STOCHASTIC CALCULUS

Kiyosi Itô
Research Institute for Mathematical Sciences
Kyoto University
Kyoto, Japan

## 1. Stochastic differentials.

Let $(\Omega, P)$ be a complete probability space and $\mathfrak{B}(P)$ denote the $\sigma$-algebra of all P-measurable sets. Let $\{\mathcal{F}_t : 0 \le t < \infty\}$ be a right continuous increasing family of sub-$\sigma$-algebras of $\mathfrak{B}(P)$ such that $\mathcal{F}_0$ includes all P-null sets, and let $\mathcal{Q}$ denote the family of all local quasi-martingales $(\mathcal{F}_t)$ with continuous sample functions. We denote by $\mathcal{M}$ and $\mathcal{A}$ respectively the local martingales $(\mathcal{F}_t)$ in $\mathcal{Q}$ and the locally bounded variation processes in $\mathcal{Q}$.

$\mathcal{Q}$ is a vector space over the reals $\mathbb{R}$ and both $\mathcal{M}$ and $\mathcal{A}$ are subspaces of $\mathcal{Q}$. $\mathcal{Q}$ is also a commutative ring over $\mathbb{R}$ and is a subring of $\mathcal{Q}$.

The differential $dX$ of $X \in \mathcal{Q}$ is defined to be the random interval function induced from X:

$$(dX)(I) = X(t) - X(s) \qquad \text{for} \qquad I = (s,t] .$$

Let $d\mathcal{Q}$, $d\mathcal{M}$ and $d\mathcal{A}$ denote

$$\{dX : X \in \mathcal{Q}\}, \qquad \{dM : M \in \mathcal{M}\} \text{ and } \qquad \{dA : A \in \mathcal{A}\}$$

respectively. $d\mathcal{Q}$ is a vector space over $\mathbb{R}$ and both $d\mathcal{M}$ and $d\mathcal{A}$ are subspaces of $d\mathcal{Q}$ and

$$d\mathcal{Q} = d\mathcal{M} \oplus d\mathcal{A}.$$

$d\mathcal{Q}$ is also a $\mathcal{Q}$-module with respect to the $\mathcal{Q}$-multiplication:

$$(Y dX)(I) = \int_I Y(t) dX(t) \qquad (\text{stochastic integral } [1],[2],[3])$$

and both $d\mathcal{M}$ and $d\mathcal{A}$ are submodules of $d\mathcal{Q}$. Introducing product in $d\mathcal{Q}$ by

$$dX\, dY = d(XY) - XdY - YdX,$$

we can regard $d\mathcal{Q}$ as a commutative ring over $\mathcal{Q}$. Then $d\mathcal{A}$ is a subring of $d\mathcal{Q}$. It is easy to check that

$$d\mathcal{Q}\, d\mathcal{Q} \subset d\mathcal{A} \quad \text{and} \quad d\mathcal{Q}\, d\mathcal{A} = = 0, \text{ so } \quad (d\mathcal{Q})^3 = 0.$$

The <u>stochastic chain rule</u> is formulated as follows: If $F \in C^2(\mathbb{R}^n)$ and $X^i \in Q$ $(i=1,2,\ldots,n)$, then $F = F(X^1 \ldots X^n) \in Q$ and

$$(C) \quad dF = \sum_{i=1}^{n} \partial_i F \, dX^i + \tfrac{1}{2} \sum_{i,j=1}^{n} \partial_i \partial_j F \, dX^i dX^j.$$

The <u>symmetric $Q$-multiplication</u> $Y \circ dX$ is defined by

$$Y \circ dX = YdX + \tfrac{1}{2} dY dX.$$

$dQ$ is also a $Q$-module with this multiplication. The stochastic integral $(Y \circ dX)(I)$ is often denoted by

$$S\text{-}\int_I Y(t) dX(t)$$

and is called the <u>symmetric stochastic integral</u> that was introduced by Fisk [1] and Stratonovich [4]. They proved that the above chain rule (C) can be formulated in the same form as in the ordinary differential calculus:

$$(C_s) \quad dF(X) = \sum_{i=1}^{n} \partial_i F \circ dX^i, \text{ where } F \in C^3(\mathbb{R}^n).$$

This can be derived immediately from (C) and is called the <u>symmetric stochastic chain rule</u>.

