Lecture note 14. Multidimensional Itô's formula. More on diffusion processes and partial differential equations.

Let $\boldsymbol{X}_t = (X_t^1, ..., X_t^r)$ be an r-dimensional random function having a stochastic differential:

$$dX_t^i = f_i(t, \omega)\, dt + \sum_{k=1}^{n} g_{ik}(t, \omega)\, dW_t^k. \tag{14.1}$$

Let $F(t, \boldsymbol{x}) = F(t, x^1, ..., x^r)$ be a function that is once continuously differentiable in t and twice in the space variables x^1, ..., x^r. How to write the formula for the stochastic differential of the (one-dimensional) random function $F(t, \boldsymbol{X}_t)$?

The preliminary formula should be

$$dF(t, \boldsymbol{X}_t) = \frac{\partial F}{\partial t}(t, \boldsymbol{X}_t)\, dt + \sum_{i=1}^{r} \frac{\partial F}{\partial x^i}(t, \boldsymbol{X}_t)\, dX_t^i + \frac{1}{2} \sum_{i,\, j=1}^{r} \frac{\partial F}{\partial x^i}(t, \boldsymbol{X}_t)\, dX_t^i dX_t^j. \tag{14.2}$$

Then we replace dX_t^i with $f_i(t, \omega)\, dt + \sum_{k=1}^{n} g_{ik}(t, \omega)\, dW_t^k$, and delete all terms with dt^2 and $dt dW_t^k$, and we should replace $(dW_t^k)^2$ with dt; but what should we replace $dW_t^k dW_t^l$, $k \neq l$, with?

Since by problem **6** (see its solution) $\underset{\max(t_i - t_{i-1}) \to 0}{\text{l.i.m.}} (W_{t_i}^k - W_{t_{i-1}}^k)(W_{t_i}^l - W_{t_{i-1}}^l) = 0$, we should replace $dW_t^k dW_t^l$ with 0. So we have:

$$\begin{aligned} dF(t, \boldsymbol{X}_t) = \Big[&\frac{\partial F}{\partial t}(t, \boldsymbol{X}_t) + \sum_{i=1}^{r} \frac{\partial F}{\partial x^i}(t, \boldsymbol{X}_t) \cdot f_i(t, \omega) \\ &+ \sum_{i,\, j=1}^{r} \frac{1}{2} \frac{\partial^2 F}{\partial x^i \partial x^j}(t, \boldsymbol{X}_t) \cdot \sum_{k=1}^{n} g_{ik}(t, \omega) g_{jk}(t, \omega)\Big]\, dt \\ &+ \sum_{i=1}^{r} \sum_{k=1}^{n} \frac{\partial^2 F}{\partial x^i}(t, \boldsymbol{X}_t) \cdot g_{ik}(t, \omega)\, dW_t^k; \end{aligned} \tag{14.3}$$

or, in the integral form: almost surely

$$\begin{aligned} F(t, \boldsymbol{X}_t) = F(t_0, \boldsymbol{X}_{t_0}) + \int_{t_0}^{t} \Big[&\frac{\partial F}{\partial t}(s, \boldsymbol{X}_s) + \sum_{i=1}^{r} \frac{\partial F}{\partial x^i}(s, \boldsymbol{X}_s) \cdot f_i(s, \omega) \\ &+ \sum_{i,\, j=1}^{r} \frac{1}{2} \frac{\partial^2 F}{\partial x^i \partial x^j}(s, \boldsymbol{X}_s) \cdot \sum_{k=1}^{n} g_{ik}(s, \omega)\, g_{jk}(s, \omega)\Big]\, ds \\ &+ \sum_{i=1}^{r} \sum_{k=1}^{n} \int_{t_0}^{t} \frac{\partial^2 F}{\partial x^i}(s, \boldsymbol{X}_s) \cdot g_{ik}(s, \omega)\, dW_s^k. \end{aligned} \tag{14.4}$$

I am not stopping to prove this formula.

For a multidimensional diffusion process $\boldsymbol{X}_t$, satisfying stochastic equations

$$dX_t^i = b_i(t,\omega)\,dt + \sum_{k=1}^{n} \sigma_{ik}(t,\omega)\,dW_t^k, \tag{14.5}$$

we have:

$$\begin{aligned} dF(t,\boldsymbol{X}_t) = \big[\frac{\partial F}{\partial t}(t,\boldsymbol{X}_t) &+ \sum_{i=1}^{r} b_i(t,\boldsymbol{X}_t)\cdot\frac{\partial F}{\partial x^i}(t,\boldsymbol{X}_t) \\ &+ \sum_{i,\,j=1}^{r} \frac{a_{ij}(t,\boldsymbol{X}_t)}{2}\,\frac{\partial^2 F}{\partial x^i\partial x^j}(t,\boldsymbol{X}_t)\big]\,dt \\ &+ \sum_{i=1}^{r}\sum_{k=1}^{n} \sigma_{ik}(t,\boldsymbol{X}_t)\,\frac{\partial^2 F}{\partial x^i}(t,\boldsymbol{X}_t)dW_t^k, \end{aligned} \tag{14.6}$$

where

$$a_{ij}(t,\boldsymbol{x}) = \sum_{k=1}^{n} \sigma_{ik}(t,\boldsymbol{x})\,\sigma_{jk}(t,\boldsymbol{x}). \tag{14.7}$$

