An Introduction to Stochastic Ordinary Differential Equations (Part II)

Thomas Wanner

Department of Mathematical Sciences
George Mason University
Fairfax, Virginia 22030, USA

Introduction to Stochastic Differential Equations

IMA, Minneapolis, Minnesota

January 11, 2013

1. Probabilistic Background

Throughout this lecture, we assume the following framework:

- The triple $(F, \mathcal{F}, \mathbb{P})$ denotes a probability space over a set F, with σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$.
- The stochastic process $W : \mathbb{R}_0^+ \times F \to \mathbb{R}$ denotes a standard Wiener process over $(F, \mathcal{F}, \mathbb{P})$, i.e., we have:
  - The process satisfies $W(0, \omega) = 0$ for all $\omega \in F$.
  - For every $\omega \in F$ the path $W(\cdot, \omega)$ is continuous.
  - For every $0 \leq s \leq t$ the random variable $W(t, \cdot) - W(s, \cdot)$ has a Gaussian distribution with mean 0 and variance $t - s$.
  - The process W has independent increments, i.e., the m random variables $W(t_k, \cdot) - W(t_{k-1}, \cdot)$ for $k = 1, \ldots, m$ are independent, for any $0 \leq t_0 < \ldots < t_m$.

Random Variables

Recall that a random variable X is a mapping $X : F \to \mathbb{R}$ which is measurable with respect to $\mathcal{F}$. Its expected value is defined as

$$\mathbb{E}(X) = \int_F X \, d\mathbb{P} \,,$$

and its variance via

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2 \,.$$

Furthermore, if the random variables X and Y are independent, then we have

$$\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y) \,.$$

Throughout, we always assume that the indicated integrals exist and are finite.

Random Variables with Finite Second Moment

It will turn out to be very useful in the following to consider random variables in a Hilbert space setting. For this, let $\mathcal{G} \subset \mathcal{F}$ denote a σ-algebra. We consider the set of all $\mathcal{G}$-measurable random variables with finite second moment, i.e., we consider

$$L^2(\mathbb{P}, \mathcal{G}) = \left\{ X : F \to \mathbb{R} \;\middle|\; X \text{ is } \mathcal{G}\text{-measurable}, \; \int_F X^2 \, d\mathbb{P} < \infty \right\}$$

This space is a Hilbert space with norm and inner product given by

$$\|X\|_{L^2(\mathbb{P},\mathcal{G})} = \sqrt{\int_F X^2 \, d\mathbb{P}} \quad \text{and} \quad (X, Y)_{L^2(\mathbb{P},\mathcal{G})} = \int_F XY \, d\mathbb{P} \,,$$

and will be central for the definition of the stochastic integral. For the special case $\mathcal{G} = \mathcal{F}$ we use the abbreviation

$$L^2(\mathbb{P}) = L^2(\mathbb{P}, \mathcal{F}) \,.$$

Conditional Expectation

Let $X : F \to \mathbb{R}$ denote a random variable, and let $\mathcal{G} \subset \mathcal{F}$ denote a σ-algebra. Then the conditional expectation $\mathbb{E}(X|\mathcal{G})$ is defined as the (unique up to measure zero) random variable $Y : F \to \mathbb{R}$ which is measurable with respect to $\mathcal{G}$, and which satisfies

$$\int_G X \, d\mathbb{P} = \int_G Y \, d\mathbb{P} \quad \text{for all} \quad G \in \mathcal{G} \, .$$

If $X \in L^2(\mathbb{P})$, then one can show that the conditional expectation $\mathbb{E}(X|\mathcal{G})$ is the orthogonal projection of $X \in L^2(\mathbb{P})$ onto the closed subspace $L^2(\mathbb{P}, \mathcal{G})$.

For example, if $\mathcal{G} = \{\emptyset, F\}$, then the conditional expectation of X is the constant random variable $\mathbb{E}(X|\mathcal{G})(\omega) \equiv \mathbb{E}(X)$.

2. The Stochastic Ito Integral

We now turn our attention to the primary goal of this second lecture: How can we define an integral of the form

$$\int_S^T f(\tau,\omega)\, dW(\tau,\omega)$$

for suitable stochastic processes $f : \mathbb{R}_0^+ \times F \to \mathbb{R}$? This integral should be a random variable over $(F, \mathcal{F}, \mathbb{P})$.

Remarks:

- Recall that it is impossible to define this integral pathwise for fixed $\omega \in F$ in the Riemann-Stieltjes sense, since the paths of the Wiener process are not of bounded variation.
- Our goal will be to start by defining the integral for simple stochastic processes f first, and then to extend this definition to a larger class of integrands via approximation in $L^2(\mathbb{P})$.

