Chapter 1

Brownian motion

Note: This chapter and the next are adapted from related chapters of the following two books as

[1] *Fima C Klebaner,* INTRODUCTION TO STOCHASTIC CALCULUS WITH APPLICATIONS*, Imperial College Press, 1998.*

[2] *Steven E. Shreve,* STOCHASTIC CALCULUS FOR FINANCE II: CONTINUOUS-TIME MODELS*, Springer, 2004.*

The text notes are only for the Course "Introduction to Modern Mathematics" for PhD students of Zhejiang University.

Zhang Li-Xin

1.1 Basic concepts on stochastic processes

A stochastic process X is an umbrella term for any collection of random variables $\{X(t,\omega)\}$ depending on time t, which is defined on the same probability space $(\Omega, \mathscr{F}, \mathsf{P})$. Time can be discrete, for example, $t = 0, 1, 2, \ldots$, or continuous, $t \geq 0$. For fixed time t, the observation is described by a random variable which we denote by X_t or $X(t)$. For fixed $\omega \in \Omega$, $X(t)$ is a single realization (single path) of this process. Any single path is a function of time t, $x_t = x(t)$, $t \geq 0$.

At a fixed time t, properties of the random variable $X(t)$ are described by a probability distribution of $X(t)$, $\mathsf{P}(X(t) \leq x)$.

A stochastic process is determined by all its finite dimensional distributions, that is, probabilities of the form

$$\mathsf{P}\big(X(t_1) \le x_1, X(t_2) \le x_2, \ldots, X(t_n) \le x_n\big), \tag{1.1.1}$$

for any choice of time points $0 \le t_1 < t_2 < \ldots < t_n$, any $n \ge 1$ with $x_1, \ldots, x_n \in \mathbb{R}$.

1.1.1 Gaussian process

If all finite dimensional distributions of a stochastic process is Gaussian (multi normal), then the process is called a Gaussian process. Because, a multivariate normal distribution is determined by its mean and covariance matrix, a Gaussian process is determined by it mean function $m(t) = \mathsf{E}X(t)$ and covariance function $\gamma(t,s) = \mathsf{Cov}\{X(t), X(s)\}$.

1.2 Brownian motion

1.2.1 Definition of Brownian motion

Definition 1.2.1 *Brownian motion $\{B(t)\}$ is a stochastic process with the following properties.*

1. *(Independence of increments) For all $0 = t_0 < t_1 < \ldots < t_m$, the increments*

$$B(t_1) - B(t_0),\ \ B(t_2) - B(t_1),\ \ \ldots,\ \ B(t_m) - B(t_{m-1})$$

are independent.

2. *(Stationary normal increments) $B(t) - B(s)$ has normal distribution with mean zero and variance $t - s$.*

3. *(Continuity of paths) $B(t)$,$t \ge 0$ are continuous functions of t.*

If the process is started at x, then $B(t)$ has the $N(x,t)$ distribution. This can be written as

$$P_x\big(B(t) \in (a,b)\big) = \int_a^b \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} dy.$$

P_x denotes the probability of events when the process starts at x. The function under the above integral is called the transition probability density of Brownian motion,

$$p_t(x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}.$$

If $B^x(t)$ denotes a Brownian motion started at x, then $B^x(t) - x$ is a Brownian motion started at 0, and $B^0(t) + x$ is Brownian motion started at$x$, in other words

$$B^x(t) = x + B^0(t).$$

So, usually we also assume $B(0) = 0$ if not specified, that is, the process is started at 0.

Example 1.2.1 *Calculate* $P(B(1) \le 0, B(2) \le 0)$.

Solution.

$$\begin{aligned}
& P(B(1) \le 0, B(2) \le 0) \\
=& P(B(1) \le 0, B(1) + B(2) - B(1) \le 0) \\
=& \int_{-\infty}^{\infty} P(B(1) \le 0, B(1) + B(2) - B(1) \le 0 | B(1) = y_1) \frac{1}{\sqrt{2\pi}} e^{-\frac{y_1^2}{2}} dy_1 \\
=& \int_{-\infty}^{0} P(B(1) + B(2) - B(1) \le 0 | B(1) = y_1) \frac{1}{\sqrt{2\pi}} e^{-\frac{y_1^2}{2}} dy_1 \\
=& \int_{-\infty}^{0} P(B(2) - B(1) \le -y_1) \frac{1}{\sqrt{2\pi}} e^{-\frac{y_1^2}{2}} dy_1 \\
=& \int_{-\infty}^{0} \Phi(-y_1) d\Phi(y_1) = \int_{-\infty}^{0} [1 - \Phi(y_1)] \, d\Phi(y_1) = \frac{1}{2} - \int_0^{1/2} y dy = \frac{3}{8}.
\end{aligned}$$

1.2.2 Distribution of Brownian motion

Brownian motion is a Gaussian process

Because the increments

$$B(t_1) = B(t_1) - B(0), \quad B(t_2) - B(t_1), \quad \ldots, \quad B(t_m) - B(t_{m-1})$$

are independent and normal distributed, as their linear transform, the random variables $B(t_1), B(t_2), \ldots, B(t_m)$ are jointly normally distributed, that is, the finite dimensional of Brownian motion is multivariate normal. So Brownian motion is a Gaussian process with mean 0 and covariance function

$$\gamma(t,s) = \mathsf{Cov}\{B(t), B(s)\} = \mathsf{E} B(t) B(s).$$

If $t < s$, then $B(s) = B(t) + B(s) - B(t)$, and

$$\mathsf{E}B(t)B(s) = \mathsf{E}B^2(t) + \mathsf{E}B(t)(B(s) - B(t)) = \mathsf{E}B^2(t) = t.$$

Similarly if $t > s$, $\mathsf{E}B(t)B(s) = s$. Therefore

$$\gamma(t, s) = \min(t, s).$$

On the other hand, a continuous mean zero Gaussian process with covariance function $\gamma(t, s) = \min(t, s)$ is a Brownian motion.

Example 1.2.2 *1. For any $T > 0$, $\{T^{-1/2}B(Tt)\}$ is Brownian motion.*

2. The process

$$\xi_0 \frac{t}{\sqrt{\pi}} + \frac{2}{\pi} \sum_{j=1}^{\infty} \frac{\sin(jt)}{j} \xi_j,$$

where ξ_j's, $, = 0, 1, \dots,$ are independent standard nomal random variables, is Brownian motion on $[0, \pi]$.

3. $\{-B(t), t \geq 0\}$ is also a Brownian motion.

4. $\{tB(\frac{1}{t}), t > 0\}$ is also a Brownian motion.

5. If $B(t)$ is a Brownian motion on $[0, 1]$, then $(t+1)B(\frac{1}{t+1}) - B(1)$ is a Brownian motion on $[0, \infty)$.

The first statement is the self-similarity property of the Brownian motion. The second is the random series representation of Brownian motion. The third is the symmetry of Brownian motion. The fourth allows to transfer results on the behavior of the paths of Brownian motion for large t to that of small t. The second and the the last show the existence of Brownian motion. Each of the above can be shown by checking the mean and covariance function.

The finite dimensional distribution of Brownian motion

Notice, the jointly density function of increments

$$B(t_1) = B(t_1) - B(0), \;\; B(t_2) - B(t_1), \;\; \dots, \;\; B(t_m) - B(t_{m-1})$$

is

$$p_{t_1}(0,x_1)p_{t_2-t_1}(0,x_2)\cdots p_{t_m-t_{m-1}}(0,x_m).$$

The jointly density function of

$$\begin{aligned}B(t_1) =&B(t_1),\\ B(t_2) =&B(t_1)+[B(t_2)-B(t_1)],\\ &\cdots\cdots\\ B(t_m) =&B(t_1)+[B(t_2)-B(t_1)]+\ldots+[B(t_m)-B(t_{m-1})],\end{aligned}$$

is

$$p_{t_1}(0,y_1)p_{t_2-t_1}(y_1,y_2)\cdots p_{t_m-t_{m-1}}(y_{m-1},y_m).$$

So,

$$\begin{aligned}&\mathsf{P}\Big(B(t_1)\le x_1, B(t_2)\le x_2,\ldots,B(t_m)\le x_m)\Big)=\\ &\quad\int_{-\infty}^{x_1}p_{t_1}(0,y_1)dy_1\int_{-\infty}^{x_2}p_{t_2-t_1}(y_1,y_2)dy_2\cdots\int_{-\infty}^{x_m}p_{t_m-t_{m-1}}(y_{m-1},y_m)dy_m.\end{aligned}$$

In general, if the process starts at x, the

$$\begin{aligned}&P_x\Big(B(t_1)\le x_1, B(t_2)\le x_2,\ldots,B(t_m)\le x_m)\Big)=\\ &\quad\int_{-\infty}^{x_1}p_{t_1}(x,y_1)dy_1\int_{-\infty}^{x_2}p_{t_2-t_1}(y_1,y_2)dy_2\cdots\int_{-\infty}^{x_m}p_{t_m-t_{m-1}}(y_{m-1},y_m)dy_m.\end{aligned}$$

Example 1.2.3 *We know that*

$$\frac{1}{\sqrt{n}}S_{[nt]}\overset{D}{\to}B(t).$$

Then

$$n^{-3/2}\sum_{m=1}^{n-1}S_m=\int_0^1\frac{1}{\sqrt{n}}S_{[nt]}\overset{D}{\to}\int_0^1 B(t)dt.$$

Next, we want to find the distribution of $\int_0^1 B(t)dt$.

Notice,

$$\int_0^1 B(t)dt=\lim\sum B(t_i)(t_{i+1}-t_i),$$

and $\sum B(t_i)(t_{i+1}-t_i)$ are normal random variables with mean zeros. So $\int_0^1 B(t)dt$ is a normal random variable with mean zero. On the other hand,

$$\begin{aligned}\mathsf{Var}\left\{\int_0^1 B(t)dt\right\} =& \mathsf{E}\left[\int_0^1 B(t)dt\int_0^1 B(s)ds\right] = \mathsf{E}\left[\int_0^1\int_0^1 B(t)B(s)dtds\right]\\ =& \int_0^1\int_0^1 \mathsf{E}[B(t)B(s)]dtds = \int_0^1\int_0^1 \min(t,s)dtds = 1/3.\end{aligned}$$

Exchanging the integrals and expectation is justified by Fubini's theorem since

$$\int_0^1\int_0^1 E|B(t)B(s)|dtds \leq \int_0^1\int_0^1 \sqrt{\mathsf{E}B^2(t)\mathsf{E}B^2(s)}dsdt < 1.$$

Thus $\int_0^1 B(t)dt$ has $N(0,1/3)$ distribution.

1.2.3 Filtration for Brownian motion

In addition to the Brownian motion itself, we will need some notation for the amount of information available at each time, We do that with a filtration.

Definition 1.2.2 *Let $(\Omega,\mathscr{F},P)$ be a probability space on which is defined a Brownian motion $B(t)$, $t\geq 0$. A filtration for the Brownian motion is a collection of σ-algebras $\mathscr{F}_t$,$t\geq 0$, satisfying*

(i) *(Information accumulates) For $0\leq s<t$, every $\mathscr{F}_s$ is also in $\mathscr{F}_t$. In other words, there is at least as much information available at the later time $\mathscr{F}_t$ as there is at the earlier time $\mathscr{F}_s$.*

(ii) *(Adaptivity) For each $t\geq 0$, the Brownian motion $B(t)$ at time t is $\mathscr{F}_t$-measurable. In other words, the information available at time t is sufficient to evaluate the Bownian motion $B(t)$ at that time.*

(iii) *(Independence of future increments) For $0\leq t<u$, the increment $B(u)-B(t)$ is independent of $\mathscr{F}_t$. In other words, any increments of the Brownian motion after time t is independent of the information available at time t.*

Properties (i) and (ii) in the definition above guarantee that the information available at each time t is at least as much as one would learn from observing the Brownian motion up to time t. Property (iii) says that this information is of no use in predicting future movements of the Brownian motion.

If $\mathscr{F}_t = \sigma(B(u) : u \leq t)$, then $\mathscr{F}_t$ is called the nature filtration of the Brownian motion.

Here, for a stochastic process $X(t)$,$t \geq 0$, $\sigma(X(u), u \leq t)$ is the smallest σ-field that contains sets of the form $\{a \leq X(u) \leq b\}$ for all $0 \leq u \leq t$, $a, b \in \mathbb{R}$. It is the information available to an observer of X up to time t.

In general,

Definition 1.2.3 *A family $\mathbb{F} = \{\mathscr{F}_t\}$ of increasing sub σ-fields on $(\Omega, \mathscr{F})$ is called a filtration. Usually, $\mathscr{F}_0$ is defined to be $\{\emptyset, \Omega\}$. $(\Omega, \mathscr{F}, \mathbb{F}, P)$ is called the filtered probability space.*

A stochastic process $X(t)$, $t \geq 0$, is called adapted if for all t, $X(t)$ is $\mathscr{F}_t$-measurable, that is, if for any t, $\mathscr{F}_t$ contains all the information about $X(t)$ (as well as all $X(u)$, $u \leq t$) but it may contain some extra information as well.