## <u>2. Stochastic differential equations.</u>

Let $B(t) = (B^1(t), B^2(t), \ldots, B^n(t))$ be an n-dimensional Brownian motion and $\tilde{\mathcal{F}}_t$ denote the $\sigma$-algebra generated by $B^i(s)$ $(i=1,2,\ldots m,$ $s \le t)$ and all P-null sets. Then each $B^i(t)$ $(i=1,2,\ldots,m)$ is a martingale relative to the family $\mathcal{F}_t = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$. Applying the results of Section 1 we have

$$dB^i \in d\mathcal{M}, \; dt \in d\mathcal{A}, \; dB^i dB^j = \delta_{ij} \, dt \; \text{ and } \; dB^i dt = 0.$$

Consider the following two stochastic differential equations on $X_\alpha$ $(\alpha = 1,2,\ldots,m)$:

$$(E) \quad dX^\alpha = a^\alpha(t,X) \, dt + \sum_{i=1}^{n} \sigma_i^\alpha(t,X) dB^i$$

and

$$(E_s) \quad dX^\alpha = \bar{a}^\alpha(t,X) \circ dt + \sum_{i=1}^{n} \bar{\sigma}_i^\alpha(t,X) \circ dB^i,$$

where $a^\alpha(t,x)$, $\sigma_i^\alpha(t,x)$, $\bar{a}^\alpha(t,x)$ and $\bar{\sigma}_i^\alpha(t,x)$ belong to class $C^3$.

From now on the summation sign $\sum$ will be omitted as is usual in differential geometry.

Theorem 3.1. The equation (E) can be written in the form $(E_s)$ with the coefficients

(1) $\bar{a}^\alpha = a^\alpha - \tfrac{1}{2}(\partial_\beta \sigma_i^\alpha)\sigma_i^\beta$ and $\bar{\sigma}_i^\alpha = \sigma_i^\alpha$,

and the equation $(E_s)$ can be written in the form (E) with the coefficients

(i) $a^\alpha = \bar{a}^\alpha + \tfrac{1}{2}(\partial_\beta \bar{\sigma}_i^\alpha)\bar{\sigma}_i^\beta$ and $\sigma_i^\alpha = \bar{\sigma}_i^\alpha$.

Proof. Using the results of Section 1 we have

$$\begin{aligned}
& a^\alpha dt + \sigma_i^\alpha dB^i \\
&= a^\alpha dt + \sigma_i^\alpha \circ dB^i - \tfrac{1}{2} d\sigma_i^\alpha \, dB^i \\
&= a^\alpha dt + \sigma_i^\alpha \circ dB^i - \tfrac{1}{2}(\partial_\beta \sigma_i^\alpha \, dX^\beta)\, dB^i \\
&= a^\alpha dt + \sigma_i^\alpha \circ dB^i - \tfrac{1}{2} \partial_\beta \sigma_i^\alpha (\sigma_j^\beta dB^j)\, dB^i \\
&= (a^\alpha - \tfrac{1}{2}\sigma_j^\beta \partial_\beta \sigma_i^\alpha)\, dt + \sigma_i^\alpha \circ dB^i \\
&= \bar{a}^\alpha \circ dt + \bar{\sigma}_i^\alpha \circ dB^i,
\end{aligned}$$

proving the first assertion. Similarly for the second.