Elementary Processes and Their Integral

Consider fixed times $0 \leq S < T$. We say that a stochastic process $f : [S, T] \times F \to \mathbb{R}$ is elementary, if it is piecewise constant in the following sense. There exists a partition

$$S = t_0 < t_1 < \ldots < t_n = T$$

as well as random variables $e_k : F \to \mathbb{R}$ such that

$$f(t, \omega) = e_k(\omega) \quad \text{for all} \quad t_k \leq t < t_{k+1} \quad \text{and} \quad \omega \in F \ ,$$

for all $k = 0, \ldots, n-1$. Then it is natural to define the integral of f with respect to the Wiener process as the sum

$$\int_S^T f(t, \omega)\, dW(t, \omega) \ = \ \sum_{k=0}^{n-1} e_k(\omega) \cdot (W(t_{k+1}, \omega) - W(t_k, \omega)) \ .$$

An Illustrative Example

Suppose now that we would like to approximate the value of the stochastic integral

$$\int_S^T W(t,\omega)\, dW(t,\omega)\,.$$

Since the paths of the Wiener process are continuous, both of the following approximations of the integrand W via elementary processes seem reasonable:

(L) For a partition $S = t_0 < \ldots < t_n = T$, consider $e_k(\omega) = W(t_k, \omega)$ for all k, ω, i.e., evaluate the integrand at the left endpoint of the partition interval.

(R) For a partition $S = t_0 < \ldots < t_n = T$, consider $e_k(\omega) = W(t_{k+1}, \omega)$ for all k, ω, i.e., evaluate the integrand at the right endpoint of the partition interval.

An Illustrative Example

For sufficiently fine partitions, the resulting approximations $A_L(\omega)$ and $A_R(\omega)$ should be close to each other, where

$$A_L(\omega) = \sum_{k=0}^{n-1} W(t_k,\omega)\cdot(W(t_{k+1},\omega) - W(t_k,\omega)) ,$$

$$A_R(\omega) = \sum_{k=0}^{n-1} W(t_{k+1},\omega)\cdot(W(t_{k+1},\omega) - W(t_k,\omega)) .$$

Yet, regardless of the choice of partition they cannot get arbitrarily close, which follows from the properties of the Wiener process:

$$\mathbb{E}(A_L) = \sum_{k=0}^{n-1} \mathbb{E}W(t_k)\cdot\mathbb{E}\left(W(t_{k+1}) - W(t_k)\right) = 0 ,$$

$$\mathbb{E}(A_R) = \sum_{k=0}^{n-1} \underbrace{\mathbb{E}\left(\left(W(t_{k+1}) - W(t_k)\right)^2\right)}_{=t_{k+1}-t_k} = T - S \gg \mathbb{E}(A_L) \, !$$

Specifying the Evaluation Point

The example indicates that when approximating a more general integrand f using elementary functions, one has to specify at which point $t_k^* \in [t_k, t_{k+1}]$ the integrand is being evaluated in the form

$$e_k(\omega) = f\left(t_k^*, \omega\right) \quad \text{for} \quad k = 0, \ldots, n-1 \; .$$

There are many possibilities, but two have proved to be useful:

- The Ito stochastic integral uses the choice

$$t_k^* = t_k \quad \text{for} \quad k = 0, \ldots, n-1 \; .$$

- The Stratonovich stochastic integral uses

$$t_k^* = \frac{t_k + t_{k+1}}{2} \quad \text{for} \quad k = 0, \ldots, n-1 \; .$$

We will only consider the Ito version of the stochastic integral.

Wiener Process Filtration and Adapted Processes

The approximation idea only works if we restrict the class of admissible integrands f. Intuitively, one needs to make sure that the random variable $f(t, \cdot)$ depends only on the behavior of the Wiener process W up to time t. More precisely, we need:

Definition (Wiener Process Filtration)

Let W denote a Wiener process over $(F, \mathcal{F}, \mathbb{P})$. For each $t \geq 0$ we define the σ-algebra $\mathcal{F}_t$ as the smallest σ-algebra $\mathcal{F}_t \subset \mathcal{F}$ generated by the random variables $W(s, \cdot)$ for $0 \leq s \leq t$. In other words, $\mathcal{F}_t$ is the smallest σ-algebra $\mathcal{F}_t \subset \mathcal{F}$ such that the random variables $W(s, \cdot)$ are $\mathcal{F}_t$-measurable for every $0 \leq s \leq t$.

Definition ($\mathcal{F}_t$-Adapted Process)

A stochastic process $f : \mathbb{R}_0^+ \times F \to \mathbb{R}$ is called adapted to $\mathcal{F}_t$ if the random variable $f(t, \cdot)$ is $\mathcal{F}_t$-measurable for all t.

The Filtration Associated with a Wiener Process

One can think of the σ-algebra $\mathcal{F}_t$ as the history of the Wiener process W up to time t. Note that we have

$$\mathcal{F}_s \subset \mathcal{F}_t \quad \text{for all} \quad 0 \leq s \leq t \, .$$

It can be shown that a random variable X is $\mathcal{F}_t$-measurable if and only if it is the pointwise almost everywhere limit of sums of functions of the form

$$g_1\left(W(s_1,\omega)\right) \cdot g_2\left(W(s_2,\omega)\right) \cdot \ldots \cdot g_m\left(W(s_m,\omega)\right)$$

where $g_1, \ldots, g_m$ are bounded continuous functions and $0 \leq s_k \leq t$ for all $k = 1, \ldots, m$ and $m \in \mathbb{N}$. In other words, the random variable X is $\mathcal{F}_t$-measurable if its values can be decided from the values of $W(s, \cdot)$ for $0 \leq s \leq t$.