In the matrix form, we can write, denoting

$$\sigma(t,\boldsymbol{x}) = \big(\sigma_{ik}(t,\boldsymbol{x})\big)_{\substack{1\le i\le r,\\ 1\le k\le n}}, \qquad a(t,\boldsymbol{x}) = \big(a_{ij}(t,\boldsymbol{x})\big)_{1\le i,\,j\le r}: \tag{14.8}$$

$$a(t,\boldsymbol{x}) = \sigma(t,\boldsymbol{x})\cdot\sigma(t,\boldsymbol{x})^T \tag{14.9}$$

(the superscript ^{T} denoting the transposed matrix). The matrix $a(t,\boldsymbol{x})$ is clearly non-negative definite.

So the r-dimensional differential operator L_t associated with an r-dimensional diffusion process is given by

$$L_t f(\boldsymbol{x}) = \sum_{i=1}^{r} b_i(t,\boldsymbol{x})\cdot\frac{\partial f}{\partial x^i}(\boldsymbol{x}) + \frac{1}{2}\sum_{i,\,j=1}^{r} a_{ij}(t,\boldsymbol{x})\cdot\frac{\partial^2 f}{\partial x^i\partial x^j}(\boldsymbol{x}). \tag{14.10}$$

The coefficients $b_i(t,\boldsymbol{x})$ are called *the drift coefficients*, the vector $\boldsymbol{b}(t,\boldsymbol{x}) = \big(b_1(t,\boldsymbol{x}),\ldots, b_r(t,\boldsymbol{x})\big)$, the *drift vector*; the matrix $a(t,\boldsymbol{x})$ *the diffusion matrix*. It is natural (but not necessary) to look for a matrix $\sigma(t,\,\boldsymbol{x})$ satisfying (14.9) in the class of square $r\times r$ matrices.

Examples: The r-dimensional Wiener process is associated with one-half the *Laplace operator* Δ:

$$L_t f(\boldsymbol{x}) = \frac{1}{2}\,\Delta f(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{r}\frac{\partial^2 f}{(\partial x^i)^2}: \tag{14.11}$$

the drift is equal to 0, and the diffusion matrix is the identity matrix I.

The process (X_t, V_t) of the position and velocity of the "physical Brownian motion" (see Lecture note 11) satisfies the system of stochastic equations

$$\begin{aligned} dX_t &= V_t\,dt, \\ dV_t &= -\frac{\mu}{m} V_t\,dt + \sigma\,dW_t. \end{aligned} \tag{14.12}$$

Here the dimension r of the stochastic process (X_t, V_t) is 2, and the dimension n of the Wiener process in the stochastic equation is 1; the drift is

$$\boldsymbol{b}(t, x, v) = \begin{pmatrix} v \\ -(\mu/m)v \end{pmatrix}, \tag{14.13}$$

$$\sigma(t, x, v) = \begin{pmatrix} 0 \\ \sigma \end{pmatrix}, \tag{14.14}$$

the diffusion matrix

$$a(t,\, x,\, v) = \sigma(t,\, x,\, v) \cdot \sigma(t,\, x,\, v)^T = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2 \end{pmatrix} \tag{14.15}$$

is a degenerate one, not *positive* definite, only nonnegative; the corresponding differential operator $L_t \equiv L$ is given by

$$Lf(x, v) = v \cdot \frac{\partial f}{\partial x} - \frac{\mu}{m} v \cdot \frac{\partial f}{\partial v} + \frac{\sigma^2}{2} \cdot \frac{\partial^2 f}{\partial v^2}. \tag{14.16}$$

A linear partial differential equation

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i,\,j=1}^{r} \alpha_{ij}(t, \boldsymbol{x}) \cdot \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{i=1}^{r} \beta_i(t, x) \cdot \frac{\partial u}{\partial x^i} + \gamma(t, \boldsymbol{x}) \cdot u + f(t, \boldsymbol{x}) \tag{14.17}$$

is called parabolic if the matrix $\big(\alpha_{ij}(t, \boldsymbol{x})\big)$ is everywhere positive definite:

$$\sum_{i,\,j=1}^{r} \alpha_{ij}(t, \boldsymbol{x}) \cdot \lambda_i \lambda_j > 0 \qquad \text{for any vector } \boldsymbol{\lambda} \neq \mathbf{0}, \tag{14.18}$$

or everywhere negative definite:

$$\sum \alpha_{ij}(t, \boldsymbol{x}) \cdot \lambda_i \lambda_j < 0; \tag{14.19}$$

if this matrix is nonnegative definite:

$$\sum \alpha_{ij}(t, \boldsymbol{x}) \cdot \lambda_i \lambda_j \geq 0, \tag{14.20}$$

or non-positive definite:

$$\sum \alpha_{ij}(t, \boldsymbol{x}) \cdot \lambda_i \lambda_j \le 0, \tag{14.21}$$

we can call the equation *degenerate parabolic.*

In the cases (14.18), (14.20) we solve the Cauchy problem with an *initial* condition, up from the time at which this condition is prescribed; for (14.19), (14.21), with a *final* condition, *down* in time from the time moment T at which this condition is prescribed.