For the Brownian motion, the extra information contained in $\mathscr{F}_t$ is not allowed to give clues about the future increments of B because of Property (iii).

1.3 Properties of Brownian motion paths

1.3.1 Continuity and differentiability

Almost every sample path $B(t)$, $0 \leq t \leq T$

1. is a continuous function of t,

2. is not monotone in any interval, no matter how small the interval is;

3. is not differentiable at any point.

Property 1 is know. Property 2 follows from Property 3. For Property 3, one can show that

$$\lim_{h \to 0} \inf_{0 \leq t \leq T} \max_{0 < s \leq h} \frac{|B(t+s) - B(t)|}{h} = \infty \quad a.s.$$

We will not prove this result. Instead, we prove a weaker one.

Theorem 1.3.1 *For every t_0,*

$$\limsup_{t\to t_0}\left|\frac{B(t)-B(t_0)}{t-t_0}\right| = \infty \ \ a.s.,$$

which implies that for any t_0, almost every sample $B(t)$ is not differentiable at this point.

Proof. Without loss of generality, we assume $t_0 = 0$. If one considers the event

$$A(h,\omega) = \left\{ \sup_{0<s\le h}\left|\frac{B(s)}{s}\right| > D \right\},$$

then for any sequence $\{h_n\}$ decreasing to 0, we have

$$A(h_n,\omega) \supset A(h_{n+1},\omega),$$

and

$$A(h_n,\omega) \supset \left\{ \left|\frac{B(h_n)}{h_n}\right| > D \right\}.$$

So,

$$\mathsf{P}(A(h_n)) \ge \mathsf{P}\left(|B(h_n)/\sqrt{h_n}| > D\sqrt{h_n}\right) = \mathsf{P}(|B(1)| > D\sqrt{h_n}) \to 1 \ \ n\to\infty.$$

Hence,

$$\mathsf{P}\left(\bigcap_{n=1}^{\infty} A(h_n)\right) = \lim P(A(h_n)) = 1.$$

It follows that

$$\sup_{0<s\le h_n}\left|\frac{B(s)}{s}\right| \ge D \ \ a.s. \text{ for all } n \text{ and } D>0.$$

Hence

$$\lim_{h\to 0}\sup_{0<s\le h}\left|\frac{B(s)}{s}\right| = \infty \ \ a.s.$$

1.3.2 Variation and quadratic variation

If g is a function real variable, define its quadratic variation over the interval $[0,t]$ as limit (when it exists)

$$[g,g](t) = [g,g]([0,t]) = \lim\sum_{i=1}^{n}(g(t_i^n) - g(t_{i-1}^n))^2, \qquad (1.3.1)$$

where limit is taken over partitions:

$$0 = t_0^n < t_1^n < \ldots < t_n^n = t,$$

with $\delta = \max_{1\le i\le n}(t_i^n - t_{i-1}^n) \to 0$.

The variation of g on $[0,t]$ is

$$V_g(t) = \sup \sum_{i=1}^{n} \left|g(t_i^n) - g(t_{i-1}^n)\right|, \tag{1.3.2}$$

where supremum is taken over partitions:

$$0 = t_0^n < t_1^n < \ldots < t_n^n = t.$$

Notice that the summation in (1.3.2) is non-decreasing on the partitions, that is, the thinner is the partition, the larger is the summation. So

$$V_g(t) = \lim \sum_{i=1}^{n} \left|g(t_i^n) - g(t_{i-1}^n)\right|,$$

where limit is taken over partitions:

$$0 = t_0^n < t_1^n < \ldots < t_n^n = t,$$

with $\delta = \max_{1\le i\le n}(t_i^n - t_{i-1}^n) \to 0$. However, the quadratic variation has not such property.

From the definition of the quadratic variation, the follow theorem follows.

Theorem 1.3.2 *1. Let $a \le c \le b$. Then the quadratic variation of g over the interval $[a,b]$ exists if and only if its quadratic variations on both the intervals $[a,c]$ and $[c,b]$ exist. Further*

$$[g,g]([a,b]) = [g,g]([a,c]) + [g,g]([c,b]), \quad a \le c \le b.$$

2. If $g(t) = A + B(f(t) - C)$ for $t \in [a,b]$, then

$$[g,g]([a,b]) = B^2 \cdot [f,f]([a,b]).$$

Theorem 1.3.3 *If g is continuous and of finite variation then its quadratic variation is zero.*

Proof.

$$\begin{aligned}
[g,g](t) &= \lim \sum_{i=0}^{n-1} (g(t_{i+1}^n) - g(t_i^n))^2 \\
&\le \lim \max_{1\le i\le n-1} \left|g(t_{i+1}^n) - g(t_i^n)\right| \cdot \sum_{i=0}^{n-1} \left|g(t_{i+1}^n) - g(t_i^n)\right| \\
&\le \lim \max_{1\le i\le n-1} \left|g(t_{i+1}^n) - g(t_i^n)\right| \cdot V_g(t) = 0.
\end{aligned}$$

For Brownian motion B (or other stochastic process), we can similarly define its variation and quadratic variation. In such case, the variation and quadratic variation are both random variables. Also, in the definition of quadratic variation, the limit in (1.3.1) must be specified in which sense, convergence in probability, almost sure convergence or others, because the summation is now a random variables. We take the weakest one, limit in probability.

Definition 1.3.1 *The quadratic variation of Brownian motion $B(t)$ is defined as*

$$[B,B](t) = [B,B]([0,t]) = \lim \sum_{i=1}^{n} \left| B(t_i^n) - B(t_{i-1}^n) \right|^2, \tag{1.3.3}$$

where for each n, $\{t_i^n, 0 \le i \le n\}$ is a partition of $[0,t]$, and the limit is taken over all partitions with $\delta_n = \max_i (t_{i+1}^n - t_i^n) \to 0$ as $n \to \infty$, and in the sense of convergence in probability.

Theorem 1.3.4 *Quadratic variation of a Brownian motion over $[0,t]$ is t, that is, the Brownian motion accumulates quadratic variation at rate one per unit time.*

Proof. Let $T_n = \sum_{i=1}^n \left| B(t_i^n) - B(t_{i-1}^n) \right|^2$. Then

$$\mathsf{E}T_n = \sum_{i=1}^{n} \mathsf{E} \left| B(t_i^n) - B(t_{i-1}^n) \right|^2 = \sum_{i=1}^{n} (t_i^n - t_{i-1}^n) = t.$$

We want to show that $T_n - \mathsf{E}T_n \to 0$ in probability. Notice for a standard normal random variable N,

$$\mathsf{E}N^4 = 3, \quad \mathsf{E}N^2 = 1.$$

So,

$$\mathsf{Var}(N^2) = \mathsf{E}N^4 - (\mathsf{E}N^2)^2 = 3 - 1 = 2.$$

It follows that for a normal $N(0,\sigma^2)$ random variable X,

$$\mathsf{Var}\big(X^2\big) = \sigma^4 \mathsf{Var}\big((X/\sigma)^2\big) = \sigma^4 \cdot 2\big(\mathsf{Var}(X/\sigma)\big)^2 = 2\big(\mathsf{Var}X\big)^2.$$

Hence

$$\mathsf{Var}\big((B(t) - B(s))^2\big) = 2\big(\mathsf{Var}(B(t) - B(s))\big)^2 = 2(t-s)^2.$$

The variance of T_n is

$$\begin{aligned}\mathsf{Var}(T_n) =& \sum_{i=1}^n \mathsf{Var}\left(\left|B(t_i^n) - B(t_{i-1}^n)\right|^2\right)\\ =& \sum_{i=1}^n 2(t_i^n - t_{i-1}^n)^2\\ \le& 2\max_i (t_i^n - t_{i-1}^n)\sum_{i=1}^n (t_i^n - t_{i-1}^n) \le 2\delta_n t.\end{aligned}$$

So

$$T_n - \mathsf{E}T_n \to 0 \quad \text{in probability.}$$

The proof is completed.

Actually, we have shown that $T_n \to t$ in L_2, that is

$$\mathsf{E}|T_n - t|^2 \to 0.$$

If one uses the following inequality for independent mean-zero random variables $\{X_i\}$ as

$$\mathsf{E}\left|\sum_{i=1}^n X_i\right|^p \le C_p\left\{\Big(\sum_{i=1}^n \mathsf{Var}X_i\Big)^{p/2} + \sum_{i=1}^n \mathsf{E}|X_i|^p\right\} \text{ for } p \ge 2,$$

it can be showed that

$$\mathsf{E}|T_n - \mathsf{E}T_n|^p \le C_p\left\{(\delta_n t)^{p/2} + \delta_n^{p-1}t\right\} \to 0 \quad \text{for all } p.$$

So, for Brownian motion, the limit in the definition of quadratic variation can be taken in sense of any L_p convergence. However, in the sense of almost sure convergence, the limit does not exists unless additional condition on δ_n is assumed for example, $\delta_n = o(1/(\log n)^{1/2})$.

Combing Theorem 1.3.3 and Theorem 1.3.4 yields the following corollary.

Corollary 1.3.1 *Almost every path $B(t)$ of Brownian motion has infinite variation on any interval, no matte how small it is.*

The quadratic variation property of the Brownian motion is important for considering the stochastic calculus. It means that

$$\big(B(t_{j+1}) - B(t_j)\big)^2 \approx t_{j+1} - t_j.$$

Although this approximating is not true locally, because

$$\frac{\big(B(t_{j+1}) - B(t_j)\big)^2}{t_{j+1} - t_j}$$

is a normal $N(0,1)$ random variable and so not near 1, no matter how small we make $t_{j+1} - t_j$, the approximating makes sense if we consider the summation over a partition on a interval since the summation of the left hand can be approximated by the summation of the right hand.

We may write the approximation informally in the differential notations

$$dB(t)dB(t) = (dB(t))^2 = dt.$$

Recall that for a real differential function $f(x)$, we have

$$f(t_{j+1}) - f(t_j) \approx f'(t_j)(t_{j+1} - t_j),$$

that is

$$df(t) = f'(t)dt.$$

And so

$$df(t)df(t) = (df(t))^2 = (f'(t))^2(dt)^2 = 0.$$

This difference will be the main different issue between the stochastic calculus and usual real calculus.

1.3.3 Law of large numbers and law of the iterated logarithm

Theorem 1.3.5 *(Law of Large Numbers)*

$$\lim_{t\to\infty} \frac{B(t)}{t} = 0 \ \ a.s.$$

Theorem 1.3.6 *(Law of the Iterated Logarithm)*

$$\limsup_{t\to\infty} \frac{B(t)}{\sqrt{2t\log\log t}} = 1, \quad a.s.$$

$$\liminf_{t\to\infty} \frac{B(t)}{\sqrt{2t\log\log t}} = -1 \ \ a.s.$$

Notice $\widehat{B}(t) = t(B(\frac{1}{t})$ is also a Brownian motion.

Theorem 1.3.7 *(Law of the Iterated Logarithm)*

$$\limsup_{t\to 0} \frac{B(t)}{\sqrt{2t\log\log(1/t)}} = 1, \quad a.s.$$

$$\liminf_{t\to 0} \frac{B(t)}{\sqrt{2t\log\log(1/t)}} = -1 \quad a.s.$$

1.4 Martingale property for Brownian motion

Definition 1.4.1 *A stochastic process $\{X(t), t \geq 0\}$ is a martingale if for any t it is integrable, $E|X(t)| < \infty$, and for any $s > 0$*

$$E[X(t+s)|\mathscr{G}_t] = X(t) \quad a.s., \tag{1.4.1}$$

where $\mathscr{G}_t$ is the information about the process up to time t, that is, $\{\mathscr{G}_t\}$ is is a collection of σ-algebras such that

1. *$\mathscr{G}_u \subset \mathscr{G}_t$, if $u \leq t$;*
2. *$X(t)$ is $\mathscr{G}_t$ measurable.*

Theorem 1.4.1 *Let $B(t)$ be a Brownian motion. Then*

1. *$B(t)$ is a martingale.*
2. *$B^2(t) - t$ is a martingale.*
3. *For any u, $e^{uB(t)-\frac{u^2}{2}t}$ is a martingale.*

Proof. Let $\mathscr{F}_t$ is a filtration for $B(t)$.

1. By definition, $B(t) \sim N(0,t)$, so that $B(t)$ is integrable with $\mathsf{E}B(t) = 0$. Then, for $s < t$,

$$\begin{aligned} \mathsf{E}[B(t)|\mathscr{F}_s] =& \mathsf{E}[B(s) + \big(B(t) - B(s)\big)|\mathscr{F}_s] \\ =& \mathsf{E}[B(s)|\mathscr{F}_s] + \mathsf{E}[B(t) - B(s)|\mathscr{F}_s] \\ =& B(s) + \mathsf{E}[B(t) - B(s)] = B(s). \end{aligned}$$

So, $B(t)$ is a martingale.

2. Notice $\mathsf{E}B^2(t) = t < \infty$. Therefore $B^2(t) - t$ is integrable. Since,

$$\begin{aligned}\mathsf{E}[B^2(t) - t|\mathscr{F}_s] =& \mathsf{E}\left[B(s)^2 + \big(B(t) - B(s)\big)^2 + 2B(s)\big(B(t) - B(s)\big)|\mathscr{F}_s\right] - t \\ =& B(s)^2 + \mathsf{E}\big(B(t) - B(s)\big)^2 + 2B(s)\mathsf{E}[B(t) - B(s)] - t \\ =& B(s)^2 + (t - s) + 0 - t = B^2(t) - s.\end{aligned}$$

$B^2(t) - t$ is a martingale.

