Lecture 18 : Itō Calculus

1. Ito's calculus

In the previous lecture, we have observed that a sample Brownian path is nowhere differentiable with probability 1. In other words, the differentiation

$$\frac{dB_t}{dt}$$

does not exist. However, while studying Brownain motions, or when using Brownian motion as a model, the situation of estimating the difference of a function of the type

$$f(B_t)$$

over an infinitesimal time difference occurs quite frequently (suppose that f is a smooth function). To be more precise, we are considering a function $f(t, B_t)$ which depends only on the second variable. Hence there exists an implicit dependence on time since the Brownian motion depends on time.

If the differentiation $\frac{dB_t}{dt}$ existed, then we can easily do this by using chain rule:

$$df = \left(\frac{dB_t}{dt} \cdot f'(B_t)\right) dt.$$

We already know that the formula above makes no sense.

One possible way to work around this problem is to try to describe the difference df in terms of the difference dB_t. In this case, the equation above becomes

$$df = f'(B_t) dB_t. \tag{1.1}$$

Our new formula at least makes sense, since there is no need to refer to the differentiation $\frac{dB_t}{dt}$ which does not exist. The only problem is that it does not quite work. Consider the Taylor expansion of f to obtain

$$f(x + \Delta x) - f(x) = (\Delta x) \cdot f'(x) + \frac{(\Delta x)^2}{2} f''(x) + \frac{(\Delta x)^3}{6} f'''(x) + \cdots .$$

To deduce Equation (1.1) from this formula, we must be able to say that the significant term is the first term $(\Delta x) \cdot f'(x)$ and all other terms are of smaller order of magnitude. Is this true for $x = B_t$? For $x = B_t$, we have

$$\Delta f = (\Delta B_t) \cdot f'(B_t) + \frac{(\Delta B_t)^2}{2} f''(x) + \frac{(\Delta B_t)^3}{6} f'''(x) + \cdots .$$

Now consider the term $(\Delta B_t)^2$. Since B_t is a Brownian motion, we know that $\mathbb{E}[(\Delta B_t)^2] = \Delta t$. Since a difference in B_t is necessarily accompanied by a difference in t, we see that the second term is no longer negligable. The theory of Ito calculus essentially tells us that we can make the substitution

$(\Delta B_t)^2 = \Delta t$, and the remaining terms are negligable. Hence the equation above becomes

$$\Delta f = (\Delta B_t) \cdot f'(B_t) + \frac{\Delta t}{2} f''(x) + \cdots ,$$

which in terms of infinitesimals becomes

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2} f''(B_t)dt. \tag{1.2}$$

This equation known as the *Ito's lemma* is the main equation of Ito's calculus.

More generally, consider a smooth function $f(t, x)$ which depends on two variables, and suppose that we are interested in the differential of $f(t, B_t)$. In classical calculus, we will get

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx,$$

but in Ito calculus, we will have

$$\begin{aligned} df(t, B_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dB_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dB_t)^2 \\ &= \left(\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \frac{\partial f}{\partial x} dB_t. \end{aligned}$$

Theorem 1.1. *(Ito's lemma) Let $f(t, x)$ be a smooth function of two variables, and let X_t be a stochastic process satisfying $dX_t = \mu_t dt + \sigma_t dB_t$ for a Brownian motion B_t. Then*

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2} \sigma_t^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \frac{\partial f}{\partial x} dB_t.$$

Proof. We have

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_t)^2 \\ &= \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2} \sigma_t^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} dB_t + ..dtdB_T + ..(dt)^2. \end{aligned}$$