In the example above, the equation for the expectation

$$u(t, x, v) = E\Big(\varphi(X_T^{t,\,x,v},\, V_T^{t,\,x,v}) - \int_t^T g(s, X_s^{t,\,x,v}, V_s^{t,\,x,v})\, ds\Big) \tag{14.22}$$

is a degenerate parabolic one, and is to be solved *down* from the final condition $\varphi(x, v)$.

When we first spoke of parabolic equations, we wrote a formula for a linear second-order operator $\mathcal{L}_t$ with a term $\gamma(t, x) \cdot u(t, x)$; but we haven't spoken of how to solve such equations probabilistically. Let us do it now.

Solving parabolic equations of the form $\dfrac{\partial u}{\partial t}(t, x) + \dfrac{a(t, x)}{2}\,\dfrac{\partial^2 u}{\partial x^2}(t, x) + b(t, x) \cdot \dfrac{\partial u}{\partial x}(t, x) = g(t, x)$ (for simplicity's sake, we are considering the one-dimensional case) is done by applying Itô's formula to the random function $u(t, X_t)$, $t \ge t_0$. Let us look what we get if we apply the same formula to the random function $u(t, X_t) \cdot \exp\Big\{\int_{t_0}^t c(s, X_s)\, ds\Big\}$.

If we denote $Y_t = Y_t^{t_0,\,x_0} = \displaystyle\int_{t_0}^t c(s, X_s)\, ds$, we have:

$$\begin{aligned} dX_t &= b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t, \\ dY_t &= c(t, X_t)\, dt, \end{aligned} \tag{14.23}$$

and applying Itô's formula to the function $F(t, x, y) = u(t, x) \cdot e^y$ we get

$$\begin{aligned} & d[u(t, X_t^{t_0,\,x_0}) \cdot e^{Y_t^{t_0,\,x_0}}] \\ &\quad = \Big[\frac{\partial u}{\partial t}(t, X_t^{t_0,\,x_0}) \cdot e^{Y_t^{t_0,\,x_0}} + \frac{\partial u}{\partial x}(t, X_t^{t_0,\,x_0}) \cdot e^{Y_t^{t_0,\,x_0}} \cdot b(t, X_t^{t_0,\,x_0}) \\ &\qquad + u(t, X_t^{t_0,\,x_0}) \cdot e^{Y_t^{t_0,\,x_0}} \cdot c(t, X_t^{t_0,\,x_0}) \\ &\qquad + \frac{1}{2}\,\frac{\partial^2 u}{\partial x^2}(t, X_t^{t_0,\,x_0}) \cdot e^{Y_t^{t_0,\,x_0}} \cdot \sigma(t, X_t^{t_0,\,x_0})^2\Big]\, dt \\ &\qquad + \frac{\partial u}{\partial x}(t, X_t^{t_0,\,x_0}) \cdot e^{Y_t^{t_0,\,x_0}} \cdot \sigma(t, X_t^{t_0,\,x_0})\, dW_t. \end{aligned} \tag{14.24}$$

If we denote

$$\frac{\partial u}{\partial t}(t, x) + \frac{a(t, x)}{2}\,\frac{\partial^2 u}{\partial x^2}(t, x) + b(t, x)\,\frac{\partial u}{\partial x}(t, x) + c(t, x)\, u(t, x) = g(t, x), \quad u(T, x) = \varphi(x), \tag{14.25}$$

rewrite (14.24) in the integral form, and take into account that the expectation of a stochastic integral is zero (provided that certain expectations are finite), we get:

$$u(t_0, x_0) = E\Big(\exp\Big\{ \int_{t_0}^{T} c(t, X_t^{t_0, x_0})\ dt\Big\} \cdot \varphi(T, X_T^{t_0, x_0}) \\ - \int_{t_0}^{T} \exp\Big\{\int_{t_0}^{t} c(s, X_s^{t_0, x_0})\ ds\Big\} \cdot g(t, X_t^{t_0, x_0})\ dt\Big). \tag{14.26}$$

Here is the formula for solving the Cauchy problem (14.25); we can rewrite it for (t, x) instead of (t_0, x_0) (and the integration variable t changed to s and the integration variable in the integral of c to some other letter, say v).