3. The moment generating function of the $N(0, t)$ variable $B(t)$ is

$$\mathsf{E}e^{uB(t)} = e^{tu^2/2} < \infty.$$

Therefore $e^{uB(t) - tu^2/2}$ is integrable with

$$\mathsf{E}e^{uB(t) - tu^2/2} = 1.$$

Further for $s < t$

$$\begin{aligned}\mathsf{E}\left[e^{uB(t)}|\mathscr{F}_s\right] =& \mathsf{E}\left[e^{uB(s)}e^{u(B(t) - B(s))}|\mathscr{F}_s\right] \\ =& e^{uB(s)}\mathsf{E}\left[e^{u(B(t) - B(s))}|\mathscr{F}_s\right] \\ =& e^{uB(s)}\mathsf{E}\left[e^{u(B(t) - B(s))}\right] \\ =& e^{uB(s)}e^{(t-s)u^2/2}.\end{aligned}$$

It follows that

$$\mathsf{E}\left[e^{uB(t) - tu^2/2}|\mathscr{F}_s\right] = e^{uB(s) - su^2/2},$$

and then $e^{uB(t) - tu^2/2}$ is a martingale.

All three martingales have a central place in the theory. The martingale $B^2(t) - t$ provides a characterization (Levy's characterization) of Brownian motion. It can be shown that if a process $X(t)$ is a continuous martingale such that $X^2(t) - t$ is also a martingale, then $X(t)$ is Brownian motion. The martingale $e^{uB(t) - tu^2/2}$ is known as the exponential martingale. Actually, a continuous process $B(t)$ with property 3 and $B(t) = 0$ must be a Brownian motion.

Theorem 1.4.2 *Let $X(t)$ be a continuous process such that for any u, $e^{uX(t) - tu^2/2}$ is a martingale. Then $X(t)$ is a Brownian motion.*

Proof. Since $e^{uX(t)-tu^2/2}$ is a martingale,

$$\mathsf{E}[e^{uX(t)-tu^2/2}|\mathscr{F}_s] = e^{uX(s)-su^2/2}.$$

It follows that

$$\mathsf{E}[e^{u\{X(t)-X(s)\}}|\mathscr{F}_s] = e^{(t-s)u^2/2}.$$

Taking expectation yields the moment generating function of $X(t) - X(s)$ is

$$\mathsf{E}[e^{u\{X(t)-X(s)\}}] = e^{(t-s)u^2/2}.$$

So, $X(t)-X(s) \sim N(0, t-s)$. The above conditional expectation also tells us that the conditional moment generating function given $\mathscr{F}_s$ is a non-random function, which implies that $X(t) - X(s)$ is independent of $\mathscr{F}_s$ and so independent of all $X(u)$, $u \leq s$. So, by definition, $X(t)$ is a Brownian motion.

1.5 Markov Property of Brownian motion

1.5.1 Markov Property

Definition 1.5.1 *Let $X(t), t \geq 0$ be a stochastic process on filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}, P)$. The process is called a Markov process if for any t and $s > 0$, the conditional distribution of $X(t+s)$ given $\mathscr{F}_t$ is the same as the conditional distribution of $X(t+s)$ given $X(t)$, that is,*

$$P\big(X(t+s) \leq y|\mathscr{F}_t\big) = P\big(X(t+s) \leq y|X(t)\big);$$

or equivalently, if for any t and $s > 0$ and every nonnegative Borel-measurable function f, there is another Borel-measurable function g such that

$$E\left[f\big(X(t+s)\big)|\mathscr{F}_t\right] = g\big(X(t)\big).$$

Theorem 1.5.1 *Brownian motion $B(t)$ process Markov property.*

Proof. It is easy to see by using the moment generating function that the conditional distribution of $B(t+s)$ given $\mathscr{F}_t$ is the same as that given $B(t)$. Indeed,

$$\begin{aligned}
&\mathsf{E}\left[e^{uB(t+s)}|\mathscr{F}_t\right]\\
=&e^{uB(t)}\mathsf{E}\left[e^{u\{B(t+s)-B(t)\}}|\mathscr{F}_t\right]\\
=&e^{uB(t)}\mathsf{E}\left[e^{u\{B(t+s)-B(t)\}}\right] \quad \text{since } e^{u\{B(t+s)-B(t)\}} \text{ is independent of } \mathscr{F}_t\\
=&e^{uB(t)}\mathsf{E}\left[e^{u\{B(t+s)-B(t)\}}|B(t)\right] \quad \text{since } e^{u\{B(t+s)-B(t)\}} \text{ is independent of } B(t)\\
=&\mathsf{E}\left[e^{uB(t+s)}|B(t)\right].
\end{aligned}$$

Remark 1.5.1 *With the conditional characteristic function taking the place of conditional moment generating function, one can show that every stochastic process with independent increments is a Markov process.*

Transition probability function of a Markov process X is defined as

$$P(y,t;x,s) = P(X(t) \le y|X(s) = x)$$

the conditional distribution function of the process at time t, given that it is at point x at time s.

In the case of Brownian motion the ransition probability function is given by the distribution of the normal $N(x, t-s)$ distribution

$$\begin{aligned}
P(y,t;x,s) &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{-\frac{(u-x)^2}{2(t-s)}\right\} du\\
&= \int_{-\infty}^{y} p_{t-s}(x,u)du.
\end{aligned}$$

It is easy seen that, for Brownian motion $B(t)$, $P(y,t;x,s) = P(y,t-s;x,0)$. In other words,

$$\mathsf{P}(B(t) \le y|B(s) = x) = P(B(t-s) \le y|B(0) = x).$$

The above property states that Brownian motion is time-homogeneous, that is, its distributions do not change with a shift in time.

Now, we calculate $\mathsf{E}[f(B(t))|\mathscr{F}_s]$.

$$\mathsf{E}[f(B(t))|B(s) = x] = \int f(y)P(dy,t;x,s) = \int f(y)p_{t-s}(x,y)dy.$$

So,

$$\mathsf{E}[f(B(t))|\mathscr{F}_s] = \mathsf{E}[f(B(t))|B(s)] = \int f(y)p_{t-s}(B(s),y)dy.$$

1.5.2 Stopping time

Definition 1.5.2 *A random time T is called a stopping for $B(t)$, $t \geq 0$, if for any t it is possible to decide whether T has occurred or not by observing $B(s), 0 \leq s \leq t$. More rigorously, for any t the sets $\{T \leq t\} \in \mathscr{F}_t = \sigma(B(s), 0 \leq s \leq t)$, the σ-field generated by $B(s), 0 \leq s \leq t$.*

If T is a stopping time, events observed before or at time T are described by σ-field $\mathscr{F}_T$, defined as the collection of sets

$$\{A \in \mathscr{F} : \text{ for any } t,\ A \cap \{T \leq t\} \in \mathscr{F}_t\}.$$

Example 1.5.1

1. *Any nonrandom time T is a stopping time. Formally, $\{T \leq t\}$ is either the $\emptyset$ or Ω, which are members of $\mathscr{F}_t$ for any t.*

2. *The first passage time of level a, $T_a = \inf\{t > 0 : B(t) = a\}$ is a stoping time. Clearly, if we know $B(s)$ for all $s \leq t$ then we know whether the Brownian motion took the value a before or at t or not. Thus we know that $\{T_a \leq t\}$ has occurred or not just by observing the past of the process prior to t. Formally,*

$$\{T_a \leq t\} = \{\max_{u \leq t} B(u) \geq a\} \in \mathscr{F}_t.$$

3. *Let T be the time when Brownian motion reaches its maximum on the interval $[0, 1]$. Then clearly, to decide whether $\{T \leq t\}$ has occurred or not, it is not enough to know the values of the process prior to time t, one needs to know all the values on the interval $[0, 1]$. So that T is not a stopping time.*

1.5.3 Strong Markov property

Strong Markov property is similar to the Markov property, except that in the definition a fixed time t is replaced by a stopping time.

Theorem 1.5.2 *Recall that P_x denotes the probability of events when the Brownian motion starts at x. Brownian motion $B(t)$ has the strong Markov property: for a finite stopping time T the regular conditional distribution of $B(T+t), t \geq 0$ given $\mathscr{F}_T$*

is $P_{B(T)}$, that is, given $\mathscr{F}_T$, $\{B(T+t), t \geq 0\}$ is a Brownian motion which starts at $B(T)$. In particular,

$$\mathsf{P}\big(B(T+t) \leq y | \mathscr{F}_T\big) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y - B(T))^2}{2t}\right\} dy. \tag{1.5.1}$$

From (1.5.1) it follows that given $\mathscr{F}_T$, $B(T+t) - B(S) \sim N(0,t)$. And so, $B(T+t) - B(S)$ is a $N(0,t)$ random variable and is independent of $\mathscr{F}_T$. So,

$$\mathsf{P}\big(B(T+t) \leq y | \mathscr{F}_T\big) = \mathsf{P}\big(B(T+t) \leq y | B(T)\big),$$

which is the strong Markov property of $B(t)$.

An equivalent result of Theorem 1.5.2 is the following corollary.

Corollary 1.5.1 *Let T be a finite stopping time. Define the new process in $t \geq 0$ by*

$$\widehat{B}(t) = B(T+t) - B(T).$$

Then $\widehat{B}(t)$ is a Brownian motion started at zero and independent of $\mathscr{F}_T$.

Proof. Fix $0 = t_0 \leq t_1 \leq \cdots \leq t_n$, and reals $\theta_1, \cdots, \theta_n$ and take some bounded $\mathscr{F}_T$ measurable random variables ξ. Let $s_j = t_j - t_{j-1}$, $Z_j = \widehat{B}(t_j) - \widehat{B}(t_{j-1})$, $j = 1, \cdots, n$. We want to show that

$$\mathsf{E}\left[\xi \exp\left\{\sum_{j=1}^{n} \theta_j Z_j\right\}\right] = \mathsf{E}[\xi] \exp\left\{\frac{1}{2}\sum_{j=1}^{n} \theta_j^2 s_j\right\}.$$

Let S be a stopping time. Note that

$$M(t) = \exp\left\{\theta B(t) - \frac{1}{2}\theta^2 t\right\}$$

is a martingale. By applying the optional stopping time theorem to the stopping time $S \wedge N$, we obtain

$$\mathsf{E}\left[\exp\left\{\theta B(S \wedge N + t) - \frac{1}{2}\theta^2 (S \wedge N + t)\right\} \Big| \mathscr{F}_{S \wedge N}\right] = \exp\left\{\theta B(S \wedge N) - \frac{1}{2}\theta^2 (S \wedge N)\right\},$$

i.e.,

$$\mathsf{E}\left[\exp\left\{\theta\big(B(S \wedge N + t) - B(S \wedge N)\big)\right\} \Big| \mathscr{F}_{S \wedge N}\right] = \exp\left\{\frac{1}{2}\theta^2 t\right\}.$$

Letting $N \nearrow \infty$ yields

$$\mathsf{E}\left[\exp\left\{\theta\big(B(S+t) - B(S)\big)\right\} \Big| \mathscr{F}_S\right] = \exp\left\{\frac{1}{2}\theta^2 t\right\}.$$

In fact, let $A \in \mathscr{F}_T$. For $n \le N$, let $B_n = A \cap \{S \le n\}$. Then $B_n \cap \{S \wedge N \le t\} = A \cap \{S \le t \wedge n\} \in \mathcal{F}_{t\wedge n} \subset \mathcal{F}_t$. So, $B_n \in \mathscr{F}_{S\wedge N}$. It follows that

$$\mathsf{E}\left[I_{A\cap\{S\le n\}} \exp\left\{\theta\big(B(S\wedge N+t) - B(S\wedge N)\big)\right\}\right] = \mathsf{E}\left[I_{A\cap\{S\le n\}}\right] \exp\left\{\frac{1}{2}\theta^2 t\right\}.$$

Letting $N \to \infty$ yields

$$\mathsf{E}\left[I_{A\cap\{S\le n\}} \exp\left\{\theta\big(B(S+t) - B(S)\big)\right\}\right] = \mathsf{E}\left[I_{A\cap\{S\le n\}}\right] \exp\left\{\frac{1}{2}\theta^2 t\right\}.$$

Then, letting $n \to \infty$ yields

$$\mathsf{E}\left[I_A \exp\left\{\theta\big(B(S+t) - B(S)\big)\right\}\right] = \mathsf{E}\left[I_A\right] \exp\left\{\frac{1}{2}\theta^2 t\right\}, \quad \forall A \in \mathscr{F}_T.$$

Now, taking $S = T + t_{n-1}$ and $t = s_n$ yields

$$\begin{aligned}
&\mathsf{E}\left[\xi \exp\left\{\sum_{j=1}^{n} \theta_j Z_j\right\}\right] \\
=&\mathsf{E}\left[\xi \exp\left\{\sum_{j=1}^{n-1} \theta_j Z_j\right\} \mathsf{E}\left[\exp\left\{\theta_n\big(B(S+s_n) - B(S)\big)\right\} \Big| \mathscr{F}_{T+t_{n-1}}\right]\right] \\
=&\mathsf{E}\left[\xi \exp\left\{\sum_{j=1}^{n-1} \theta_j Z_j\right\}\right] \exp\left\{\frac{1}{2}\theta_n^2 s_n\right\}.
\end{aligned}$$

Conditioning successively on $\mathscr{F}_{T+t_k}$ $(k = n-2, \cdots, 0)$ gives

$$\mathsf{E}\left[\xi \exp\left\{\sum_{j=1}^{n} \theta_j Z_j\right\}\right] = \mathsf{E}[\xi] \exp\left\{\frac{1}{2}\sum_{j=1}^{n} \theta_j^2 s_j\right\}.$$

The proof is completed. □

In the above proof, we applied the following stopping theorem of a martingale, the proof is omitted.