We can ignore the terms $dtdB_t$ and $(dt)^2$. $\square$

Definition 1.2. We define integration as an inverse of differentiation, i.e.,

$$F(t, B_t) = \int f(t, B_t) dB_t + \int g(t, B_t) dt,$$

if and only if

$$dF = f(t, B_t)dB_t + g(t, B_t)dt.$$

Example 1.3. (i) The stochastic process $X_t = \mu t + \sigma B_t$ is known as the Brownian motion with drift μ and variance σ. For this process, we have

$$dX_t = \mu dt + \sigma dB_t.$$

(ii) Consider the function $f(x) = \frac{1}{2}x^2$. We see that

$$df(B_t) = B_t dB_t + \frac{1}{2}dt.$$

Equiavlently,

$$\frac{1}{2}B_T^2 = \int_0^T B_t dB_t + \int_0^T \frac{1}{2}dt = \int_0^T B_t dB_t + \frac{T}{2}.$$

This implies that

$$\int_0^T B_t dB_t = \frac{1}{2}B_T^2 - \frac{T}{2}.$$

Note how this 'violates' the fundamental theorem of calculus.

(iii) Let $f(t, x) = exp(\mu t + \sigma x)$. Then

$$df(t, B_t) = (\mu + \frac{1}{2}\sigma^2)f(t, B_t)dt + \sigma f(t, B_t)dB_t.$$

We can now answer the question of finding the stochastic process $X_t(t, B_t)$ such that

$$dX_t = \sigma X_t \cdot dB_t.$$

To do this we can just set $\mu = -\frac{1}{2}\sigma^2$ in the function above, i.e., $X_t(t, B_t) = exp(-\frac{1}{2}\sigma^2 t + \sigma B_t)$.

(iv) Let $f(t, x) = t^2 + x^2$, and $X_t = \mu t + \sigma B_t$. Then

$$\begin{aligned}
df(t, X_t) &= 2tdt + 2X_t dX_t + (dX_t)^2 \\
&= 2tdt + 2(\mu dt + \sigma dB_t) + \sigma^2 dt \\
&= (2t + 2\mu + \sigma^2)dt + 2\sigma dB_t.
\end{aligned}$$

We discussed an elegant way to work around the fact that B_t is not differentiable to understand the difference of a function of a Brownian motion over a small period of time. For the remainder of this class, we will study the properties of Ito calculus by addressing the following type of questions:

(1) Given $g(t, B_t) = \int a\, dB_t + \int b\, dt$ for some functions a and b, is there a simple way to describe the variance of g?
(2) Given $g(t, B_t)$ as above, when is g a martingale?
(3) Suppose that $b = 0$. Then when is $g(t, B_t)$ normally distributed at time t?

Remark. The theory of calculus can be extended to cover Brownian motions in several different ways which are all 'correct' (in other words, there can be several different versions of Ito's calculus). For example, there exists a theory of calculus where

$$df = f'(B_t)dB_t - \frac{1}{2}f''(B_t)dt.$$

However, Ito integral is the most natural one in the context of how the time variable fits into the theory, because the fact that we cannot see the future is the basis of the whole theory. We will further study this in next section.

2. Properties of Ito calculus

First theorem can be seen as an extension of the fact that the sum of independent normal random variables is a random variable.

Theorem. *Let $B(t)$ be a Brownian motion, and let $\Delta(t)$ be a nonrandom function of time. Suppose that a stochastic process $I(t)$ satisfies*

$$dI = \Delta(s) dB_s \quad i.e. \quad I(t) = \int \Delta(s) dB_s.$$

where $I(0) = 0$. Then for each $t \geq 0$, the random variable $I(t)$ is normally distributed.

What happens when we allow $\Delta(t)$ to be a random function of time? Here is an interesting and natrual class of random varialbes $\Delta(t)$ that we consider.

Definition 2.1. Let X_t be a stochastic process. A process Δ_t is called an *adapted process* (with respect to X_t) if for all $t \geq 0$, the random variable Δ_t depends only on X_s for $s \leq t$.

Example 2.2. Let X_t be a stochastic process.
(i) The process $\Delta(t) = X_t$ is an adapted process.
(ii) The process $\Delta(t) = \min\{X_t, c\}$ is an adapted process (where c is a constant).
(iii) The process $\Delta(t) = \max_{0 \leq t \leq T} X_t$ is not an adapted process.
(iv) If τ is a stopping time, then X_τ is an adapted process.

