Lecture note 44. Generalization of Itô's formula.

In the previous lecture I said that the function defined by formula (43.26) is continuous; and this is true, but needs some proving. I could have left it to you, but here it is: Suppose that $\delta > 0$ is chosen so that for z and z' in the same interval (x_{i-1}, x_i), or $(-\infty, x_1)$, or (x_m, ∞)

$$|z' - z| < \delta \quad \Rightarrow \quad |F''(z') - F''(z)| < \varepsilon. \tag{44.1}$$

Then for $0 < x' - x < \delta$ we have:

$$\begin{aligned} \Big| \int_l^r \varphi(y) \cdot F''(x' + y)\ dy - \int_l^r \varphi(y) \cdot F''(x + y)\ dy \Big| \\ \le \int\limits_{[l,\,r] \setminus [x_i - x,\, x_i - x']} |\varphi(y)| \cdot |F''(x' + y) - F''(x + y)|\ dy \\ + \int\limits_{[l,\,r] \setminus [x_i - x,\, x_i - x']} |\varphi(y)| \cdot |F''(x' + y) - F''(x + y)|\ dy \\ \le (r - l) \cdot \max_y |\varphi(y)| \cdot \varepsilon + \delta \cdot \max_y |\varphi(y)| \cdot 2 \sup_z |F''(z)| \end{aligned} \tag{44.2}$$

(here x_i is the (at most unique) discontinuity point of the function F'' in the interval $[l + x', r + x]$). The right-hand side of (44.2) can be made arbitrarily small by choosing δ small enough (we could have made it less than ε), which proves the continuity.

Now, we have for arbitrary real α and ξ:

$$F'(\xi) = F'(\alpha) + \int_\alpha^\xi F''(\zeta)\ d\zeta \tag{44.3}$$

(if there are no discontinuity points x_i of F'' between α and ξ, it's clear; if there is one, we have $F'(\xi) - F'(\alpha) = F'(x_i) - F'(\alpha) + F'(\xi) - F'(x_i) = \int_\alpha^{x_i} F''(\zeta)\ d\zeta + \int_{x_i}^\xi F''(\zeta)\ d\zeta$; same if there are two, etc.). We want to prove that

$$F'_\varphi(x) = F'_\varphi(a) + \int_a^x \Big[\int_l^r \varphi(y) \cdot F''(z + y)\ dy \Big]\ dz. \tag{44.4}$$

After using formula (43.25) and changing the order of integration, the equality to be proved turns into

$$\int_l^r \varphi(y) \cdot F'(x + y)\ dy = \int_l^r \varphi(y) \cdot F'(a + y)\ dy + \int_l^r \varphi(y) \cdot \Big[\int_a^x F''(z + y)\ dz \Big]\ dy. \tag{44.5}$$

If we take, in formula (44.3), $\alpha = a + y$, $\xi = x + y$, we get:

$$F'(x + y) = F'(a + y) + \int_{a+y}^{x+y} F''(\zeta)\ d\zeta = F'(a + y) + \int_a^x F''(z + y)\ dz, \tag{44.6}$$

from which (44.5) immediately follows. (By the way, this is one of the ways to prove Microtheorem 43.1; and I had no intention of proving it...)

Now let us remember that we are in the process of proving Theorem 43.1. We take $l = 0$, $r = \varepsilon > 0$, $\varphi(y) \equiv 1/\varepsilon$:

$$F_{+\varepsilon}(x) = \frac{1}{\varepsilon} \int_0^\varepsilon F(x + y)\ dy. \tag{44.7}$$

This function is twice continuously differentiable, so

$$dF_{+\varepsilon}(X_t) = F'_{+\varepsilon}(X_t) \cdot g(t, \omega)\, dW_t + \big[\frac{1}{2}\, F''_{+\varepsilon}(X_t) \cdot g(t, \omega)^2 + F'_{+\varepsilon}(X_t) \cdot f(t, \omega)\big]\, dt, \tag{44.8}$$

which means that for $t_0 \le t$

$$\begin{aligned} F_{+\varepsilon}(X_t) = F_{+\varepsilon}(X_{t_0}) + \int_{t_0}^t F'_{+\varepsilon}(X_s) \cdot g(s, \omega)\, dW_s \\ + \int_{t_0}^t \big[\frac{1}{2}\, F''_{+\varepsilon}(X_s) \cdot g(s, \omega)^2 + F'_{+\varepsilon}(X_s) \cdot f(s, \omega)\big]\, ds. \end{aligned} \tag{44.9}$$

As $\varepsilon \to 0$, we have $F_{+\varepsilon}(x) \to F(x)$,

$$F'_{+\varepsilon}(x) = \frac{1}{\varepsilon} \int_0^\varepsilon F'(x + y)\ dy \to F'(x), \tag{44.10}$$

uniformly in x, but

$$F''_{+\varepsilon}(x) = \frac{1}{\varepsilon} \int_0^\varepsilon F''(x + y)\ dy \to F''(x^+) \qquad (= F''(x) \text{ for } x \neq x_1, ..., x_m), \tag{44.11}$$

and this not uniformly in x.