Lemma 1.5.1 *(i) (Stopped martingale) If $M(t)$ is a martingale with filtration $\{\mathscr{F}_t\}$ and τ is a stopping time, then the stopped process $M(t\wedge\tau)$ is a martingale, in particular for any t, $\mathsf{E}M(t\wedge\tau) = \mathsf{E}M(0)$.*

(ii) (Optional stopping time theorem) If $M(t)$ is a right-continuous martingale with filtration $\{\mathscr{F}_t\}$ and τ, σ are bounded stopping times such that $P(\tau \le \sigma) = 1$. Then

$$\mathsf{E}[M(\sigma)|\mathscr{F}_\tau] = M(\tau).$$

1.6 Functions of Brownian motion

1.6.1 The first passage time

Let x be a real number, the first passage time of Brownian motion $B(t)$ is

$$T_x = \inf\{t > 0 : B(t) = x\}$$

where $\inf \emptyset = \infty$.

Theorem 1.6.1 $P_a(T_b < \infty) = 1$ *for all* a *and* b.

Proof of Theorem 1.6.1. Notice

$$\begin{aligned} P_a(T_b < \infty) =& \mathsf{P}(\inf\{t > 0 : B(t) = b\} < \infty | B(0) = a\} \\ =& \mathsf{P}(\inf\{t > 0 : B(t) - B(0) = b - a\} < \infty | B(0) = a\} \\ =& P_0(T_{b-a} < \infty) \end{aligned}$$

and

$$\begin{aligned} P_0(T_{b-a} < \infty) =& \mathsf{P}(\inf\{t > 0 : B(t) = b - a\} < \infty | B(0) = 0\} \\ =& \mathsf{P}(\inf\{t > 0 : -B(t) = a - b\} < \infty | - B(0) = 0\} \\ =& P_0(T_{a-b} < \infty) \end{aligned}$$

by the symmetry of Brownian motion. So, without loss of generality we assume $a = 0$, $b \geq 0$ and that the Brownian motion starts at 0. For $u > 0$, let

$$Z(t) = \exp\left\{uB(t) - \frac{u^2}{2}t\right\}.$$

Then $Z(t)$ is a martingale. By the theorem for stopped martingale, $\mathsf{E}Z(0) = \mathsf{E}Z(t \wedge T_b)$, that is

$$1 = \mathsf{E}\exp\left\{uB(t \wedge T_b) - \frac{u^2}{2}(t \wedge T_b)\right\}. \tag{1.6.1}$$

We will take $t \to \infty$ in (1.6.1). First, notice that the Brownian motion is always at or below level b for $t \leq T_b$ and so

$$0 \leq \exp\left\{uB(t \wedge T_b) - \frac{u^2}{2}(t \wedge T_b)\right\} \leq \exp\left\{uB(t \wedge T_b)\right\} \leq e^{ub},$$

that is, the random variables in (1.6.1) is bounded by e^{ub}. Next, on the event $\{T_b = \infty\}$,

$$\exp\left\{uB(t \wedge T_b) - \frac{u^2}{2}(t \wedge T_b)\right\} \le e^{ub} \exp\left\{-\frac{u^2}{2}t\right\} \to 0, \quad t \to \infty,$$

and on the event $\{T_b < \infty\}$,

$$\exp\left\{uB(t \wedge T_b) - \frac{u^2}{2}(t \wedge T_b)\right\} = \exp\left\{ub - \frac{u^2}{2}(t \wedge T_b)\right\} \to \exp\left\{ub - \frac{u^2}{2}T_b\right\}, \quad t \to \infty.$$

Taking the limits in (1.6.1) yields

$$\mathsf{E}\left[\mathbb{I}\{T_b < \infty\} \exp\left\{ub - \frac{u^2}{2}T_b\right\}\right] = 1 \tag{1.6.2}$$

by the Dominated Convergence Theorem. Again, random variables in (1.6.2) are bounded by 1. Taking $\mu \to 0$ yields

$$\mathsf{P}(T_b < \infty) = \mathsf{E}\left[\mathbb{I}\{T_b < \infty\}\right] = 1.$$

The proof is now completed. And also from (1.6.2), it also holds that

$$\mathsf{E}\left[\exp\left\{-\frac{u^2}{2}T_b\right\}\right] = e^{-ub}.$$

Replacing $u^2/2$ by α yields

$$\mathsf{E}e^{-\alpha T_b} = e^{-b\sqrt{2\alpha}}.$$

This is the Laplace transform of T_b.

Theorem 1.6.2 *For real number b, let the first passage time of Brownian motion $B(t)$ be T_b. Then the Laplace transform of the distribution of T_b is given by*

$$\mathsf{E}e^{-\alpha T_b} = e^{-|b|\sqrt{2\alpha}} \quad \text{for all } \alpha > 0, \tag{1.6.3}$$

and the density of T_b is

$$f_{T_b}(t) = \frac{|b|}{\sqrt{2\pi}} t^{-3/2} \exp\left\{-\frac{b^2}{2t}\right\}, \quad t \ge 0 \tag{1.6.4}$$

for $b \neq 0$.

Proof. For non-negative level b, (1.6.3) is proved. If b is negative, then because Brownian motion is symmetric, the first passage time T_b and $T_{|b|}$ have the same distribution. Equation (1.6.3) for negative b follows. (1.6.4) follows because

$$\int_0^\infty e^{-\alpha t} f_{T_b}(t)dt = e^{-|b|\sqrt{2\alpha}} \quad \text{for all } \alpha > 0.$$

Checking this equality is another hard work. We omit it.

Differentiation of (1.6.3) with respect to α results in

$$\mathsf{E}\left[T_b e^{-\alpha T_b}\right] = \frac{|b|}{\sqrt{2\alpha}} e^{-|b|\sqrt{2\alpha}} \quad \text{for all } \alpha > 0.$$

Letting $\alpha \downarrow 0$, we obtain $\mathsf{E}T_b = \infty$ so long as $b \neq 0$.

Corollary 1.6.1 *$\mathsf{E}T_b = \infty$ for any $b \neq 0$.*

The next result looks at the first passage time T_x as a process in x.

Theorem 1.6.3 *The process of the first passage times $\{T_x, x > 0\}$, has increments independent of the past, that is, for any $0 < a < b$, $T_b - T_a$ is independent of $B(t), t \leq T_a$, and the distribution of the increment $T_b - T_a$ is the same as that of T_{b-a}.*

Proof. By the continuity, the Brownian motion must pass a before it passes b. So, $T_a \leq T_b$. By the strong Markov property $\widehat{B}(t) = B(T_a + t) - B(T_a)$ is Bownian motion started at zero, and independent of the past $B(t), t \leq T_a$.

$$\begin{aligned} T_b - T_a &= \inf\{t \geq 0 : B(t) = b\} - T_a = \inf\{t \geq T_a : B(t) = b\} - T_a \\ &= \inf\{t \geq T_a : B(t - T_a + T_a) - B(T_a) = b - B(T_a)\} - T_a \\ &= \inf\{t \geq 0 : \widehat{B}(t) = b - a\}. \end{aligned}$$

Hence $T_b - T_a$ is the same as first hitting time of $b - a$ by $\widehat{B}$. The conclusion follows.

1.6.2 Maximum and Minimum

Let $B(t)$ be Brownian motion (which starts at zero). Define

$$M(t) = \max_{0 \leq s \leq t} B(s) \quad \text{and } m(t) = \min_{0 \leq s \leq t} B(s).$$

Theorem 1.6.4 *For any $x > 0$,*

$$\mathsf{P}(M(t) \geq x) = 2\mathsf{P}(B(t) \geq x) = 2\left(1 - \Phi(\frac{x}{\sqrt{t}})\right),$$

where $\Phi(x)$ stands for the standard normal distribution function.

Proof. The second equation is obvious. For the first one, notice

$$\begin{aligned}
&\mathsf{P}(M(t) \geq x) = \mathsf{P}(T_x \leq t)\\
=&\int_0^t \frac{x}{\sqrt{2\pi}} u^{-3/2} \exp\left\{-\frac{x^2}{2u}\right\} du\\
&(\text{letting } \frac{x^2}{u} = \frac{y^2}{t}, y > 0, \text{ then } u^{-3/2}du = -2du^{-1/2} = -\frac{1}{xt^{1/2}}dy)\\
=&\int_x^\infty \sqrt{\frac{2}{\pi t}} \exp\left\{-\frac{y^2}{2t}\right\} dy.
\end{aligned}$$

The proof is now completed.

Another proof. Observe that the event $\{M(t) \geq x\}$ and $\{T_x \leq t\}$ are the same. Since

$$\{B(t) \geq x\} \subset \{T_x \leq t\},$$

$$\mathsf{P}(B(t) \geq x) = \mathsf{P}(B(t) \geq x, T_x \leq t).$$

As $B(T_x) = x$,

$$\mathsf{P}(B(t) \geq x) = \mathsf{P}(T_x \leq t, B(T_x + (t - T_x)) - B(T_x) \geq 0).$$

Since T_x is a finite stopping time. By the strong Markov property, $\widehat{B}(s) = B(T_x + s) - B(T_x)$ is a Brownian motion which is independent of $\mathscr{F}_{T_x}$. So when $t \geq T_x$,

$$\mathsf{P}\big(\widehat{B}(t - T_x) \geq 0 | \mathscr{F}_{T_x}\big) = \frac{1}{2}.$$

So,

$$\begin{aligned}
\mathsf{P}(B(t) \geq x) =&\mathsf{E}\left[\mathbb{I}\{T_x \leq t\} \mathsf{P}\big(\widehat{B}(t - T_x) \geq 0 | \mathscr{F}_{T_x}\big)\right]\\
=&\mathsf{E}\left[\mathbb{I}\{T_x \leq t\}\frac{1}{2}\right]\\
=&\frac{1}{2}\mathsf{P}(T_x \leq t) = \frac{1}{2}\mathsf{P}(M(t) \geq x).
\end{aligned}$$

To find the distribution of the minimum of Brownian motion $m(t) = \min_{0 \leq s \leq t} B(s)$ we use the symmetry argument, and that

$$-\min_{0 \leq s \leq t} B(s) = \max_{0 \leq s \leq t}(-B(s)).$$

Notice $-B(t)$ is also a Brrownian motion which has the property as $B(t)$. It follows that for $x < 0$,

$$\begin{aligned}
\mathsf{P}(\min_{0 \leq s \leq t} B(s) \leq x) =&\mathsf{P}(\max_{0 \leq s \leq t}(-B(s)) \geq -x)\\
=&2\mathsf{P}(-B(t) \geq -x) = 2\mathsf{P}(B(t) \leq x).
\end{aligned}$$

Theorem 1.6.5 *For any $x < 0$,*

$$P(\min_{0\le s\le t} B(s) \le x) = 2P(B(t) \le x) = 2P(B(t) \ge -x).$$

1.6.3 Reflection principle and joint distribution

Theorem 1.6.6 *(Reflection principle) Let T be a stopping time. Define $\widehat{B}(t) = B(t)$ for $t \le T$, and $\widehat{B}(t) = 2B(T) - B(t)$ for $t \ge T$. Then $\widehat{B}$ is also Brownian motion.*

Note that $\widehat{B}$ defined above coincides with $B(t)$ for $t \le T$, and then for $t \ge T$ it is the reflected path about the horizontal line passing though $(T, B(T)$, that

$$\widehat{B}(t) - B(T) = -(B(t) - B(T)),$$

which gives the name to the result.