For example, suppose that we model the price of a stock using a stochastic process, and are trying to find a strategy which has positive expected return. Consider a simple strategy where at each time t, we either buy or sell one stock, hence $\Delta_t = 1$ or -1. Our strategy only makes sense if Δ_t is an adapted process, since otherwise it contradicts the fact that we cannot see the future.

Our second theorem asserts that for a Brownian motion B_t, the Ito integral of an adapted process with respect to B_t is also a martingale.

Theorem 2.3. *Let B_t be a Brownian motion. Then for all adapted processes $g(t, B_t)$, the integral*

$$\int g(t, B_t)\, dB_s$$

is a martingale, as long as g is a 'reasonable function'. Formally, if $g \in L^2$, i.e.,

$$\int \int_0^t g^2(t, B_t) dt dB_t < \infty.$$

Example 2.4. The process B_t itself is an adapted process. Recall that

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{t}{2},$$

and $\mathbb{E}[B_t^2] = t$. Hence

$$\mathbb{E}[\int_0^t B_s dB_s] = 0.$$

More generally,

$$\mathbb{E}[\int_{t_1}^{t_2} B_s dB_s \,|\, \mathcal{F}_{t_1}] = \mathbb{E}[\left(\frac{1}{2}B_{t_2}^2 - \frac{t_2}{2}\right) \,|\, \mathcal{F}_1] - \left(\frac{1}{2}B_{t_1}^2 - \frac{t_1}{2}\right)$$
$$= \frac{1}{2}(t_2 - t_1) + \frac{1}{2}B_{t_1}^2 - \frac{t_2}{2} - \frac{1}{2}B_{t_1}^2 + \frac{t_1}{2} = 0.$$

This confirms the theorem above for $\Delta(t) = B_t$.

Here is another useful fact about the Ito integral of an adapted process known as Ito isometry. It can be used to compute the variance of the Ito integral.

Theorem 2.5. *(Ito isometry) Let B_t be a Brownian motion. Then for all adapted processes $\Delta(t)$, we have*

$$\mathbb{E}\left[\left(\int_0^t \Delta(s) dB_s\right)^2\right] = \mathbb{E}\left[\int_0^t \Delta(s)^2 ds\right].$$

Example 2.6. Let $\Delta(t) = 1$. Then the left hand side of the theorem above is

$$\mathbb{E}\left[\left(\int_0^t \Delta(s) dB_s\right)^2\right] = \mathbb{E}[B_t^2] = t,$$

and the right hand side of the above is

$$\mathbb{E}\left[\int_0^t \Delta(s)^2 ds\right] = \mathbb{E}[t] = t.$$

Note that $\int_0^t \Delta(s) dB_s = 0$ by Theorem 2.1 given above. Hence Ito isometry tell us how to compute the variance of this integral .

3. Change of measure

Following is a quote from [3].

> Can a stochastic process with drift also be viewed as a process without drift? This modestly paradoxical question is no mere curiosity. It has many important consequences, the most immediate of which is the discovery that almost any question about Brownian motion with drift may be rephrased as a parallel question about standard Brownian motion.

Recall that a stochastic process is a probability distribution over a set of paths. A *change of measure* of a stochastic process is a method of shifting the probability distribution into another probability distribution. In this section, we fix a final time T and suppose that all paths are defined over the time $0 \leq t \leq T$. Also, we take a more abstract view of defining stochastic processes in terms of an underlying space Ω. Hence a stochastic process X is a function $X : \Omega \to [0,T]^\infty$ for $\omega \in \Omega$, where $\mathbf{P}$ is a probability distribution over Ω and describes the probability distribution of paths through X^{-1}, i.e. $\mathbf{P}(A) = \mathbf{P}(X^{-1}(A))$ for all set of paths $A \subset [0,T]^\infty$. We denote a particular realization of a stochastic process as X_ω.