The Class of Admissible Integrands

After these preparations, we can now define the class of possible integrands for the stochastic Ito integral.

Definition (Admissible Integrands)

Let $0 \leq S < T$ be fixed reals. Then the set of admissible integrands is defined as

$$\mathcal{V}(S,T) = \{f : [S,T] \times F \to \mathbb{R} \,|\, f \text{ satisfies (i), (ii), (iii) below}\},$$

where

(i) the process f is $\mathcal{B} \times \mathcal{F}$-measurable,

(ii) the random variable $f(t,\cdot)$ is $\mathcal{F}_t$-measurable for $t \in [S,T]$, i.e., the process f is $\mathcal{F}_t$-adapted,

(iii) we have

$$\mathbb{E}\left(\int_S^T f(t,\omega)^2 dt\right) < \infty\,.$$

The Class of Admissible Integrands

Remarks:

- The requirements *(ii)* and *(iii)* in the definition of $\mathcal{V}(S,T)$ can be relaxed significantly. But for this introductory lecture we stick to the above stricter situation.
- Under the assumptions given in the definition, the set $\mathcal{V}(S,T)$ has a Hilbert space structure with the inner product

$$(f,g)_{\mathcal{V}(S,T)} = \mathbb{E}\left(\int_S^T f(t,\omega)g(t,\omega)\,dt\right) .$$

- If $f \in \mathcal{V}(S,T)$ is an elementary process, i.e., if we have

$$f(t,\omega) = e_k(\omega) \quad \text{for all} \quad t_k \le t < t_{k+1} \quad \text{and} \quad \omega \in F ,$$

then e_k has to be measurable with respect to $\mathcal{F}_{t_k}$.

Definition of the Ito Integral

We can finally turn our attention to the definition of the Ito integral. For any integrand $f \in \mathcal{V}(S, T)$, this integral will be denoted by

$$\mathcal{I}[f](\omega) = \int_S^T f(t,\omega)\, dW(t,\omega) \;, \quad \text{and} \quad \mathcal{I}[f] \in L^2(\mathbb{P}) \;.$$

The definition proceeds in three steps:

(1) If $f \in \mathcal{V}(S, T)$ is an elementary process, define

$$\mathcal{I}[f](\omega) \;=\; \sum_{k=0}^{n-1} e_k(\omega) \cdot (W(t_{k+1},\omega) - W(t_k,\omega)) \;.$$

(2) Show that the following Ito isometry holds:

$$\|\mathcal{I}[f]\|_{L^2(\mathbb{P})} = \|f\|_{\mathcal{V}(S,T)} \quad \text{for all elementary } \; f \in \mathcal{V}(S, T) \;.$$

Definition of the Ito Integral

(3) Use the density of elementary processes in $\mathcal{V}(S,T)$ with respect to the norm $\|\cdot\|_{\mathcal{V}(S,T)}$ and the Ito isometry to extend the integral to all of $\mathcal{V}(S,T)$.
More precisely, one can show that for every $f \in \mathcal{V}(S,T)$ there exists a sequence of elementary processes $f_n \in \mathcal{V}(S,T)$ such that $\|f - f_n\|_{\mathcal{V}(S,T)} \to 0$ as $n \to \infty$. Then define

$$\mathcal{I}[f] = \lim_{n\to\infty} \int_S^T f_n(t,\omega)\, dW(t,\omega) \quad \text{in} \quad L^2(\mathbb{P}) .$$

Remarks:

- Notice that in contrast to the deterministic integral, one actually obtains a one-to-one correspondence between the integrand f and the stochastic Ito integral $\mathcal{I}[f]$.
- The Ito isometry relies heavily on the properties of W and $\mathcal{V}(S,T)$.

Proof of the Ito Isometry

We briefly sketch the proof of the Ito isometry, which makes use of the abbreviation $\Delta W_j = W(t_{j+1}) - W(t_j)$. Then one has:

- The central identity is given by

$$\mathbb{E}\left(e_i e_j \Delta W_i \Delta W_j\right) = \begin{cases} 0 & \text{for} \quad i \neq j \\ \mathbb{E}\left(e_i^2\right)\left(t_{i+1} - t_i\right) & \text{for} \quad i = j \end{cases}$$

  To see that this expression vanishes for $i < j$, one just has to note that since e_i is $\mathcal{F}_{t_i}$-measurable, e_j is $\mathcal{F}_{t_j}$-measurable, and ΔW_i is $\mathcal{F}_{t_{i+1}}$-measurable, the inequality $t_{i+1} \leq t_j$ shows that $e_i e_j \Delta W_i$ is $\mathcal{F}_{t_j}$-measurable. But ΔW_j is independent of $\mathcal{F}_{t_j}$ and has mean zero, which implies the first part of the identity. On the other hand, for $i = j$ one obtains

$$\mathbb{E}\left(e_i^2 \left(\Delta W_i\right)^2\right) = \mathbb{E}\left(e_i^2\right)\mathbb{E}\left(\Delta W_i\right)^2 = \mathbb{E}\left(e_i^2\right)\left(t_{i+1} - t_i\right) .$$