Corollary 1.6.2 *For any $y > 0$, let $\widehat{B}(t)$ be $B(t)$ reflected at T_y, that is, $\widehat{B}(t)$ equals $B(t)$ before $B(t)$ hits y, and is the reflection of $B(t)$ after the first hitting time. Then $\widehat{B}(t)$ is also a Brownian motion.*

Proof. Consider the process

$$Y(t) =: B(t)\ (0 \le t \le T),\ \ Z(t) = B(t+T) - B(T)\ (t \ge 0).$$

By the strong Markov property, Z is a Brownian motion independent of Y and T. So, $-Z$ is also a Brownian motion, also independent of Y and T. Thus $(Y, T, Z) \overset{d}{=} (Y, T, -Z)$. The map

$$\varphi : (Y, T, Z) \to \Big(Y(t)I\{t \le T\} + (Y(T) + Z(t-T))I\{t > T\}\Big)_{t\ge 0}$$

produces a continuous process, which will therefore have the same law as $\varphi(Y, T, -Z)$. But $\varphi(Y, T, -Z) = B$, $\varphi(Y, T, -Z) = \widehat{B}$. □

Theorem 1.6.7 *The joint distribution of $(B(t), M(t))$ has the density*

$$f_{B,M}(x, y) = \sqrt{\frac{2}{\pi}}\frac{2y - x}{t^{3/2}} \exp\left\{-\frac{(2y - x)^2}{2t}\right\}, \ \ \textit{for } y \ge 0, x \le y.$$

Proof. Let $y > 0$ and $y > x$. Let $\widehat{B}(t)$ be $B(t)$ reflected T_y. Then

$$\begin{aligned}
&\mathsf{P}(B(t) \le x, M(t) \ge y) = \mathsf{P}(B(t) \le x, T_y \le t)\\
=&\mathsf{P}(T_y \le t, \widehat{B}(t) \ge 2y - x) \ \ (\text{on } \{T_y \le t\}, \widehat{B}(t) = 2y - B(t)\ )\\
=&\mathsf{P}(T_y \le t, B(t) \ge 2y - x)\ ;\ (\text{since } T_y \text{ is the same for } \widehat{B} \text{ and } B\ )\\
=&\mathsf{P}(B(t) \ge 2y - x) \ \ (\text{since } y - x > 0, \text{ and } \{B(t) \ge 2y - x\} \subset \{T_y \le t\})\\
=&1 - \Phi(\frac{2y - x}{\sqrt{t}}).
\end{aligned}$$

That is

$$\int_{-\infty}^{x}\int_{y}^{\infty} f_{B,M}(u,v)dudv = 1 - \Phi(\frac{2y - x}{\sqrt{t}}).$$

The density is obtained by differentiation.

Corollary 1.6.3 *The conditional distribution of $M(t)$ given $B(t) = x$ is*

$$f_{M|B}(y|x) = \frac{2(2y - x)}{t} \exp\left\{-\frac{2y(y - x)}{t}\right\}, \quad y > 0, x \le y.$$

Proof. The density of $B(t)$ is

$$f_B(x) = \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{x^2}{2t}\right\}.$$

So for $y > 0$ and $x \le y$,

$$\begin{aligned}
f_{M|B}(y|x) =&\frac{f_{B,M}(x,y)}{f_B(x)}\\
=&\sqrt{\frac{2}{\pi}}\frac{2y - x}{t^{3/2}} \cdot \sqrt{2\pi t} \exp\left\{-\frac{(2y - x)^2}{2t} + \frac{x^2}{2t}\right\}\\
=&\frac{2(2y - x)}{t} \exp\left\{-\frac{2y(y - x)}{t}\right\}.
\end{aligned}$$

1.6.4 Arcsine law

A time point τ is called a zero of Brownian motion if $B(\tau) = 0$. Let $\{B^x(t)\}$ denotes Brownian motion started at x.

Theorem 1.6.8 *For any $x \ne 0$, the probability that $\{B^x(t)\}$ has at least one zero in the interval $(0,t)$, is give by*

$$\frac{|x|}{\sqrt{2\pi}}\int_0^t u^{-3/2} \exp\left\{-\frac{x^2}{2u}\right\} du = \sqrt{\frac{2}{\pi t}}\int_{|x|}^{\infty} \exp\left\{-\frac{y^2}{2t}\right\} dy,$$

that is the same probability of $P_0(T_{|x|} \le t)$.

Proof. If $x < 0$, then due to continuity of $B^x(t)$, the events $\{B^x$ has at least one zero between 0 and t $\}$ and $\{\max_{0\le s\le t} B^x(t) \ge 0\}$ are the same. Since $B^x(t) = B(t)+x$. So

$$\begin{aligned}
&\mathsf{P}(B^x \text{ has at least a zero between 0 and } t) = \mathsf{P}(\max_{0\le s\le t} B^x(t) \ge 0) \\
&= P_0(\max_{0\le s\le t} B(t) + x \ge 0) = P_0(\max_{0\le s\le t} B(t) \ge -x) \\
&= 2P_0(B(t) \ge -x) = P_0(T_{|x|} \le t).
\end{aligned}$$

If $x > 0$, then $-B^x(t)$ is a Brownian motion started at $-x$ by the symmetry of Brownian motion. The result follows.

Theorem 1.6.9 *The probability that Brownian motion $B(t)$ has no zeros in the time interval (a, b) is given by*

$$\frac{2}{\pi} \arcsin \sqrt{\frac{a}{b}}.$$

Proof. Denote by

$$h(x) = \mathsf{P}(B \text{ has at least one zero in } (a, b)|B_a = x).$$

By the Markov property $\mathsf{P}(B$ has at least one zero in $(a, b)|B_a = x)$ is the same as $\mathsf{P}(B^x$ has at least one zero in $(0, b-a))$. So

$$h(x) = \frac{|x|}{\sqrt{2\pi}} \int_0^{b-a} u^{-3/2} \exp\left\{-\frac{x^2}{2(b-a)}\right\} du.$$

By conditioning

$$\begin{aligned}
&\mathsf{P}(B \text{ has at least one zero in } (a, b)) \\
&= \int_{-\infty}^{\infty} \mathsf{P}(B \text{ has at least one zero in } (a, b)|B_a = x)\mathsf{P}(B_a \in dx) \\
&= \int_{-\infty}^{\infty} h(x)\mathsf{P}(B_a \in dx) \\
&= \sqrt{\frac{2}{\pi a}} \int_0^{\infty} h(x) \exp\left\{-\frac{x^2}{2a}\right\} dx.
\end{aligned}$$

Putting in the expression for $h(x)$ and performing the necessary calculations we obtain

$$\mathsf{P}(B \text{ has at least one zero in } (a, b)) = \frac{2}{\pi} \arccos \sqrt{\frac{a}{b}}.$$

The result follows.

Let

$$\gamma_t = \sup\{s \le t : B(s) = 0\} = \text{ last zero before } t.$$
$$\beta_t = \inf\{s \ge t : B(s) = 0\} = \text{ first zero after } t.$$

Note that β_t is a stopping time but γ_t is not.

Corollary 1.6.4

$$P(\gamma_t \le x) = \frac{2}{\pi} \arcsin \sqrt{\frac{x}{t}}, \quad x \le t,$$
$$P(\beta_t \ge y) = \frac{2}{\pi} \arcsin \sqrt{\frac{t}{y}}, \quad y \ge t,$$
$$P(\gamma_t \le x, \beta_t \ge y) = \frac{2}{\pi} \arcsin \sqrt{\frac{x}{y}}, \quad x \le t \le y.$$

Proof.

$$\{\gamma_t \le x\} = \{B \text{ has no zeros in } (x,t)\},$$
$$\{\beta_t \ge y\} = \{B \text{ has no zeros in } (t,y)\},$$
$$\{\gamma_t \le x, \beta_t \ge y\} = \{B \text{ has no zeros in } (x,y)\}.$$

Chapter 2

Brownian motion calculus

2.1 Introduction

Let $B(t)$ be a Brownian motion, together with a filtration $\mathscr{F}_t, t \geq 0$. Our goal to define stochastic integral

$$\int_0^T X(t)dB(t),$$

the integrand $X(t)$ can also be a stochastic process. The integral should be well-defined for at least all non-random continuous functions on $[0, T]$. When the integrand is random, we will assume that it is an adapted stochastic process.

Riemann-Stieltjes integral

Stieltjes integral of a function f with respect to a function g over interval $(a, b]$ is defined as

$$\int_a^b f dg = \int_a^b f(t)dg(t) = \lim_{\delta} \sum_{i=1}^{n} f(\xi_i^n)\big[g(t_i^n) - g(t_{i-1}^n)\big],$$

where $\{t_i^n\}$ represent partitions of the interval,

$$a = t_0^n < t_1^n < \ldots < t_n^n = b, \quad \delta = \max_{1 \leq i \leq n}(t_i^n - t_{i-1}^n), \quad \text{and } \xi_i^n \in [t_{i-1}^n, t_i^n],$$

the limit is taken over all partitions and all choice of ξ_i^n's with $\delta \to 0$.

When g is a function of finite variation, then any continuous function function f is stieltjes-integrable with respect to g.

Theorem 2.1.1 *Let* $\delta = \max_{1\le i\le n}(t_i^n - t_{i-1}^n)$ *denote the length of the largest interval in partition of* $[a,b]$. *If*

$$\lim_{\delta} \sum_{i=1}^{n} f(\xi_i^n)\big[g(t_i^n) - g(t_{i-1}^n)\big]$$

exists for any continuous function f *then* g *must be of finite variation on* $[a,b]$.

This shows that if g (for example, the Brownian motion B) has infinite variation then the limit of the approximating sums may not exist even when the integrand function f is continuous. Therefore integrals with respect to functions of infinite variation (stochastic integrals) must be defined in another way.

2.2 Definition of Itô integral

2.2.1 Itô's Integral for simple integrands

The integral integral $\int_0^T X(t)dB(t)$ should have the properties:

- If $X(t) = 1$, then
$$\int_0^T X(t)dB(t) = B(T) - B(0);$$
- If $X(t) = c$ in $(a,b] \subset [0,T]$ and zero otherwise, then
$$\int_0^T X(t)dB(t) = c(B(b) - B(a));$$
- For real α and β,
$$\int_0^T \big(\alpha X(t) + \beta Y(t)\big)dB(t) = \alpha\int_0^T X(t)dB(t) + \beta\int_0^T Y(t)dB(t).$$

Deterministic simple process

Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0,T]$; i.e.,

$$0 = t_0 < t_1 < \ldots < t_n = T.$$

Assume that $X(t)$ is a non-random constant in t on each subinterval $(t_i, t_{i+1}]$, i.e.,

$$X(t) = \begin{cases} c_0 & \text{if } t = 0, \\ c_i & \text{if } t_i < t \le t_{i+1}, i = 0, \ldots, n-1. \end{cases}, \qquad (2.2.1)$$

or in one formula,

$$X(t) = c_0 I_0(t) + \sum_{i=0}^{n-1} c_i I_{(t_i,t_{i+1}]}(t).$$

Such a process is a deterministic simple process. The Itô integral $\int_0^T X(t)dB(t)$ is defined as a sum

$$\int_0^T X(t)dB(t) = \sum_{i=0}^{n-1} c_i\big(B(t_{i+1}) - B(t_i)\big).$$

It is easy to see that the integral is a Gaussian random variable with mean zero and variance

$$\begin{aligned}\mathsf{Var}\Big(\int XdB\Big) &= \sum_{i=0}^{n-1} c_i^2 \mathsf{Var}\big(B(t_{i+1}) - B(t_i)\big)\\ &= \sum_{i=0}^{n-1} c_i^2 (t_{i+1} - t_i) = \int_0^T X^2(t)dt.\end{aligned}$$

Simple process

Since we would like to integrate random process, it is important to allow constants c_i in (2.2.1) to be random. If c_i's are replaced by random variables ξ_i's, then in order to carry out calculations, and have convenient properties of the integral, the random variable ξ_i are allowed to depend on the values of $B(t)$ for $t \le t_i$, that is, they are allowed to be $\mathscr{F}_{t_i}$-measurable, and independent of the future increments of the Brownian motion. In finance, $X(t)$ is regraded as an asset at time t, it depends on the information of the past. So, $X(t)$ is assumed to be adapted at least. So, ξ_i is $\mathscr{F}_{t_i}$- measurable.

A simple process is

$$X(t) = \xi_0 I_0(t) + \sum_{i=0}^{n-1} \xi_i I_{(t_i,t_{i+1}]}(t), \text{ with } \xi_i \text{ is } \mathscr{F}_{t_i} - \text{measurable.}$$

For a simple process, the Itô integral $\int_0^T XdB$ is defined as a sum

$$\int_0^T X(t)dB(t) = \sum_{i=0}^{n-1} \xi_i\big(B(t_{i+1}) - B(t_i)\big). \tag{2.2.2}$$

In general, for $0 \le t \le T$, if $t_k < t \le t_{k+1}$, then

$$\int_0^t X(u)dB(u) = \sum_{j=0}^{k-1} \xi_j\big(B(t_{j+1}) - B(t_j)\big) + \xi_k\big(B(t) - B(t_k)\big)$$
$$= \sum_j \xi_j\big(B(t \wedge t_{j+1}) - B(t \wedge t_j)\big).$$

Properties of Itô integral of simple processes

(i) Linearity. If $X(t)$ and $Y(t)$ are simple processes and α and β are constant, then

$$\int_0^T \big(\alpha X(t) + \beta Y(t)\big)dB(t) = \alpha \int_0^T X(t)dB(t) + \beta \int_0^T Y(t)dB(t).$$

(ii) For all $[a, b] \subset [0, T]$,

$$\int_0^T I_{[a,b]}(t)dB(t) = B(b) - B(a).$$

The next properties hold if ξ_i's are square integrable, $\mathsf{E}\xi_i^2 < \infty$.