Of course, $F_{+\varepsilon}(X_t) \to F(X_t)$, $F_{+\varepsilon}(X_{t_0}) \to F(X_{t_0})$; but what about *the integrals* in (44.9)?

We have:

$$\begin{aligned} E\Big(\Big(\int_{t_0}^t F'_{+\varepsilon}(X_s) \cdot g(s, \omega)\ dW_s - \int_{t_0}^t F'(X_s) \cdot g(s, \omega)\ dW_s\Big)^2\Big) \\ = \int_{t_0}^t E\big(\big(F'_{+\varepsilon}(X_s) - F'(X_s)\big)^2 \cdot g(s, \omega)^2\big)\ ds \\ \le \max_x \big(F'_{+\varepsilon}(x) - F'(x)\big)^2 \cdot \sup\{g(s, \omega)^2 \colon\ X_s(\omega) \in [A - 1,\ B]\} \to 0 \end{aligned} \tag{44.12}$$

as $\varepsilon \to 0^+$, where $[A,\ B]$ is the interval outside which $F(x) \equiv 0$. Similarly,

$$\begin{aligned} \Big\| \int_{t_0}^t F'_{+\varepsilon}(X_s) \cdot f(s, \omega)\ ds - \int_{t_0}^t F'(X_s) \cdot f(s, \omega)\ ds \Big\|_2 \\ \le \int_{t_0}^t \|F'_{+\varepsilon}(X_s) \cdot f(s, \omega) - F'(X_s) \cdot f(s, \omega)\|_2\, ds \to 0. \end{aligned} \tag{44.13}$$

But to get that

$$\begin{aligned} \Big\| \int_{t_0}^{t} F''_{+\varepsilon}(X_s) \cdot g(s,\omega)^2 \, ds - \int_{t_0}^{t} F''(X_s) \cdot g(s,\omega)^2 \, ds \Big\|_2 \\ \le \int_{t_0}^{t} \| F''_{+\varepsilon}(X_s) \cdot g(s,\omega)^2 - F''(X_s) \cdot g(s,\omega)^2 \|_2 \, ds \to 0, \end{aligned} \tag{44.14}$$

where $F''(x)$ is replaced with its right-hand limit for $x = x_i$, we need to use the dominated-convergence theorem, and not only in the formulation given in Lecture note 2 (Theorem 2.2), but also in the following one:

Theorem 2.2′. *If* $f_n(x)$ *is a sequence of functions*, $\lim_{n\to\infty} f_n(x) = f(x)$, *and all* $|f_n(x)| \le g(x)$, $\int_a^b g(x)\, dx < \infty$, *then* $\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b f(x)\, dx$.

From the point of view of theory of measure and integration, Theorems 2.2 and 2.2′ are particular cases of one general dominated-convergence theorem.

We have, for every $s \in [t_0, t]$ and every $\omega \in \Omega$:

$$\begin{aligned} & [F''_{+\varepsilon}(X_s) \cdot g(s,\omega)^2 - F''(X_s) \cdot g(s,\omega)^2]^2 \to 0, \\ & [F''_{+\varepsilon}(X_s) \cdot g(s,\omega)^2 - F''(X_s) \cdot g(s,\omega)^2]^2 \\ & \quad \le 4 \max_x F''(x)^2 \cdot \sup\{g(s,\omega)^4 \colon X_s(\omega) \in [A-1,\, B]\}, \end{aligned} \tag{44.15}$$

so by Theorem 2.2 for every s

$$\begin{aligned} E\big([F''_{+\varepsilon}(X_s) \cdot g(s,\omega)^2 - F''(X_s) \cdot g(s,\omega)^2]^2\big) \to 0, \\ \|F''_{+\varepsilon}(X_s) \cdot g(s,\omega)^2 - F''(X_s) \cdot g(s,\omega)^2\|_2 \to 0; \end{aligned} \tag{44.16}$$

these functions for every $\varepsilon \in (0,1]$ are dominated by the constant $2\max_x |F''(x)| \times$ $\times \sup\{g(s,\omega)^2 \colon X_s(\omega) \in [A-1,\, B]\}$, and by Theorem 2.2′ we have (44.14).

So passing to the mean-square limits, we get from formula (44.9) that almost surely holds the equality

$$\begin{aligned} F(X_t) = F(X_{t_0}) + \int_{t_0}^{t} F'(X_s) \cdot g(s,\omega)\, dW_s \\ + \int_{t_0}^{t} \big[\frac{1}{2} F''(X_s) \cdot g(s,\omega)^2 + F'(X_s) \cdot f(s,\omega)\big]\, ds, \end{aligned} \tag{44.17}$$

where $F''(x_i)$ are replaced with $F''(x_i^+)$.

If we take

$$F_{-\varepsilon}(x) = \int_{-\varepsilon}^{0} \frac{1}{\varepsilon} \cdot F(x+y)\, dy, \tag{44.18}$$

we obtain in the same way that we can replace $F''(x_i)$ in equality (44.17) with $F''(x_i^-)$.