Proof of the Ito Isometry

- Recalling that

$$\int_S^T f(t,\omega)\,dW(t,\omega) = \sum_{i=0}^{n-1} e_i(\omega)\left(W(t_{i+1},\omega) - W(t_i,\omega)\right),$$

the central identity from the previous slide then implies

$$\begin{aligned}
\mathbb{E}\left(\left(\int_S^T f(t,\omega)dW(t,\omega)\right)^2\right) &= \sum_{i,j=0}^{n-1} \mathbb{E}\left(e_i e_j \Delta W_i \Delta W_j\right) \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left(e_i^2\right)\left(t_{i+1} - t_i\right) \\
&= \mathbb{E}\left(\int_S^T f(t,\omega)^2 dt\right),
\end{aligned}$$

and this completes the proof. □

Properties of the Ito Integral I

- Linearity: For $\alpha_1, \alpha_2 \in \mathbb{R}$ and $f_1, f_2 \in \mathcal{V}(S, T)$ we have

$$\int_S^T (\alpha_1 f_1(t,\omega) + \alpha_2 f_2(t,\omega))\, dW(t,\omega) = \alpha_1 \int_S^T f_1(t,\omega) dW(t,\omega) + \alpha_2 \int_S^T f_2(t,\omega) dW(t,\omega)\,.$$

- Additivity: For $0 \leq S < T < U$ and $f \in \mathcal{V}(S, U)$ we have

$$\int_S^U f(t,\omega) dW(t,\omega) = \int_S^T f(t,\omega) dW(t,\omega) + \int_T^U f(t,\omega) dW(t,\omega)$$

Properties of the Ito Integral II

- Measurability: The stochastic integral $\int_S^T f(t,\omega)dW(t,\omega)$ is measurable with respect to $\mathcal{F}_T$.
- Continuity: There exists an $\mathcal{F}_t$-adapted stochastic process J which is continuous with respect to t and which satisfies

$$\mathbb{P}\left(J(t,\omega) = \int_S^t f(\tau,\omega)dW(\tau,\omega)\right) = 1 \ .$$

- Approximation: If $f, f_n \in \mathcal{V}(S,T)$ satisfy $\|f - f_n\|_{\mathcal{V}(S,T)} \to 0$ as $n \to \infty$, then

$$\int_S^T f(t,\omega)dW(t,\omega) \overset{L^2(\mathbb{P})}{=} \lim_{n\to\infty} \int_S^T f_n(t,\omega)dW(t,\omega) \ .$$

Properties of the Ito Integral III

- Expected Value: For all $f \in \mathcal{V}(S, T)$ we have

$$\mathbb{E}\left(\int_S^T f(t,\omega)dW(t,\omega)\right) = 0 \ .$$

- Variance: For all $f \in \mathcal{V}(S, T)$ we have

$$\mathbb{V}\left(\int_S^T f(t,\omega)dW(t,\omega)\right) = \mathbb{E}\left(\int_S^T f(t,\omega)^2 dt\right) \ .$$

  In other words, the Ito isometry is valid on all of $\mathcal{V}(S, T)$.

- Martingale Property: For a suitable process $f : \mathbb{R}_0^+ \times F \to \mathbb{R}$ define $M_t(\omega) = \int_0^t f(\tau,\omega)dW(\tau,\omega)$ for all $t \geq 0$. Then

$$\mathbb{E}\left(M_t \,|\mathcal{F}_s\right) = M_s \quad \text{for all} \quad 0 \leq s \leq t \ .$$

Properties of the Ito Integral IV

- Martingale Inequalities: Define $M_t(\omega) = \int_0^t f(\tau,\omega) dW(\tau,\omega)$ for all $t \geq 0$ as before, and assume without loss of generality that M is continuous with respect to t. Then for all λ, $T > 0$ we have

$$\begin{aligned} \mathbb{P}\left(\sup_{0\leq t\leq T} |M_t| \geq \lambda \right) &\leq \frac{1}{\lambda^2} \cdot \mathbb{E}\left(M_T^2\right) \\ &= \frac{1}{\lambda^2} \cdot \mathbb{E}\left(\int_0^T f(s,\omega)^2 \, ds \right) \,, \end{aligned}$$

as well as

$$\mathbb{E}\left(\sup_{0\leq t\leq T} M_t^2 \right) \leq 4\mathbb{E}\left(M_T^2\right) = 4\mathbb{E}\left(\int_0^T f(s,\omega)^2 \, ds \right) .$$

Properties of the Ito Integral V

- Gaussianity: If the integrand function f is deterministic, i.e., if f is independent of ω, then the stochastic integral

$$\int_S^T f(t)dW(t,\omega) \quad \text{is a Gaussian random variable}$$

with

$$\text{mean } 0 \qquad \text{and} \qquad \text{variance } \int_S^T f(t)^2 dt \, .$$

If the integrands depends on ω, then in general the Ito integral is not a Gaussian random variable.

A First Example

We now demonstrate how the properties of the Ito integral can be used to show that

$$\int_0^T W(t,\omega)dW(t,\omega) \;=\; \frac{W(T,\omega)^2}{2} - \frac{T}{2} \, .$$

The idea is to approximate the integral by elementary processes in $\mathcal{V}(0, T)$. For this, let $0 = t_0 < \ldots < t_n = T$ denote a partition P of $[0, T]$ and consider the elementary process

$$f_P(t,\omega) = \sum_{k=0}^{n-1} W(t_k,\omega) \cdot \chi_{[t_k,t_{k+1})}(t) \, ,$$

where $\chi_A(t)$ denotes the characteristic function of a set A.

