the reasonable strategy in stock market is juggling several different kinds of stock; but one should start with the simplest case.

The coefficients in the equations (11.7), (11.10) are clearly continuous and satisfy a Lipschitz condition: so the solutions do exist (and are unique given the initial value at some time t_0). This is already good for a mathematical model.

Also these equations are *linear*: their right-hand sides depend on their unknown functions linearly.

It turns out that, in a rare exception, we can "solve" these 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 (11.7) means that

$$X_t = X_{t_0} - \frac{\mu}{m} \int_{t_0}^{t} X_s \, ds + \int_{t_0}^{t} \sigma \, dW_s. \tag{11.11}$$

I have mentioned that in general, a stochastic integral of a function $g(t, \omega)$ cannot be defined as a Stieltjes integral, and we cannot operate with it the same way as if W_t were a smooth function. *But* the exception is made for non-random functions $g(t, \omega) = g(t)$. In (11.11) the integrand in the stochastic integral is just a constant; so we may try to solve the equation (11.11) – which is equivalent to the stochastic differential equation (11.7) plus the initial condition at the time point t_0 just the way we would were the function W_t differentiable.

The solution X_t of the initial-value problem

$$\frac{dX_t}{dt} = -\frac{\mu}{m} X_t + \sigma \cdot \frac{dW_t}{dt}, \qquad X_{t_0} = x_0 \tag{11.12}$$

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

$$X_t = x_0 \cdot e^{-\frac{\mu}{m} t} + \sigma \int_{t_0}^{t} e^{-\frac{\mu}{m}(t-s)} \frac{dW_s}{ds} \, ds; \tag{11.13}$$

this can be rewritten as

$$X_t = x_0 \cdot e^{-\frac{\mu}{m} t} + \sigma \int_{t_0}^{t} e^{-\frac{\mu}{m}(t-s)} \, dW_s. \tag{11.14}$$

While formula (11.13) did not really make sense for the Wiener process, because it is not differentiable, formula (11.14) 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

$$Y_t = \int_{t_0}^{t} e^{\frac{\mu}{m}s}\, dW_s; \tag{11.15}$$

its stochastic differential is

$$dY_t = e^{\frac{\mu}{m}t}\, dW_t. \tag{11.16}$$

Taking the function

$$F(t,x) = x_0 \cdot e^{-\frac{\mu}{m}t} + \sigma e^{-\frac{\mu}{m}t} \cdot x, \tag{11.17}$$

we can write that

$$X_t = F(t, Y_t). \tag{11.18}$$

The rest is left to you as a problem (see the new ones in the list of problems).

In the stochastic integral equation corresponding to (11.10) the integrand in the stochastic integral is a true random function, not a function of the time t only, so it's not plausible that we should be able to handle the equation as if it were about differentiable functions. If we tried to do so, we'd obtain:

$$X_t = X_{t_0} \cdot \exp\{\mu \cdot (t - t_0) + \sigma \cdot (W_t - W_{t_0})\}. \tag{11.19}$$

It turns out that (11.19) is *not* true.

Let us apply Itô's formula to stochastically differentiate the right-hand side of (11.19). For simplicity's sake, let us assume that $W_{t_0} = 0$ (all our formulas, all our results contain only increments of the Wiener process, so this doesn't really matter; and if we didn't take $W_{t_0} = 0$, we would have to introduce a special notation for W_{t_0}, and drag it along in all formulas). Then the right-hand side of (11.19) can be written as $F(t, W_t)$, where $F(t,x) = x_0 e^{\mu(t-t_0)+\sigma x}$. Itô's formula yields:

$$\begin{aligned} dF(t,W_t) &= x_0\big[\mu e^{\mu(t-t_0)+\sigma W_t} + \frac{\sigma^2}{2}\, e^{\mu(t-t_0)+\sigma W_t}\big]\, dt + \sigma x_0 e^{\mu(t-t_0)+\sigma W_t}\, dW_t \\ &= \big(\mu + \frac{\sigma^2}{2}\big) F(t,W_t)\, dt + \sigma F(t,W_t)\, dW_t. \end{aligned} \tag{11.20}$$

So $F(t, W_t)$ does *not* satisfy equation (11.10).

It becomes clear that the solution has the form

$$X_t = x_0 \exp\big\{\big(\mu - \frac{\sigma^2}{2}\big)(t - t_0) + \sigma(W_t - W_{t_0})\big\} \tag{11.21}$$

(checked again using Itô's formula).

The exponent $\big(\mu - \frac{\sigma^2}{2}\big)(t - t_0) + \sigma(W_t - W_{t_0})$ has the normal distribution with parameters $\big(\mu_0(t - t_0),\, \sigma^2(t - t_0)\big)$, $\mu_0 = \mu - \sigma^2/2$; and the random variable X_t has what they call *lognormal* (logarithmic normal) distribution.

Let me show how this is applied to calculating prices of stock *options*.

At a time moment t_0 we can acquire a *buying option* allowing (but not forcing) us to buy a specified amount of stock at time $t_1 > t_0$ at the price C pr unit. A *selling* option is the right, acquired at time t_0, to sell a specified amount of stock at time t_1 at the price D.

What should be the fair price of these options per stock unit?

The fair price should be equal to the expectation of our gain from using our option (we don't speak now of the risk: that our gain with expectation a and variance b being preferable to the gain with the same expectation and a variance $b' > b$; this can be taken care of in more sophisticated versions of the theory).