(iii) Zero mean property.

$$\mathsf{E}\int_0^T X(t)dB(t) = 0.$$

(iv) Isometry property.

$$\mathsf{E}\left(\int_0^T X(t)dB(t)\right)^2 = \int_0^T \mathsf{E}X^2(t)dt.$$

(v) Martingale property. Let $I(t) = \int_0^t X(u)dB(u)$. Then $I(t)$ is a continuous martingale.

(vi) The quadratic variation accumulated up to time t by the Itô integral is

$$[I, I](t) = \int_0^t X^2(u)du.$$

(vii) $I^2(t) - \int_0^t X^2(u)du$ is a martingale, or equivalently,

$$\mathsf{E}\left[\left(\int_s^t X(u)dB(u)\right)^2 \Big| \mathscr{F}_s\right] = \int_s^t \mathsf{E}\left[X^2(u)\big|\mathscr{F}_s\right] du, \quad \forall\, s \le t.$$

Proof. Property (i) and (ii) are easy to be verified directly from the definition. Property (iii) follows from Property (v).

Property (iv). Write $D_i = B(t_{i+1}) - B(t_i))$. Then

$$(I(T))^2 = (\sum_{i=0}^{n} \xi_i D_i)^2 = \sum_{i=0}^{n-1} \xi_i^2 D_i^2 + 2\sum_{j<i} \xi_i \xi_j D_i D_j.$$

Notice that ξ_i is $\mathscr{F}_{t_i}$-measurable and $D_i = B(t_{i+1}) - B(t_i)$ is independent of $\mathscr{F}_{t_i}$. So

$$\mathsf{E}[\xi_i^2 D_i^2 | \mathscr{F}_{t_i}] = \xi_i^2 \mathsf{E}[D_i^2] = \xi_i^2 (t_{i+1} - t_i).$$

It follows that

$$\mathsf{E}[[\xi_i^2 D_i^2] = \mathsf{E}\big[\mathsf{E}[\xi_i^2 D_i^2 | \mathscr{F}_{t_i}]\big] = \mathsf{E}[\xi_i^2](t_{i+1} - t_i).$$

Also, for $j < i$, ξ_j, D_j are $\mathscr{F}_{t_i}$-measurable. So

$$\mathsf{E}[\xi_i \xi_j D_i D_j | \mathscr{F}_{t_i}] = \xi_j D_j \xi_i \mathsf{E}[D_i] = 0.$$

It follows that

$$\mathsf{E}[\xi_i \xi_j D_i D_j] = 0.$$

Hence

$$\mathsf{E}(I(T))^2 = \sum_{i=0}^{n-1} \mathsf{E}[\xi_i^2](t_{i+1} - t_i) = \int_0^T \mathsf{E}(X^2(t))dt.$$

Property (v). For $0 \le s \le t \le T$, we want to show that $\mathsf{E}[I(t)|\mathscr{F}_s] = I(s)$. It suffices to show that $\mathsf{E}[I(T)|\mathscr{F}_s] = I(s)$, which implies

$$\mathsf{E}[I(t)|\mathscr{F}_s] = \mathsf{E}\left[\mathsf{E}[I(T)|\mathscr{F}_t]\middle|\mathscr{F}_s\right] = \mathsf{E}[I(T)|\mathscr{F}_s] = I(s).$$

Assume $s \in (t_k, t_{k+1}]$, write $I(T)$ as

$$I(T) = \sum_{i=0}^{k-1} \xi_i D_i + \xi_k [B(t_{k+1}) - B(t_k)] + \sum_{i=k+1}^{n-1} \xi_i D_i.$$

For $i \le k-1$, $t_{i+1} \le s$, ξ_i and D_i are $\mathscr{F}_{t_{i+1}}$-measurable, and so $\mathscr{F}_s$-measurable. It follows that

$$\mathsf{E}\left[\sum_{i=0}^{k-1} \xi_i D_i \middle| \mathscr{F}_s\right] = \sum_{i=0}^{k-1} \xi_i D_i.$$

For $i \ge k+1$, $\mathscr{F}_{t_i} \supset \mathscr{F}_s$ and $D_i = B(t_{i+1}) - B(t_i)$ is independent of $\mathscr{F}_{t_i}$. So

$$\mathsf{E}\left[\xi_i D_i \middle| \mathscr{F}_s\right] = \mathsf{E}\left[\mathsf{E}\left[\xi_i D_i | \mathscr{F}_{t_i}\right] | \mathscr{F}_s\right] = \mathsf{E}\left[\xi_i \mathsf{E}[D_i] \middle| \mathscr{F}_s\right] = 0.$$

Finally,

$$\mathsf{E}\left[\xi_k[B(t_{k+1}) - B(t_k)]|\mathscr{F}_s\right] = \xi_k[\mathsf{E}\left[B(t_{k+1})|\mathscr{F}_s\right] - B(t_k)] = \xi_k[B(s) - B(t_k)].$$

It follows that

$$\mathsf{E}[I(T)|\mathscr{F}_s] = \sum_{i=0}^{k-1} \xi_i D_i + \xi_k[B(s) - B(t_k)] = I(s).$$

<u>Property (vi)</u>. Assume $t_k \le t \le t_{k+1}$. Notice,

$$I(s) = \sum_{j=0}^{i-1} \xi_j \big(B(t_{j+1}) - B(t_j)\big) + \xi_i \big(B(s) - B(t_i)\big), \quad \text{for } t_i < s \le t_{j+1}.$$

It follows that

$$[I, I]([t_i, t_{i+1}]) = \xi_i^2 \cdot [B, B]([t_i, t_{i+1}]) = \xi_i^2 (t_{i+1} - t_i) = \int_{t_i}^{t_{i+1}} X^2(u) du.$$

Similarly,

$$[I, I]([t_k, t]) = \xi_k^2 \cdot [B, B]([t_k, t]) = \xi_k^2 (t - t_k) = \int_{t_k}^{t} X^2(u) du.$$

Adding up all these pieces, we obtain

$$[I, I](t) = [I, I]([0, t]) = \int_0^t X^2(u) du.$$

<u>Property (vii)</u>. Let $M(t) = I^2(t) - [I, I](t)$. It suffices to show that $\mathsf{E}[M(T)|\mathscr{F}_s] = M(s)$. Assume $s \in (t_k, t_{k+1}]$. Then

$$\int_s^T X(u) dB(u) = \xi_k \big[B(t_{k+1}) - B(s)\big] + \sum_{j=k+1}^{n-1} \xi_j \big[B(t_{j+1}) - B(t_j)\big].$$

Following the same lines in the proof of Property 4, we have

$$\begin{aligned} &\mathsf{E}\left[\left(\int_s^T X(u) dB(u)\right)^2 \Big| \mathscr{F}_s\right] \\ =&\mathsf{E}[\xi_k^2|\mathscr{F}_s](t_{k+1} - s) + \sum_{j=k+1}^{n-1} \mathsf{E}[\xi_j^2|\mathscr{F}_s](t_{j+1} - t_j) \\ =& \int_s^T \mathsf{E}\big[X^2(u)|\mathscr{F}_s\big] du. \end{aligned}$$

Hence

$$\begin{aligned}\mathsf{E}\left[I^2(T)\middle|\mathscr{F}_s\right] =& I^2(s) + 2I(s)\mathsf{E}\left[I(T)\middle|\mathscr{F}_s\right] + \mathsf{E}\left[\left(\int_s^T X(u)du\right)^2 \middle|\mathscr{F}_s\right]\\ =& I^2(s) + \int_s^T \mathsf{E}\big[X^2(u)|\mathscr{F}_s\big]du.\end{aligned}$$

On the other hand, it is obvious that

$$\mathsf{E}\left[[I,I](T)\middle|\mathscr{F}_s\right] = [I,I](s) + \int_s^T \mathsf{E}\big[X^2(u)|\mathscr{F}_s\big]du.$$

It follows that

$$\mathsf{E}\left[I^2(T) - [I,I](T)\middle|\mathscr{F}_s\right] = I^2(s) - [I,I](s), \quad \text{that is } \mathsf{E}[M(T)|\mathscr{F}_s] = M(s).$$

2.2.2 Itô's Integral for general integrands

Let $X^n(t)$ be a sequence of simple processes convergent in probability to the process $X(t)$. Then, under some conditions, the sequence of their integrals $\int_0^T X^n(t)dB(t)$ also convergence in probability. That limit is taken to be the integral $\int_0^T X(t)dB(t)$.

Example 2.2.1 *Find* $\int_0^T B(t)dB(t)$.

Solution. Let $0 = t_0^n < t_1^n < \ldots < t_n^n = T$ be a partition of $[0,T]$, and let

$$X^n(t) = \sum_{i=0}^{n-1} B(t_i^n) I_{(t_i^n, t_{i+1}^n]}(t).$$

Then, for each n, X^n is a simple process. Take the sequence of partitions such that $\delta_n = \max_i(t_{i+1} - t_i) \to 0$. Then $X^n(t) \to B(t)$ almost surely, be the continuity of the Brownian paths. We will show that $\int_0^T X^n(t)dB(t)$ will converges in probability.

$$\int_0^T X^n(t)dB(t) = \sum_{i=0}^{n-1} B(t_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big).$$

Observer

$$B(t_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big) = \frac{1}{2}\left(B^2(t_{i+1}^n) - B^2(t_i^n) - \big(B(t_{i+1}^n) - B(t_i^n)\big)^2\right),$$

and then

$$\begin{aligned}\int_0^T X^n(t)dB(t) =& \sum_{i=0}^{n-1}\frac{1}{2}\big(B^2(t_{i+1}^n) - B^2(t_i)\big) - \frac{1}{2}\sum_{i=0}^{n-1}\big(B(t_{i+1}^n) - B(t_i^n)\big)^2\\ =& \frac{1}{2}B^2(T) - \frac{1}{2}B^2(0) - \frac{1}{2}\sum_{i=0}^{n-1}\big(B(t_{i+1}^n) - B(t_i^n)\big)^2.\end{aligned}$$

The last summation converges to the quadratic variation T of Brownian motion on $[0, T]$ in probability. So, $\int_0^T X^n(t)dB(t)$ converges in probability, and the limit is

$$\int_0^T B(t)dB(t) = \lim \int_0^T X^n(t)dB(t) = \frac{1}{2}B^2(T) - \frac{1}{2}T.$$

Itô's Integral for square-integrable adapted processes

Theorem 2.2.1 *Let $X(t)$ be an adapted process such that*

$$\int_0^T \mathsf{E}[X^2(t)]dt < \infty. \tag{2.2.3}$$

Then Itô integral $\int_0^T X(t)dB(t)$ is defined and satisfied Properties (i)-(vii).

Proof. For an adapted process $X(t)$ there exists a sequence of simple process $X^n(t)$ such that

$$\lim_{n\to\infty} \mathsf{E} \int_0^T |(X(t) - X^n(t)|^2 dt = 0.$$

Indeed, for a continuous process X stratifying (2.2.3), such $X^n(t)$ can be taken as

$$X(0) + \sum_{k=0}^{2^n-1} X(\frac{k}{2^n}T) I_{(kT/2^n,(k+1)T/2^n]}(t).$$

If X is not continuous, the construction of approximating processes is more involved (see the lemma below). The Itô integrals for simple processes $X^n(t)$ is defined by (2.2.2). By the Isometry property

$$\mathsf{E}\left[\int_0^T X^n(t)dB(t) - \int_0^T X^m(t)dB(t)\right]^2 = \int_0^T \mathsf{E}[X^n(t) - X^m(t)]^2 dt \to 0 \;\; n, m \to \infty.$$

That is, the Itô integrals for simple processes $X^n(t)$ form a Cauchy sequence in L_2. By the completeness of L_2, $\int_0^T X^n(t)dB(t)$ converges to a limit in L_2. We denote the limit of $\int_0^T X^n(t)dB(t)$ by Z. If $Y^n(t)$ is also a simple processes satisfying

$$\lim_{n\to\infty} \mathsf{E} \int_0^T |(X(t) - Y^n(t)|^2 dt = 0.$$

Then

$$\lim_{n\to\infty} \mathsf{E} \int_0^T |(X^n(t) - Y^n(t)|^2 dt = 0.$$

Using the Isometry property,

$$\mathsf{E}\left[\int_0^T Y^n(t)dB(t) - \int_0^T X^n(t)dB(t)\right]^2 = \int_0^T \mathsf{E}[Y^n(t) - X^n(t)]^2 dt \to 0 \;\; n, m \to \infty.$$

So the limit $\int_0^T X^n(t)dB(t)$ does not depend on the choice of the approximating sequence. The limit of $\int_0^T X^n(t)dB(t)$ is called $\int_0^T X(t)dB(t)$.

Properties (i)-(v) and (vii) are easy to be verified by taking limits since the Itô integral of the simple processes satisfies Property (i)-(vii).