A First Example

First we need to show that as $\max(t_{k+1} - t_k) \to 0$, the elementary function f_P converges to the Wiener process in $\mathcal{V}(0, T)$. This can be seen as follows:

$$
\begin{aligned}
\|f_P - W\|^2_{\mathcal{V}(0,T)} &= \mathbb{E}\left(\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \left(W(t_k,\omega) - W(s,\omega)\right)^2 ds\right) \\
&= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\left(\left(W(s,\omega) - W(t_k,\omega)\right)^2\right) ds \\
&= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (s - t_k)\, ds \;=\; \sum_{k=0}^{n-1} \frac{(t_{k+1} - t_k)^2}{2} \\
&\to 0 \qquad \text{as long as } \max(t_{k+1} - t_k) \to 0\,.
\end{aligned}
$$

A First Example

Finally we need that for $\max(t_{k+1} - t_k) \to 0$, the integral of the elementary function f_P converges to $(W(T)^2 - T)/2$ in $L^2(\mathbb{P})$, since $\mathcal{I}[f_P] \to \int_0^T W(t,\omega)dW(t,\omega)$. This follows from

$$
\begin{aligned}
W(T)^2 &= \sum_{k=0}^{n-1} \left(W(t_{k+1})^2 - W(t_k)^2 \right) \\
&= \underbrace{\sum_{k=0}^{n-1} (W(t_{k+1}) - W(t_k))^2}_{\to T} \\
&\qquad + \underbrace{2 \sum_{k=0}^{n-1} W(t_k) \left(W(t_{k+1}) - W(t_k) \right)}_{= \mathcal{I}[f_P] \to \int_0^T W(t,\omega) dW(t,\omega)}
\end{aligned}
$$

Comments on the Ito Integral

- For integrands f which are adapted stochastic processes with a certain integrability condition, it is possible to define the stochastic Ito integral $\int_S^T f(t,\omega)\,dW(t,\omega)$ with respect to the Wiener process as a random variable in $L^2(\mathbb{P})$.
- The stochastic integral can not be defined path-wise, i.e., for fixed ω. It is constructed via a limit process in $L^2(\mathbb{P})$.
- Ito integration establishes a one-to-one correspondence between the integrand and the integral.
- The notion of the integral discussed here is inadequate for the vector-valued case, i.e., if one would like to integrate with respect to multi-dimensional Brownian motion. For this, the condition *(ii)* in the definition of $\mathcal{V}(S,T)$ has to be weakened.

3. The Stochastic Chain Rule

Just as in the deterministic setting, stochastic integrals are usually not computed through their definition, but by means of an associated stochastic Ito calculus. Such a calculus should involve versions of a change of variable formula and integration by parts. For this, it is convenient to introduce the following notion.

Definition (Ito Processes)

A stochastic process $X(t,\omega)$ is called Ito process, if it satisfies an integral equation of the form

$$X(t,\omega) = X(0,\omega) + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dW(s,\omega) \ ,$$

for suitable adapted stochastic processes u and v. If X is an Ito process, the above integral identity is generally abbreviated as

$$dX(t) = udt + vdW(t) \ .$$

Gaussian Ito Processes

In general, Ito processes are not Gaussian processes. There is, however, one special case in which this is true.

Lemma (Gaussian Ito Processes)

Let $X(t,\omega)$ be an Ito processes satisfying $dX(t) = udt + vdW(t)$ and assume that both u and v are deterministic functions of t. Furthermore, assume that $X(0,\omega) \equiv X_0$ is constant, i.e., we have

$$X(t,\omega) = X_0 + \int_0^t u(s)ds + \int_0^t v(s)dW(s,\omega) \; .$$

Then X is a Gaussian process with independent increments and

$$\mathbb{E}(X(t)) = X_0 + \int_0^t u(s)ds \quad \textit{and} \quad \mathbb{V}(X(t)) = \int_0^t v(s)^2 ds \; .$$

Stochastic Chain Rule

Theorem (Stochastic Chain Rule, Ito's Formula)

Let X be an Ito process with $dX(t) = udt + vdW(t)$, and let $g(t,x)$ denote a C^2-function. Then the process

$$Y(t,\omega) = g(t, X(t,\omega))$$

is again an Ito process and we have

$$\begin{aligned} dY(t) &= \frac{\partial g}{\partial t}(t, X(t))dt + \frac{\partial g}{\partial x}(t, X(t))dX(t) \\ &\quad + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X(t)) \cdot (dX(t))^2 \ , \end{aligned}$$

where $(dX(t))^2$ is computed using the rules $dW(t) \cdot dW(t) = dt$ and $dt \cdot dt = dt \cdot dW(t) = dW(t) \cdot dt = 0$.