Theorem 3.2. The equation (E) determines a diffusion process with generator

$$\mathcal{G} = a^\alpha \partial_\alpha + \tfrac{1}{2}\sigma_i^\alpha \sigma_i^\beta \partial_\alpha \partial_\beta$$

and the equation $(E_s)$ determines a diffusion process with generator

$$\mathcal{G}_s = \bar{a}^\alpha \partial_\alpha + \tfrac{1}{2}(\bar{\sigma}_i^\alpha \partial_\alpha)(\bar{\sigma}_i^\beta \partial_\beta).$$

Proof. Let $f \in C^2(\mathbb{R}^m)$. Using the same argument as in the proof of the last theorem we have

(3) $$\begin{aligned} df(x) &= \partial_\alpha f \, dX^\alpha + \tfrac{1}{2}\partial_\alpha \partial_\beta f \, dX^\alpha dX^\beta \\ &= (a^\alpha \partial_\alpha f + \tfrac{1}{2}\sigma_i^\alpha \sigma_i^\beta \partial_\alpha \partial_\beta f)\, dt + \sigma_i^\alpha \partial_\alpha f \, dB^i. \end{aligned}$$

Since the last term $\sigma_i^\alpha \partial_\alpha f \, dB^i$ belongs to $d\mathcal{M}$, the first assertion follows at once. Similarly for the second.

An interesting example of $(E_s)$ is the Stroock equation [8] determining a spherical Brownian motion:

(S) $dX^\alpha = \sum_{i=1}^{n} (\delta_{i\alpha} - X^i X^\alpha |X|^{-2}) \circ dB^i, \quad \alpha = 1,2,\dots,n,$

where $|X|$ denotes the length of the vector $(X^1, X^2, \dots, X^n)$.

$$
\begin{aligned}
d|X|^2 &= 2 \sum_{\alpha=1}^{n} X^{\alpha} \circ dX^{\alpha} \\
&= 2 \sum_{\alpha=1}^{n} \sum_{i=1}^{n} \left(X^{\alpha}\delta_{\alpha i} - \frac{X^i (X^{\alpha})^2}{|X|^2}\right) \circ dB^i \\
&= 2 \sum_{i=1}^{n} (X^i - X^i) \circ dB^i \\
&= 0,
\end{aligned}
$$

so $|X| = \text{const}$. This shows that the solution of (S) is a diffusion process on a sphere $\Gamma$ with center at the origin. Let $\theta$ be any rotation around the origin and let $\sigma(X)$ denote the coefficient matrix on the right hand side of (S). Then

$$\sigma(\theta X) = \theta\, \sigma(X)\, \theta^{-1},$$

so

$$d(\theta X) = \theta\, dX = \theta\, \sigma(X) \circ dB = \theta\, \sigma(X)\, \theta^{-1} \circ \theta\, dB = \sigma(\theta X) \circ d(\theta B).$$

Since $\theta B$ is also an n-dimensional Brownian motion, the process $\theta X$ has the same probability law as $X$. Thus $X(t)$ is a rotation invariant diffusion process on the sphere $\Gamma$, i.e. a spherical Brownian motion.

## 3. Stochastic parallel transport

The parallel transport in an affinely connected $C^3$ manifold $M = (M, \Gamma^i_{jk})$ is given by the Levi-Civita equation:

$$du = \Gamma_k\, u\, dx^k$$

where $u$ is a tensor of type $(m,n)$. The stochastic analogue of this concept was introduced by Itô [5] for the special case of tensors of type $(0,n)$ in a Riemannian manifold and was extended by Dynkin [6] to the general case. We will discuss this in terms of stochastic calculus [7].

Let $C = C(t)$, $0 \le t < \infty$, be a random curve on $M$. It is said to belong to class $Q$ if for every $C^3$ function $f: M \to R$ the process $f(C(t))$, $0 \le t < \infty$, belongs to $\mathcal{Q}$. This definition can be phrased in terms of local coordinates: If for every local coordinate $(X^i(t))$ of $C$ and for every pair of previsible stopping times $(\mathcal{F}_t)$ $T_1 \le T_2$ such that $X^i(t)$ is well-defined for $T_1 \le t \le T_2$ the process $X^i((t \vee T_1) \wedge T_2)$, $0 \le t < \infty$, belongs to $\mathcal{Q}$, then $C$ is said to belong to class $Q$.