For Property (vi), let $\Pi = \{s_0, s_1, \ldots, s_m\}$ be a partition of $[0,T]$, and $\|\Pi\|$ be the largest length of the subinterval. Denote $Q(\Pi)$ and $Q_n(\Pi)$ be the sample quadratic variation of $I(t) =: \int_0^t X(u)dB(u)$ and $I_n(t) =: \int_0^t X^n(u)dB(u)$ over $[0,T]$ respectively. Notice $[I_n, I_n](T) = \int_0^T (X^n(u))^2 du \to \int_0^T (X(u))^2 du$ in probability. It suffices to show that

$$\sup_{\Pi} \mathsf{E}|Q_n(\Pi) - Q(\Pi)| \to 0 \text{ in probability as } n \to \infty.$$

Using the inequality $b^2 - a^2 = (b-a)^2 + 2a(b-a) \le (b-a)^2 + \epsilon^{-2}(b-a)^2 + \epsilon^2 a^2$, we obtain

$$\begin{aligned}
&\mathsf{E}\left|\left(\int_{s_i}^{s_{i+1}} X^n(u)dB(u)\right)^2 - \left(\int_{s_i}^{s_{i+1}} X(u)dB(u)\right)^2\right| \\
\le&(1+\epsilon^{-2})\mathsf{E}\left(\int_{s_i}^{s_{i+1}} (X^n(u) - X(u))dB(u)\right)^2 + \epsilon^2 \mathsf{E}\left(\int_{s_i}^{s_{i+1}} X(u)dB(u)\right)^2 \\
\le&(1+\epsilon^{-2})\int_{s_i}^{s_{i+1}} \mathsf{E}\left[(X^n(u) - X(u))^2\right] du + \epsilon^2 \int_{s_i}^{s_{i+1}} \mathsf{E}\left[X^2(u)\right] du.
\end{aligned}$$

Adding up all these pieces, we obtain

$$\mathsf{E}||Q_n(\Pi) - Q(\Pi)| \le (1+\epsilon^{-2})\int_0^T \mathsf{E}\left[(X^n(u) - X(u))^2\right] du + \epsilon^2 \int_0^T \mathsf{E}\left[X^2(u)\right] du.$$

The proof is now completed.

Lemma 2.2.1 *Let $X(t)$ be an adapted process such that*

$$\int_0^T E[X^2(t)]dt < \infty. \tag{2.2.4}$$

Then there exists a sequence of simple process $\{X_n(t)\}$ such that

$$\int_0^T E\left[\left(X_n(t) - X(t)\right)^2\right] dt \to 0. \tag{2.2.5}$$

Let $\mathscr{L}_2$ be a class of of adapted processes with (2.2.4). We complete the proof via three steps.

Step 1. Let $X \in \mathscr{L}_2$ be bounded and $X(\cdot, \omega)$ continuous for each ω. Then there exists a sequence of simple process $\{X_n(t)\}$ such that

$$\int_0^T \mathsf{E}\left[\left(X_n(t) - X(t)\right)^2\right] dt \to 0.$$

Proof Take

$$X_n(t) = X(0) + \sum_{k=0}^{2^n-1} X(\frac{k}{2^n}T) I_{(kT/2^n,(k+1)T/2^n]}(t).$$

Step 2. Let $X \in \mathscr{L}_2$ be bounded. Then there exist bounded functions $Y_n \in \mathscr{L}_2$ such that $Y_n(\cdot, \omega)$ is continuous for all ω and n, and

$$\int_0^T \mathsf{E}\left[\left(Y_n(t) - X(t)\right)^2\right] dt \to 0.$$

Proof. Suppose $|X(t, \omega)| \leq M$. For each n, let ψ_n be a non-negative continuous real function such that

1. $\psi_n(x) = 0$ for $x \leq -\frac{1}{n}$ and $x \geq 0$,

2. $\int_\infty^\infty \psi_n(x)dx = 1$.

Define

$$Y_n(t, \omega) = \int_0^t \psi_n(s - t) X(s, \omega) ds.$$

Then $Y_n(\cdot, \omega)$ is continuous for all ω and $|Y_n(t, \omega)| \leq M$. Since $X \in \mathscr{L}_2$ we see that $Y_n(t, \omega)$ is $\mathscr{F}_t$-measurable for all t. Moreover,

$$\int_0^T \left(Y_n(t, \omega) - X(t, \omega)\right)^2 dt \to 0, \quad \text{for each } \omega,$$

since $\{\psi_n\}$ constitutes an approximate identity.

So

$$\int_0^T \mathsf{E}\left[\left(Y_n(t) - X(t)\right)^2\right] dt \to 0$$

by bounded convergence.

Step 3. Let $X \in \mathscr{L}_2$. Then there exists a sequence $\{Y_n\} \subset \mathscr{L}_2$ such that Y_n is bounded for each n, and

$$\int_0^T \mathsf{E}\left[\left(Y_n(t) - X(t)\right)^2\right] dt \to 0.$$

Proof. Put

$$Y_n(t,\omega) = \begin{cases} -n & \text{if } X(t,\omega) < -n \\ X(t,\omega) & \text{if } -n \le X(t,\omega) \le n \\ n & \text{if } X(t,\omega) > -n. \end{cases}$$

The the conclusion follows by dominated convergence.

Further properties of Itô's integral

Let $X(t), Y(t) \in \mathscr{L}_2$, and σ, τ be stopping times such that $\tau \le \sigma$. Then for any $0 \le t \le T$,

(viii)

$$\mathsf{E}\left[\int_{t\wedge\tau}^{t\wedge\sigma} X(u)dB(u)\Big|\mathscr{F}_\tau\right] = 0,$$

$$\mathsf{E}\left[\left(\int_{t\wedge\tau}^{t\wedge\sigma} X(u)dB(u)\right)^2\Big|\mathscr{F}_\tau\right] = \int_{t\wedge\tau}^{t\wedge\sigma} \mathsf{E}\left[X^2(u)\big|\mathscr{F}_\tau\right] du.$$

(iv) For all $s \le t$,

$$\mathsf{E}\left[\left(\int_s^t X(u)dB(u)\right)\left(\int_s^t Y(u)dB(u)\right)\Big|\mathscr{F}_s\right] = \int_s^t \mathsf{E}\left[X(u)Y(u))\big|\mathscr{F}_s\right] du.$$

$$\mathsf{E}\left[\left(\int_{t\wedge\tau}^{t\wedge\sigma} X(u)dB(u)\right)\left(\int_{t\wedge\tau}^{t\wedge\sigma} Y(u)dB(u)\right)\Big|\mathscr{F}_\tau\right] = \int_{t\wedge\tau}^{t\wedge\sigma} \mathsf{E}\left[X(u)Y(u)\big|\mathscr{F}_\tau\right] du.$$

(iiv) If σ is a stopping time, then

$$\int_0^{t\wedge\sigma} X(u)dB(u) = \int_0^t X'(u)dB(u), \ \ \forall \ \ t \ge 0,$$

where $X'(t,\omega) = X(t,\omega)I_{\{\sigma(\omega)\ge t\}}$.

Proof. Write $I_X(t) = \int_0^t X(t)dB(t)$, $M_X(t) = I_X^2(t) - [I_X, I_X](t)$. (viii) is equivalent to

$$\mathsf{E}\left[I_X(t\wedge\sigma)\big|\mathscr{F}_{t\wedge\tau}\right] = I_X(t\wedge\tau),$$

$$\mathsf{E}\left[M_X(t\wedge\sigma)\big|\mathscr{F}_{t\wedge\tau}\right] = M_X(t\wedge\tau).$$

Note $I_X(t)$ and $M_X(t)$ are both martingale. The results follows from the optional stopping time theorem (see Lemma 1.5.1 (ii)).

Write $M_{X,Y}(t) = I_X(t)I_Y(t) - [I_X, I_Y](t)$. Note $\mathsf{E}[I_X(t) - I_X(s)|\mathscr{F}_s] = 0$, $\mathsf{E}[I_Y(t) - I_Y(s)|\mathscr{F}_s] = 0$. The first property in (iv) is equivalent to

$$\mathsf{E}\left[M_{X,Y}(t)\middle|\mathscr{F}_s\right] = M_{X,Y}(s),$$

which is due to the fact that

$$\begin{aligned}2M_{X,Y}(t) =& I^2_{X+Y}(t) - [I_{X+Y}, I_{X+Y}](t)\\ &- \left(I^2_X(t) - [I_X, I_X](t)\right) - \left(I^2_Y(t) - [I_Y, I_Y](t)\right)\end{aligned}$$

is a martingale.

By the optional stopping time theorem (see Lemma 1.5.1 (ii)), we have

$$\mathsf{E}\left[M_{X,Y}(t\wedge\sigma)\middle|\mathscr{F}_{s\wedge\tau}\right] = M_{X,Y}(s\wedge\tau),$$

which is equivalent to the second property in (iv).

If Property (iiv) is checked for simple processes and then the general case can be proved by approximating X with simple processes. The following proof can be refered to N. Ikeda and Watanabe (1889), *Stochastic Differential Equations and Diffusion Process*, page 50, North-Holland Publishing Company.

Suppose X is a simple process,

$$X(t) = \xi_0 I_0(t) + \sum_i \xi_i I_{(t_i,t_{i+1}]}(t).$$

Let $\{s^{(n)}_j\}$ be a refinement of subdivisions $\{t_i\}$. Suppose X has the expression

$$X(t,\omega) = \xi_0 I_0(t) + \sum_j \xi^{(n)}_j I_{(s^{(n)}_j, s^{(n)}_{j+1}]}(t).$$

Define

$$\sigma^n(\omega) = s^{(n)}_{j+1} \quad \text{if} \quad \sigma(\omega) \in (s^{(n)}_j, s^{(n)}_{j+1}].$$

It is easy to see that σ^n is an $\mathscr{F}_t$-stopping time for each $n = 1, 2, \ldots$ and $\sigma^n \downarrow \sigma$ as $n \to \infty$. If $s \in (s^{(n)}_j, s^{(n)}_{j+1}]$, then $I_{\{\sigma^n \geq s\}} = I_{\{\sigma > s^{(n)}_j\}}$ and therefore, if we set $X'_n(s,\omega) = X(s,\omega)I_{\{\sigma^n(\omega)\geq s\}}$, then

$$\begin{aligned}X'_n(s,\omega) =& \xi_0 I_0(s) + \sum_j \xi^{(n)}_j I_{\{\sigma^n \geq s\}} I_{(s^{(n)}_j, s^{(n)}_{j+1}]}(s)\\ =& \xi_0 I_0(s) + \sum_j \xi^{(n)}_j I_{\{\sigma > s^{(n)}_j\}} I_{(s^{(n)}_j, s^{(n)}_{j+1}]}(s)\end{aligned}$$

is also a simple process. Clearly, for every $t > 0$,

$$\begin{aligned}\mathsf{E}\left[I(X'_n)(t) - I(X')(t)\right]^2 &= \mathsf{E}\left[\int_0^t \left(X'_n(u) - X'(u)\right)^2 du\right] \\ &= \mathsf{E}\left[\int_0^t X^2(u) I_{\{\sigma^n \geq u > \sigma\}} du\right] \to 0,\end{aligned}$$

as $n \to \infty$. Hence $I(X'_n)(t) \to I(X')(t)$ in probability. Also

$$\begin{aligned}I(X'_n)(t) &= \sum_j \xi_j^{(n)} I_{\{\sigma > s_j^{(n)}\}} [B(t \wedge s_{j+1}^{(n)}) - B(t \wedge s_j^{(n)})] \\ &= \sum_j \xi_j^{(n)} I_{\{\sigma > s_j^{(n)}\}} [B(t \wedge \sigma^n \wedge s_{j+1}^{(n)}) - B(t \wedge \sigma^n \wedge s_j^{(n)})] \\ &\quad \big(\text{ since } \sigma^n \geq s_{j+1}^{(n)} \text{ if } \sigma > s_j^{(n)}\big) \\ &= \sum_j \xi_j^{(n)} [B(t \wedge \sigma^n \wedge s_{j+1}^{(n)}) - B(t \wedge \sigma^n \wedge s_j^{(n)})] \\ &\quad \big(\text{ since } \sigma^n \leq s_j^{(n)} \text{ if } \sigma \leq s_j^{(n)} \\ &\quad\quad \text{and so } [B(t \wedge \sigma^n \wedge s_{j+1}^{(n)}) - B(t \wedge \sigma^n \wedge s_j^{(n)})] = 0\big) \\ &= \int_0^{t \wedge \sigma^n} X(u) dB(u) = I(X)(t \wedge \sigma_n).\end{aligned}$$

Consequently,

$$I(X')(t) = \lim I(X'_n)(t) = \lim I(X)(t \wedge \sigma^n) = I(X)(t \wedge \sigma),$$

by the continuity of $I(X)(t)$.

Itô's Integral for predictable processes

The approach of defining of the Itô integral by approximation can be carried out for the class of predictable process $X(t)$, $0 \leq t \leq T$, satisfying the condition

$$\int_0^T X^2(t) dt < \infty \quad a.s. \tag{2.2.6}$$

Recall that a process X is adapted if for any t, the value $X(t)$ can depend on the past values of $B(s)$ for $s \leq t$, but not on future values of $B(s)$ for $s > t$.

Intuitively, an adapted process is predictable if for any t, the value $X(t)$ is determined by the values of values of $B(s)$ for $s < t$. X is predictable if it is

1. a left-continuous adaptive process, for example, a simple function, that is an adapted left-continues step function,

2. a limit (almost sure, in probability) of left-continuous adapted processes,

3. a Borel-measurable function of a predictable process.