Stochastic Chain Rule

Remarks:

- Notice that Ito's formula can be written equivalently in the following form

$$dY(t) = \left(\frac{\partial g}{\partial t}(t,X) + u \cdot \frac{\partial g}{\partial x}(t,X) + \frac{v^2}{2} \cdot \frac{\partial^2 g}{\partial x^2}(t,X)\right) dt + v \cdot \frac{\partial g}{\partial x}(t,X) dW(t) \,,$$

  where we omitted the argument from the Ito process X.
- While the formula reduces to the classical chain rule in the deterministic case $v \equiv 0$, the stochastic version introduces an additional term which depends on $\partial^2 g / \partial x^2$.

Stochastic Chain Rule

- Ito's formula is basically proved by assuming that u and v are elementary processes with respect to the same partition of the underlying interval, and using a Taylor approximation on each of the subintervals. Among other things, this leads to terms of the form

$$\sum_{k=0}^{n-1} v(t_k)^2 \cdot \frac{\partial^2 g}{\partial x^2}(t_k, X(t_k)) \cdot (W(t_{k+1}) - W(t_k))^2$$

One can show that in the space $L^2(\mathbb{P})$, this random variable converges to

$$\int_0^t v(s)^2 \cdot \frac{\partial^2 g}{\partial x^2}(s, X(s,\omega))ds$$

as $\max(t_{k+1} - t_k) \to 0$, which accounts for the extra term.

Example: $\int_0^T W(t,\omega)dW(t,\omega)$

From the deterministic theory we guess that the integral should include a term of the form $W(T)^2/2$. Thus, we consider

$$dX(t) = 0dt + 1dW(t) \qquad \text{and} \qquad g(t,x) = \frac{x^2}{2} \,.$$

For $Y(t) = g(t, X(t)) = W(t)^2/2$ Ito's formula then implies

$$dY(t) = \left(\underbrace{\frac{\partial g}{\partial t}}_{=0} + \underbrace{u \cdot \frac{\partial g}{\partial x}}_{=0} + \underbrace{\frac{v^2}{2} \cdot \frac{\partial^2 g}{\partial x^2}}_{=1/2} \right) dt + \underbrace{v \cdot \frac{\partial g}{\partial x} \, dW(t)}_{=W(t)dW(t)} \,,$$

which furnishes

$$\frac{W(T)^2}{2} \;=\; \frac{T}{2} + \int_0^T W(t,\omega)\, dW(t,\omega) \,.$$

Example: $\int_0^T t\,dW(t,\omega)$

From the deterministic theory we guess that the integral should include a term of the form $T \cdot W(T)$. Thus, we consider

$$dX(t) = 0dt + 1dW(t) \qquad \text{and} \qquad g(t,x) = tx\ .$$

For $Y(t) = g(t, X(t)) = t \cdot W(t)$ Ito's formula then implies

$$dY(t) = \left(\underbrace{\frac{\partial g}{\partial t}}_{=W(t)} + \underbrace{u \cdot \frac{\partial g}{\partial x}}_{=0} + \underbrace{\frac{v^2}{2} \cdot \frac{\partial^2 g}{\partial x^2}}_{=0} \right) dt + \underbrace{v \cdot \frac{\partial g}{\partial x}\, dW(t)}_{=t\, dW(t)},$$

which furnishes

$$T \cdot W(T) \ = \ \int_0^T W(t,\omega)\,dt + \int_0^T t\,dW(t,\omega)\ .$$

Generalizations of Ito's Formula

Ito's formula can be generalized in a number of ways to cover the case of vector-valued processes $X(t,\omega)$ and vector-valued Brownian motions $B(t,\omega)$. We only mention one such generalization.

Theorem (Ito's Formula for Multiple Ito Processes)

Let X_k, $k = 1, \ldots, d$, denote a family of Ito processes with respect to the same Wiener process, given by $dX_k(t) = u_k dt + v_k dW(t)$, and let $g(t, x_1, \ldots, x_d)$ denote a C^2-function. Then the process $Y(t,\omega) = g(t, X_1(t,\omega), \ldots, X_d(t,\omega))$ is an Ito process with

$$dY(t) = \frac{\partial g}{\partial t} dt + \sum_{k=1}^{d} \frac{\partial g}{\partial x_k} dX_k(t) + \frac{1}{2} \sum_{k,\ell=1}^{d} \frac{\partial^2 g}{\partial x_k \partial x_\ell} dX_k(t) dX_\ell(t),$$

where $dX_k(t) dX_\ell(t) = v_k v_\ell dt$, and the partial derivatives of g are evaluated at $(t, X_1(t,\omega), \ldots, X_d(t,\omega))$.

Stochastic Integration by Parts

Specifically for $g(t, x_1, x_2) = x_1 \cdot x_2$ one obtains the stochastic version of integration by parts.

Theorem (Stochastic Integration by Parts Formula)

Let X_1, X_2 denote two Ito processes with respect to the same Wiener process, given by $dX_k(t) = u_k dt + v_k dW(t)$ for $k = 1, 2$. Then their product $X_1 \cdot X_2$ is again an Ito process and we have

$$d\left(X_1(t) \cdot X_2(t)\right) = X_1 dX_2 + X_2 dX_1 + v_1 v_2 dt\ .$$

Notice in particular that if either X_1 or X_2 is an Ito process with paths of bounded variation, then the classical deterministic integration by parts formula holds.