The spherical Brownian motion of the last section is of class $Q$.

Now we want to define the parallel transport $U(t)$, $0 \leq t < \infty$, of a tensor $u$ of type $(m,n)$ at $a$ along a random curve $C$ of class $Q$ starting at $a$. Since the sample curve of $C$ is not smooth in general, we cannot use the Levi-Civita equation in its original form, so we define $U(t)$ as follows. Take any division points of $[0, \infty)$:

$$\Delta = (0 = t_0 < t_1 < \dots)$$

and construct a random curve $C_\Delta = C_\Delta(t, \omega)$ by connecting $C(t_i, \omega)$ with $C(t_{i+1}, \omega)$ by a geodesic curve for every $i$. Then $C$ is piecewise smooth, so we can define the parallel transport $U(t)$ of $u$ for each sample curve by the Levi-Civita equation:

$$d\,U_\Delta = \Gamma_k U_\Delta \, dX^k, \qquad U_\Delta(0) = u.$$

By a routine method we can prove the existence of the limit process $U$ of $U_\Delta$ as

$$|\Delta| \equiv \sup_i |t_i - t_{i-1}| \to 0.$$

The limit process $U$ is called the <u>stochastic parallel transport</u> of $u$ along $C$.

<u>Theorem 4.1.</u> The above stochastic parallel transport $U$ is determined by the stochastic differential equation:

$$dU = (\Gamma_k U) \circ dX^k, \qquad U(0) = u.$$

<u>Proof.</u> Routine. We should note that this differential equation has the same form as in differential geometry because of use of symmetric $Q$-multiplication.

We can use stochastic parallel transport to discuss diffusions of tensors [5], [6], [7] and rolling along Brownian motion on a Riemannian manifold [7].

<u>BIBLIOGRAPHY</u>


1. D.L. Fisk, Quasi-martingales and stochastic integrals, Technical Report No.1, Dept Mathe., Michigan State Univ., 1963.
2. P. Courrège; Intégral stochastiques et martingales de carre integrable, Sem. Th. Potential 7. Inst. Henri Poincare, Paris, 1963.
3. H. Kunita and S. Watanabe: On square integrable martingales, Nagoya Math. J. 30 (1967) 209-245.

4. R.L. Stratonovich: Conditional Markov processes and their application to optimal control, Elsevier, New York (1968).
5. K. Itô, The Brownian motion and tensor fields on Riemannian manifold, Proc. Int. Congress Math. 1962 (Stockholm), 536-539.
6. E.B. Dynkin, Diffusion of tensors, Soviet Math. Dokl. Vol.9 (1968), No.2, 532-535.
7. K. Itô, Stochastic parallel displacement, Report of Symposium on Stochastic Differential equations, Univ. Victoria, Canada, to appear in Springer Lecture Note Series.
8. D.W. Stroock: On the growth of stochastic integrals, Z. Wahr. 18 (1971) 340-344.

## QUESTIONS AND ANSWERS

Question (T. Hida). The speaker emphasizes the significance of the second definition of stochastic integral, that is the one due to Stratonovich. So far as differential (stochastic) calculus is concerned, I agree with the speaker's opinion. However as soon as we come to integral calculus such as multiple integrals with respect to martingales, the first definition of the integral, namely the Itô integral, seems to be more important. How is the speaker's opinion on this?

Answer (Itô). The first definition is more basic, but in this lecture I emphasized that there are quite a few cases in which the second one is convenient.

Question (J.L. Lebowitz). Can you also conveniently describe the Ornstein-Uhlenbeck process on a sphere?

Answer (Itô). I hope so. There is one point to be noted. Since the Ornstein-Uhlenbeck process in $\mathbb{R}^n$ is a velocity process of a random motion in $\mathbb{R}^n$, the spherical analogue is not a process on a sphere but a process on the space of tangent bundles of the sphere.