Suppose the market price of a unit of stock at time t_1 is x. If $x < C$, the best use of our buying option is *not* to use it (we can buy cheaper at the market price); our gain from this, our best, use of our option is 0. For $x > C$, we buy a thing at a fixed price C, and can immediately resell it at the market price x, and we gain the amount $x - C$ per unit. So the gain function $G_{\mathrm{b}}(x)$ ($_{\mathrm{b}}$ for "buying") is given by

$$G_{\mathrm{b}}(x) = \begin{cases} 0, & x < C, \\ x - C, & x > C. \end{cases} \tag{11.22}$$

We have to find the expectation

$$E\big(G_{\mathrm{b}}(X_{t_1})\big) = E\big(G_{\mathrm{b}}\big(x_0 \exp\{\mu_0(t_1 - t_0) + \sigma W_{t_1}\}\big)\big) \tag{11.23}$$

(remember, we took $W_{t_0} \equiv 0$).

What is the value of the function $G_{\mathrm{b}}(x)$ at $x = C$? First of all, this is not important, because in our model the random variable X_{t_1} has a continuous distribution, and almost surely does not hit the single point C; but of course, our gain $G_{\mathrm{b}}(C)$ for the price at time t_1 being exactly equal to C is equal to 0, which can be expressed by adding the possibility of equality in either one, or in both, inequalities in (11.23).

The gain $G_{\mathrm{s}}(x)$ from using a *selling* option is given by

$$G_{\mathrm{s}}(x) = \begin{cases} D - x, & x \le D, \\ 0, & x \ge D. \end{cases} \tag{11.24}$$

Let us calculate the expectation (11.23). The random variable $Y = \mu_0(t_1 - t_0) + \sigma W_{t_1}$ has the normal distribution with parameters $(a, b) = \big(\mu_0(t_1 - t_0),\ b = \sigma^2(t_1 - t_0)\big)$, so its probability density is

$$p(y) = \frac{1}{\sqrt{2\pi b}}\, e^{-(y-a)^2/2b}. \tag{11.25}$$

We have:

$$E\big(G_{\mathrm{b}}(x_0 e^Y)\big) = \int\limits_{\ln(C/x_0)}^{\infty} (x_0\, e^y - C) \cdot \frac{1}{\sqrt{2\pi b}}\, e^{-(y-a)^2/2b}\, dy. \tag{11.26}$$

As always when handling a normal distribution, make the substitution $\dfrac{y-a}{\sqrt{b}} = u$, $y = a + \sqrt{b} \cdot u$, $dy = \sqrt{b}\, du$:

$$E\big(G_{\mathrm{b}}(x_0 e^Y)\big) = \int\limits_{(\ln(C/x_0)-a)/\sqrt{b}}^{\infty} (x_0\, e^{a+\sqrt{b}\cdot u} - C) \cdot \frac{1}{\sqrt{2\pi}}\, e^{-u^2/2}\, du. \tag{11.27}$$

The integral

$$\int\limits_{(\ln(C/x_0)-a)/\sqrt{b}}^{\infty} C\cdot\frac{1}{\sqrt{2\pi}}\,e^{-u^2/2}\,du = C\cdot\big[1-\Phi\big(\frac{\ln(C/x_0)-a}{\sqrt{b}}\big)\big], \tag{11.28}$$

where Φ is the Laplace function:

$$\Phi(z) = \int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}\,e^{-u^2/2}\,du. \tag{11.29}$$

As for the first term in the parentheses in (11.27), we have:

$$\begin{aligned} x_0 &\int\limits_{(\ln(C/x_0)-a)/\sqrt{b}}^{\infty} e^{a+\sqrt{b}\cdot u}\cdot\frac{1}{\sqrt{2\pi}}\,e^{-u^2/2}\,du \\ &= x_0\,e^{a+b/2}\int\limits_{(\ln(C/x_0)-a)/\sqrt{b}}^{\infty}\frac{1}{\sqrt{2\pi}}\,e^{-(u-\sqrt{b})^2/2}\,du \\ &= x_0\,e^{a+b/2}\int\limits_{(\ln(C/x_0)-a)/\sqrt{b}-\sqrt{b}}^{\infty}\frac{1}{\sqrt{2\pi}}\,e^{-v^2/2}\,dv \\ &= x_0\,e^{a+b/2}\big[1-\Phi\big(\frac{\ln(C/x_0)-a-b}{\sqrt{b}}\big)\big]. \end{aligned} \tag{11.30}$$

So our answer is:

$$\begin{aligned} E\big(G_{\mathrm{b}}(X_{t_1})\big) = x_0\,e^{(\mu_0+\sigma^2/2)(t_1-t_0)}&\big[1-\Phi\big(\frac{\ln(C/x_0)-(\mu_0+\sigma^2)(t_1-t_0)}{\sigma\sqrt{t_1-t_0}}\big)\big] \\ -C\cdot&\big[1-\Phi\big(\frac{\ln(C/x_0)-\mu_0(t_1-t_0)}{\sigma\sqrt{t_1-t_0}}\big)\big]. \end{aligned} \tag{11.31}$$

In this long expression (that must contain some mistakes?) some things have a nice meaning, e. g., $\mu_0+\sigma^2/2=\mu$; but some other don't – say, $\mu_0+\sigma^2$.

As a matter of fact, we don't really need any theory of stochastic processes or of stochastic differential equations to find this: the only thing we need is that the random variable $Y=\ln(X_{t_1}/x_0)$ has a normal distribution with parameters proportional to t_1-t_0; and that this is true, at least approximately, follows from our assumption that the percentages of growth or decrease of stock price are independent in non-overlapping time intervals (so the random variable Y is a sum of a large number of independent random summands, to which the normal approximation – the Central Limit Theorem – can be applied). But the theory of stochastic processes *is* needed if we come to options that can be exercised *at any time up to* t_1 (I think, this kind is what is called *European* options).

And here we come to our next topic: Diffusion processes and partial differential equations.