I claimed also that we can replace it with an arbitrary number; what about this? Take the function $F(x)$ given by

$$F(x) = \begin{cases} 0, & x \le x_i, \\ (x - x_i)^2, & x_i \le x \le x_i + 1, \\ \text{smooth} & \text{for all } x > x_i, \\ 0, & x \ge x_i + 2 \end{cases} \tag{44.19}$$

(I could have written a concrete formula for x between $x_i + 1$ and $x_i + 2$, but what we really need is that it is possible to define such a function). The jump $F''(x_i^+) - F''(x_i^-)$ of the second derivative at the point x_i is equal to 2. The formula (44.17) holds both with $F''(x_i)$ replaced with $F''(x_i^+)$ and with $F''(x_i^-)$; so we have almost surely:

$$\int_{t_0}^{t} I_{\{x_i\}}(X_s) \cdot g(s, \omega)^2 \, ds = 0. \tag{44.20}$$

So whatever value we take instead of $F''(X_s)$ for $X_s = x_i$, it does not change the integral (almost surely).

By the way, if $g(s, \omega) \neq 0$ for $X_s(\omega) = x_i$, we get that the time spent by our process at one point x_i is almost surely equal to 0. In particular, this is the situation for every diffusion process: the process spends zero time at the points at which the diffusion coefficient $a(x) = \sigma(x)^2 > 0$, and can spend some positive time only at points at which the diffusion coefficient is equal to 0.

There are modifications of Theorem 43.1 to functions $F(t, x)$ depending on both time t and the spatial variable x:

Theorem 43.1′. *Let X_t be a one-dimensional stochastic process having a stochastic differential $dX_t = f(t, \omega)\, dt + g(t, \omega)\, dW_t$. Let $F(t, x)$ be a function that is continuous and once continuously differentiable in x; let its second partial derivative in x exist except on finitely many smooth curves $x = x_1(t)$, ..., $x = x_m(t)$, and let the time derivative $\frac{\partial F}{\partial t}(t, x)$ also exist except on finitely many smooth curves $x = x_i(t)$, and maybe also at finitely many separate values $t = t_j$. Suppose that for (t, x) on these lines one-sided limits as $(s, z) \to (t, x)$ of $\frac{\partial^2 F}{\partial x^2}(s, z)$, $\frac{\partial F}{\partial t}(s, z)$ exist. Suppose that the function $F(t, x)$ is equal to 0 outside some bounded set, and that for every finite interval $[A, B]$ there exists a constant such that the random functions $f(t, \omega)$, $g(t, \omega)$ remain bounded by this constant for t for which $X_t(\omega) \in [A, B]$.*

Then

$$\begin{aligned} dF(t, X_t) = {} & \frac{\partial F}{\partial x}(t, X_t) \cdot g(t, \omega)\, dW_t \\ & + \big[\frac{\partial F}{\partial x}(t, X_t) + \frac{1}{2} \frac{\partial^2 F}{\partial^2 x}(t, X_t) \cdot g(t, \omega)^2 + \frac{\partial F}{\partial x}(t, X_t) \cdot f(t, \omega)\big]\, dt, \end{aligned} \tag{44.21}$$

where the non-existent partial derivatives are replaced with arbitrary numbers.

There exist also generalizations to the case of multidimensional space.

Applying Theorem 43.1, we get a new optimal-stopping theorem instead of Theorem 42.1 (but we'll formulate it for the one-dimensional case since such was Theorem 43.1):

Theorem 42.1′. *Suppose* $\tau^x_{(a,\,b)}$ *is almost surely finite. Suppose a function* $v(x)$, *continuously differentiable everywhere in* (a, b), *and twice continuously differentiable except at finitely many points, satisfies the following equalities and inequalities*:

$$\begin{aligned} v(x) &\geq f(x), &&x \in (a, b), \\ v(x) &= f(x), &&x = a \ \textit{or} \ b, \\ Lv(x) &= 0 &&\textit{for } x \in (a, b) \textit{ such that } v(x) > f(x), \\ Lv(x) &\leq 0 &&\textit{for } x \in (a, b) \textit{ such that } v(x) = f(x); \end{aligned} \tag{44.22}$$

suppose the function $v(x)$ *can be extended smoothly* (with exception of the points mentioned above) *to the whole real line.*

Then the optimal control consists in choosing τ *as the first time*

$$\hat{\tau}^x = \min\{t \geq 0\colon\ X^x_t \in A\} \tag{44.23}$$

at which the process reaches the closed set

$$A = \{x \in [a, b]\colon\ v(x) = f(x)\}; \tag{44.24}$$

and $v(x)$ *is our expected gain under the optimal control* (optimal stopping):

$$v(x) = E\big(\mathrm{Gain}^x(\hat{\tau}^x)\big) = \max_{\text{all stopping times } \tau \leq \tau^x_{(a,\,b)}} E\big(\mathrm{Gain}^x(\tau)\big). \tag{44.25}$$

This Theorem *can be* applied to Example 42.1, and the optimal stopping strategy is that found in Lecture note 43.