Let Z be a positive random variable satisfying $\int Z(\omega)\, dX(\omega) = 1$. Then we can define a new stochastic process $\tilde{X}$ whose probability distribution is given by

$$\tilde{\mathbf{P}}(A) = \int_A Z(\omega)\, dX(\omega),$$

for all sets A. Since Z is positive, we see that

$$\mathbf{P}(A) > 0 \Leftrightarrow \tilde{\mathbf{P}}(A) > 0 \quad \forall A. \tag{3.1}$$

Hence the set of paths 'observed' under the probability measures $\mathbf{P}$ and $\tilde{\mathbf{P}}$ are the same.

Definition 3.1. We say that two probability distributions $\mathbf{P}$ and $\tilde{\mathbf{P}}$ are *equivalent* if (3.1) holds.

When changing measures, we fix a set Ω of possible paths, and change a probability distribution $\mathbf{P}$ over Ω, and make it into another equivalent (property 3.1) probability distribution over Ω. Thus the underlying space is the same but we only change our point of view. This is not true for general transformations. For example, consider the square of a path. The probability distribution of the square of a Brownian motion $B(t)^2$ is not equivalent to the probability distribution of $B(t)$.

The function Z is known as the *Radon-Nikodým derivative* of $\tilde{\mathbf{P}}$ with respect to $\mathbf{P}$, and is denoted as

$$Z = \frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}.$$

Changing measures is of theoretical importance since it provides a tool to understand the relation between two different but equivalent stochastic processes. It is also of practical importance, since converting one probability distribution into another can reveal hidden insights. For example, in finance, we can convert a non-martingale stochastic process into a martingale by changing measure, and this gives a method of pricing financial derivatives.

We now come back to the original question that we posted: 'Can a stochastic process with drift also be viewed as a process without drift?'. We now see that this is the same question as asking whether the two stochastic processes are equivalent. Our first theorem asserts that this indeed is the case for Brownian motions.

Theorem 3.2. *(Girsanov's theorem, simple version) Let $(\Omega, \mathbf{P})$ be a probability space, and let $X : \Omega \to [0,T]^\infty$ be a stochastic process which is a Brownian motion with drift μ under the probability distribution inudced by $(\Omega, \mathbf{P})$. Consider the probability distribution $\tilde{\mathbf{P}}$ over Ω defined as*

$$\frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}(\omega) = e^{-\mu\omega(T) - \mu^2 T/2}.$$

Then X is a Brownian motion with no drift under the probability distribution induced by $(\Omega, \tilde{\mathbf{P}})$.

Following is a more general version of the theorem above.

Theorem 3.3. *(Girsanov's theorem) Let $(\Omega, \mathbf{P})$ be a probability space, and let $X : \Omega \to [0,T]^\infty$ be a stochastic process which is a Brownian motion with no drift under the probability distribution induced by $(\Omega, \mathbf{P})$. Let $\Delta(t)$ be a bounded adapted process, and consider the probability distribution $\tilde{\mathbf{P}}$ over Ω defined as*

$$\frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}(\omega) = e^{-\int_0^T \Delta(u) d\omega - \frac{1}{2}\int_0^T \Delta(u)^2 du}.$$

Then the stochastic process defined as

$$Y_\omega(t) = X_\omega(t) + \int_0^t \Delta(u) du,$$

is a Brownian motion under the probability distribution induced by $(\Omega, \tilde{\mathbf{P}})$.

4. Remarks

We avoided all technalities in this lecture note. Formalizaing the theory of Ito calculus requires a solid background in measure theory.

References

[1] P. Wilmott, S. Howison, J. Dewynne, The mathematics of financial derivatives
[2] S. Shreve, Stochastic calculus for finance II: continuous-time models
[3] M. Steele, Stochastic calculus and financial applications
[4] Failure of pathwise integration for FV processes, http://almostsure.wordpress.com/2010/06/02/failure-of-pathwise-integration-for-fv-processes/
[5] R. Karandikar, On pathwise stochastic integration, *Stochastic Process. Appl. 57 (1995), p. 11-18.*
[6] M. Nutz, Pathwise construction of stochastic integrals