4. Stochastic Differential Equations

In the final part of this lecture we finally turn our attention to stochastic differential equations, which are equations of the form

$$\boxed{dX = b(t,X)dt + \sigma(t,X)dW(t)}$$

where b and σ are sufficiently regular functions. We say that a stochastic process $X(t,\omega)$ is a solution of the above stochastic differential equation with initial condition $X(0,\omega) = X_0(\omega)$ if

$$\begin{aligned} X(t,\omega) &= X_0(\omega) + \int_0^t b(s, X(s,\omega))ds \\ &\qquad + \int_0^t \sigma(s, X(s,\omega))dW(s,\omega) \end{aligned}$$

for "suitable" $t \geq 0$. Under fairly natural conditions, such equations always have unique solutions.

Existence and Uniqueness of Solutions

Theorem (Existence and Uniqueness of Solutions of SDEs)

For measurable functions b and σ consider the SDE

$$\boxed{dX = b(t,X)dt + \sigma(t,X)dW(t)}$$

$$\begin{aligned} \textit{with} \qquad |b(t,x)| + |\sigma(t,x)| &\leq C\,(1+|x|)\ , \\ |b(t,x)-b(t,y)| + |\sigma(t,x)-\sigma(t,y)| &\leq D|x-y|\ , \end{aligned}$$

for all $t \in [0,T]$ and $x, y \in \mathbb{R}$. Furthermore, let $X_0(\omega)$ denote a square-integrable random variable which is independent of $W(t)$ for all $t \geq 0$. Then the above equation has a unique t-continuous solution $X(t,\omega)$ for $t \in [0,T]$. This solution is $\mathcal{F}_t$-adapted and

$$\mathbb{E}\left(\int_0^T |X(t,\omega)|^2\, dt\right) < \infty\ .$$

Remarks and Generalizations

- The existence and uniqueness result can be proved via Picard iteration, similarly to the deterministic situation. The necessary solution estimates make use of the martingale inequalities mentioned as one of the properties of the stochastic integral.
- An analogous existence and uniqueness theorem holds in the vector-valued case $X(t,\omega) \in \mathbb{R}^n$ for equations of the form

$$dX = b(t,X)dt + \sum_{k=1}^{m} \sigma_k(t,X)dW_k(t) ,$$

where the Wiener processes $W_1, \ldots, W_m$ are independent.

Remarks and Generalizations

- One can think of the stochastic differential equation

$$dX = b(t,X)dt + \sigma(t,X)dW(t)$$

as being a perturbation of the deterministic equation

$$\dot{X} = b(t,X) \ .$$

This latter equation is perturbed by additive noise if σ does not depend on X, and it is perturbed by multiplicative noise if σ does depend explicitly on X.

A Noisy Population Growth Model

As a first example we consider the noisy population growth model

$$dZ = aZdt + bZdW(t) \quad \text{with} \quad Z(0,\omega) = Z_0 \in \mathbb{R}$$

where a and b are real constants. This is the usual deterministic exponential growth model perturbed by multiplicative white noise.

To find the solution, we use the intuition from the deterministic situation to suggest that the solution might involve the term

$$e^{\alpha t + \beta W(t,\omega)} \quad \text{for certain} \quad \alpha, \beta \in \mathbb{R} \,.$$

Therefore, it seems natural to apply Ito's formula to the Ito process $X(t,\omega) = W(t,\omega)$ which satisfies $dX = 0dt + 1dW(t)$, and the nonlinearity $g(t,x) = \exp(\alpha t + \beta x)$.

A Noisy Population Growth Model

For $Y(t,\omega) = g(t, X(t,\omega)) = e^{\alpha t + \beta W(t,\omega)}$ Ito's formula implies

$$dY(t) = \left(\underbrace{\frac{\partial g}{\partial t}}_{=\alpha Y} + \underbrace{u \cdot \frac{\partial g}{\partial x}}_{=0} + \underbrace{\frac{v^2}{2} \cdot \frac{\partial^2 g}{\partial x^2}}_{=\beta^2 Y/2} \right) dt + \underbrace{v \cdot \frac{\partial g}{\partial x}\, dW(t)}_{=\beta Y dW(t)},$$

which shows that $Y(t,\omega)$ solves the stochastic differential equation

$$dY = \left(\alpha + \frac{\beta^2}{2} \right) Y dt + \beta Y dW(t)\,.$$

Comparing this with the form of the noisy population growth model furnishes

$$\boxed{Z(t,\omega) = Z_0 \cdot e^{\left(a - \frac{b^2}{2}\right) t + b W(t,\omega)}}$$

A Noisy Population Growth Model

- Note that due to

$$Z(t,\omega) = Z_0 + \int_0^t aZ(s,\omega)ds + \int_0^t bZ(s,\omega)dW(s,\omega)$$

and the properties of the Ito integral we have

$$\mathbb{E}\left(Z(t)\right) = Z_0 \cdot e^{at} \quad \text{for all} \quad t \geq 0 \,.$$

- The typical path behavior deviates from the deterministic case. For almost all paths of the Wiener process we have

$$\lim_{t\to\infty} \frac{W(t,\omega)}{t} = 0 \,,$$

which means that typical paths of Z satisfy

$$Z(t,\omega) = Z_0 e^{\left(a-\frac{b^2}{2}\right)t+bW(t,\omega)} \sim Z_0 e^{\left(a-\frac{b^2}{2}\right)t} \quad \text{for} \quad t \to \infty \,.$$

A Noisy Population Growth Model

- One can also determine the variance of the solution process $Z(t,\omega)$. Note that we have

$$Z(t,\omega)^2 = Z_0^2 \cdot e^{\left(2a-b^2\right)t+2bW(t,\omega)} \ ,$$

and as before this implies

$$d\left(Z^2\right) = \left(\left(2a-b^2\right) + \frac{(2b)^2}{2}\right) Z^2 dt + 2bZ^2 dW(t) \ ,$$

and therefore

$$\mathbb{E}\left(Z(t)^2\right) = Z_0^2 \cdot e^{\left(2a+b^2\right)t} \quad \text{for all} \quad t \geq 0 \ .$$

This finally implies

$$\mathbb{V}\left(Z(t)\right) = Z_0^2 \cdot e^{2at} \cdot \left(e^{b^2 t} - 1\right) \quad \text{for all} \quad t \geq 0 \ .$$

The Ornstein-Uhlenbeck Process

Our second example is the so-called Langevin equation

$$dZ = -aZdt + bdW(t) \quad \text{with} \quad Z(0,\omega) = Z_0 \in \mathbb{R}$$

where $a \neq 0$ and b are real constants. This is a deterministic linear equation perturbed by additive white noise, and its solution process is called the Ornstein-Uhlenbeck process.

In the deterministic setting, equations of this type are solved by moving the term $-aZdt$ to the left, multiplying the equation by the integrating factor e^{at}, and integrating to obtain $Z(t)e^{at}$ on the left-hand side. This suggests that in the stochastic case, it makes sense to determine whether the process $e^{at} \cdot Z(t,\omega)$ is an Ito process, i.e., apply Ito's formula to $X(t,\omega) = Z(t,\omega)$ with the transformation $g(t,x) = e^{at} \cdot x$, where $u = -aZ$ and $v = b$ in the representation $dX = udt + vdW(t)$.

The Ornstein-Uhlenbeck Process

For $Y(t,\omega) = g(t, X(t,\omega)) = e^{at} Z(t,\omega)$ Ito's formula implies

$$dY(t) = \left(\underbrace{\frac{\partial g}{\partial t}}_{=ae^{at}Z} + \underbrace{u \cdot \frac{\partial g}{\partial x}}_{=-aZ \cdot e^{at}} + \underbrace{\frac{v^2}{2} \cdot \frac{\partial^2 g}{\partial x^2}}_{=0} \right) dt + \underbrace{v \cdot \frac{\partial g}{\partial x} dW(t)}_{=b \cdot e^{at} dW(t)} ,$$

which shows that $Y(t,\omega)$ satisfies

$$\underbrace{Y(t,\omega) - Y(0,\omega)}_{= e^{at} Z(t,\omega) - Z_0} = \int_0^t b e^{as} dW(s,\omega) ,$$

and the Ornstein-Uhlenbeck process is therefore given by

$$\boxed{Z(t,\omega) = Z_0 \cdot e^{-at} + b \cdot \int_0^t e^{-a(t-s)} dW(s,\omega)}$$

The Ornstein-Uhlenbeck Process

- Note that the explicit formula for the Ornstein-Uhlenbeck process is exactly what one would obtain by applying the standard deterministic variation of constants formula to the Langevin equation.
- In contrast to the noisy population growth model discussed before, the Ornstein-Uhlenbeck process fits into the category of Gaussian Ito processes which was discussed at the beginning of this section. Thus, the Ornstein-Uhlenbeck process is a Gaussian process with independent increments.
- Due to the properties of the Ito integral we have

$$\mathbb{E}\,(Z(t)) = Z_0 \cdot e^{-at} \quad \text{for all} \quad t \geq 0\,.$$

The Ornstein-Uhlenbeck Process

- One can also easily determine the variance of the Ornstein-Uhlenbeck process $Z(t,\omega)$. Note that we have

$$\begin{aligned}
\mathbb{V}(Z(t)) &= \mathbb{E}\left(Z(t,\omega) - Z_0 \cdot e^{-at}\right)^2 \\
&= \mathbb{E}\left(be^{-at} \cdot \int_0^t e^{as} dW(s,\omega)\right)^2 \\
&= b^2 e^{-2at} \cdot \mathbb{V}\left(\int_0^t e^{as} dW(s,\omega)\right) \\
&= b^2 e^{-2at} \cdot \int_0^t e^{2as} ds \,.
\end{aligned}$$

This finally implies

$$\mathbb{V}(Z(t)) = \frac{b^2}{2a} \cdot \left(1 - e^{-2at}\right) \quad \text{for all} \quad t \geq 0 \,.$$

Thank You!

References:

- Bernt Oksendal: Stochastic Differential Equations: An Introduction with Applications. Springer-Verlag, 2010.
- Ludwig Arnold: Stochastic Differential Equations: Theory and Applications. Dover Publications, 2013.
- Lawrence C. Evans: An Introduction to Stochastic Differential Equations. Online Lecture Notes, 2006.