For a predictable process X satisfying (2.2.6), one can choose simple processes X^n convergent to X in sense that

$$\int_0^T (X^n(t) - X(t))^2 dt \to 0 \text{ in probability.}$$

The sequence of Itô integrals $\int_0^T X^n(t)dB(t)$ is a Cauchy sequence in probability. It converges to a limit $\int_0^T X(t)dB(t)$. This limit does not depend on the choice of the sequence of simple processes. It is defined as the integral of X.

Theorem 2.2.2 *The Itô integral of a predicable process X satisfying the condition (2.2.6) exists and it is a limit in probability of a sequence of corresponding Itô integrals of simple processes.*

Further, the integral $I(t) = \int_0^t X(u)dB(u)$ is a locally square integrable martingale, i.e., there exists a sequence of stopping times σ_n such that $\sigma_n < \infty$, $\sigma_n \uparrow \infty$ and $I_n(t) =: I(t \wedge \sigma_n)$ is a square integrable martingale for each n.

Proof. Let $\sigma_n(\omega) = \inf\{t : \int_0^t X^2(u,\omega)du \geq n\} \wedge n$. Then σ_n is a sequence of stopping times such that $\sigma_n \uparrow \infty$ a.s.. Set $X_n(t,\omega) = X(t,\omega)I_{\{\sigma_n(\omega)\geq t\}}$. Clearly

$$\int_0^\infty X_n^2(t,\omega)dt = \int_0^{\sigma_n} X^2(t,\omega)dt \leq n,$$

and hence

$$\int_0^\infty \mathsf{E}[X_n^2(t)]dt \leq n, \quad n = 1, 2, \ldots.$$

So, $I(X_n)(t) =: \int_0^t X_n(u)dB(u)$ is well defined and a square integrable martingale. Notice $X_m(t,\omega) = X_n(t,\omega)I_{\{\sigma_m(\omega)\geq t\}}$ for $m < n$, by Property (iiv), we have for each $m < n$ that

$$I(X_n)(t \wedge \sigma_m) = I(X_m)(t).$$

Consequently if we define $I(X)(t) = \int_0^t X(u)dB(u)$ by

$$I(X)(t) = I(X_m)(t) \text{ for } t \leq \sigma_n,$$

then this definition is well-defined and determines a continuous process $I(X)$ such that

$$I(X)(t \wedge \sigma_n) = I(X_n)(t), \quad n = 1, 2, \ldots.$$

Therefore $I(X)$ is a local square integrable martingale.

Further, for each n, $I(X_n)$ is a limit in probability of integrals of a sequence of simple processes, and so is $I(X)$.

Finally, it can be shown that if

$$\int_0^T (Y_n(t) - X(t))^2 dt \to 0 \text{ in probability,}$$

and Y_n is a simple process, then

$$I(Y_n)(t) \to I(Y)(t) \text{ in probability.}$$

The proof is similar to the following theorem and omitted here.

Theorem 2.2.3 *Suppose predicable processes X_n and X satisfy the condition (2.2.6) and*

$$\int^T (X_n(t) - X(t))^2 dt \to 0 \quad \textit{in probability.}$$

Then

$$\sup_{0 \le t \le T} \left| \int_0^t X_n(t) dB(t) - \int_0^t X(t) dB(t) \right| \to 0 \quad \textit{in probability.}$$

Proof. The result follows from the following inequality immediately.

$$\mathsf{P}\left(\sup_{0 \le t \le T} \left| \int_0^t X(t) dB(t) \right| \ge \epsilon \right) \le \mathsf{P}\left(\int_0^T X^2(t) dt \ge \delta \right) + \frac{\delta}{\epsilon^2}. \tag{2.2.7}$$

For (2.2.7), we let

$$X_\delta(s) = X(s) I\{s : \int_0^s X^2(u) du < \delta\} \in \mathscr{L}_2.$$

Then $\int_0^t X_\delta(s) dB(s)$ is a square integrable martingale. According to the Doob inequality,

$$\mathsf{P}\left(\sup_{0 \le t \le T} \left| \int_0^t X_\delta(t) dB(t) \right| \ge \epsilon \right) \le \frac{\mathsf{E}\left(\int_0^T X_\delta(t) dB(t) \right)^2}{\epsilon^2}$$

$$\le \frac{\mathsf{E}\left(\int_0^T X_\delta^2(t) dt \right)}{\epsilon^2} \le \frac{\delta}{\epsilon^2}.$$

It follows that

$$\begin{aligned}
&\mathsf{P}\left(\sup_{0\le t\le T}\left|\int_0^t X(t)dB(t)\right|\ge\epsilon\right)\\
\le&\mathsf{P}\left(\sup_{0\le t\le T}\left|\int_0^t \left(X(t)-X_\delta(t)\right)dB(t)\right|\ge\epsilon\right)+\mathsf{P}\left(\sup_{0\le t\le T}\left|\int_0^t X_\delta(t)dB(t)\right|\ge\epsilon\right)\\
\le&\mathsf{P}\left(\int_0^T X^2(t)dt\ge\delta\right)+\frac{\delta}{\epsilon^2}.
\end{aligned}$$

Corollary 2.2.1 *If X is a continuous adapted process then the Itô integral $\int_0^T X(t)dB(t)$ exists. Further, if $\{t_i^n\}$ is a partition of the interval $[0,T]$,*

$$0=t_0^n<t_1^n<\ldots<t_n^n=T,$$

with $\delta_n=\max_{1\le i\le n}(t_i^n-t_{i-1}^n)\to 0$, then

$$\int_0^T X(t)dB(t)=\lim_{n\to\infty}\sum_{i=0}^{n-1}X(t_i^n)\big[B(t_{i+1}^n)-B(t_i^n)\big]\quad\textit{in probability.}$$

Proof. A continuous adapted process X is a predictable process and $\int_0^T X^2(t)dt<\infty$ by the continuity. So, the Itô integral exists. Let

$$X^n(t)=X(0)I_0(t)+\sum_{i=0}^{n-1}X(t_i^n)I_{(t_i^n,t_{i+1}^n]}(t)$$

for the partition $\{t_i^n\}$. Then X^n's are a simple processes and for any t, $X^n(t)\to X(t)$ as $n\to\infty$ by the continuity. For the simple process X^n, the Itô process is

$$\int_0^T X^n(t)dB(t)=\sum_{i=0}^{n-1}X(t_i^n)\big[B(t_{i+1}^n)-B(t_i^n)\big].$$

The result follows.

2.3 Itô's integral process and Stochastic differentials

Suppose X is a predictable process with $\int_0^T X^2(t)dt<\infty$ with probability one, then $\int_0^t X(u)dB(u)$ is well-defined for any $t\le T$. The process

$$Y(t)=\int_0^t X(u)dB(u)$$

is called the Itô integral process. It can be showed that $Y(t)$ has a version with continuous paths. In what follows it is always assumed $Y(t)$ is continuous. So Itô integral process $Y(t) = \int_0^t X(u)dB(u)$ is continuous and adapted. The Itô integral can be written in the differential notation

$$dY(t) = X(t)dB(t).$$

The integral expression and the differential expression have the mean but with different notations. It should be remembered that the differential expression does not mean that

$$\frac{Y(t+\Delta t) - Y(t)}{B(t+\Delta t) - B(t)} \to X(t) \quad \text{as } \Delta \to 0.$$

It is just a differential expression of the Itô integral, that is

$$Y(b) - Y(a) = \lim_n \sum_i X(t_i^n)\big[B(t_{i+1}^n) - B(t_i)\big], \quad \text{when } X(t) \text{ is continuous }.$$

If further $\int_0^T \mathsf{E}[X^2(t)]dt < \infty$, we have shown that $Y(t)$ having the following properties:

1. $Y(t)$ is a continuous square integrable martingale with $\mathsf{E}Y^2(t) = \int_0^t \mathsf{E}[X^2(u)]du$,
2. $[Y, Y](t) = \int_0^t X^2(u)du$,
3. $Y^2(t) - \int_0^t X^2(u)du$ is also a martingale.

When $\int_0^T \mathsf{E}[X^2(t)]dt = \infty$, Properties 1 and 2 may fail. Property 2 remains true.

Theorem 2.3.1 *Quadratic variation of Itô integral process $Y(t) = \int_0^t X(u)dB(u)$ is given by*

$$[Y, Y](t) = \int_0^t X^2(u)du.$$

Proof. Let $\sigma_m(\omega) = \inf\{t : \int_0^t X^2(u,\omega)du \geq m\} \wedge m$. Then σ_m is a sequence of stopping times such that $\sigma_m \uparrow \infty$ a.s.. Set $X_m(t,\omega) = X(t,\omega)I_{\{\sigma_m(\omega)\geq t\}}$. Clearly

$$\int_0^\infty X_m^2(t,\omega)dt = \int_0^{\sigma_m} X^2(t,\omega)dt \leq m.$$

If define $Y_m(t) = \int_0^t X_m(u)dB(u)$, then

$$Y_m(t) = \int_0^{t\wedge\sigma_m} X(u)dB(u) = Y(t\wedge\sigma_m),$$

$$[Y_m, Y_m](t) = \int_0^t X_m^2(u)du = \int_0^{t\wedge\sigma_m} X^2(u)du.$$

So, on the event $\{\sigma_m > t\}$,

$$[Y_m, Y_m](t) = \int_0^t X^2(u)du.$$

On the other hand, on the event $\{\sigma_m > t\}$,

$$\begin{aligned}[Y_m, Y_m](t) &= \lim_n \sum_i \left|Y_m(t_{i+1}^n) - Y_m(t_i^n)\right|^2 \\ &= \lim_n \sum_i \left|Y(t_{i+1}^n \wedge \sigma_m) - Y(t_i^n \wedge \sigma_m)\right|^2 \\ &= \lim_n \sum_i \left|Y(t_{i+1}^n) - Y(t_i^n)\right|^2 = [Y, Y](t).\end{aligned}$$

The proof is completed.

With the same argument, one can show that If $\int_0^T X^2(t)dt < \infty$ a.s., then $Y(t) = \int_0^t X(u)dB(u)$ have the following properties:

1. $Y(t)$ is a continuous local square integrable martingale,
2. $Y^2(t) - \int_0^t X^2(u)du$ is also a local martingale.

Recall that the Quadratic variation of Y is defined by

$$[Y, Y](t) = \lim \sum_{i=0}^{n-1} \left|Y(t_{i+1}^n) - Y(t_i^n)\right|^2 \quad \text{in probability,}$$

when partition $\{t_i^n\}$ of $[0, t]$ become finer and finer. Loosely writing, this can be expressed as

$$\int_0^t \big(dY(u)\big)^2 = \int_0^t X^2(u)du$$

or in differential notation $dY(t)dY(t) = \big(dY(t)\big)^2 = X^2(t)dt$.

Especially,

$$dB(t)dB(t) = (dB(t))^2 = dt.$$

Hence

$$d[Y, Y](t) = dY(t)dY(t) = [X(t)dB(t)][X(t)dB(t)] = X^2(t)dt.$$

Quadratic covariation of Itô integrals

If $Y_1(t)$ and $Y_2(t)$ are Itô integrals of $X_1(t)$ and $X_2(t)$ with respect to the same Brownian motion $B(t)$, we define quadratic covariation of Y_1 and Y_2 on $[0,t]$ by

$$[Y_1, Y_2](t) = \frac{1}{2}\Big([Y_1 + Y_2, Y_1 + Y_2](t) - [Y_1, Y_1](t) - [Y_2, Y_2](t)\Big).$$

Then it follows that

$$[Y_1, Y_2](t) = \int_0^t X_1(s)X_2(s)ds. \tag{2.3.1}$$

It can be shown that quadratic covariation is given by the limit in probability of crossproducts of increments of processes

$$[Y_1, Y_2](t) = \lim \sum_{i=0}^{n-1} \left[Y_1(t_{i+1}^n) - Y_1(t_i^n)\right]\left[Y_2(t_{i+1}^n) - Y_2(t_i^n)\right], \tag{2.3.2}$$

when partition $\{t_i^n\}$ of $[0,t]$ become finer and finer. (2.3.1) can be expressed as in differential notations

$$d[Y_1, Y_2](t) = dY_1(t)dY_2(t) = X_1(t)dB(t)X_2(t)dB(t) = X_1(t)X_2(t)dt.$$

In the limit in (2.3.2), if one of the function say Y_1 is of finite variation and the other is continuous, then it can be showed that the limit is zero.

Theorem 2.3.2 *Suppose f is a real function of finite variation on interval $[a,b]$, and g is a continuous function on $[a,b]$. Then*

$$[f, g]([a,b]) = \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \left[f(t_{i+1}^n) - f(t_i^n)\right]\left[g(t_{i+1}^n) - g(t_i^n)\right] = 0,$$

where the limit is taken over all partitions $\Pi = \{t_i^n\}$ of $[a,b]$ with $\|\Pi\| = \max_i(t_{i+1} - t_i) \to 0$.

Proof. In fact, the summation is

$$\begin{aligned} &\leq \max_i \left|g(t_{i+1}^n) - g(t_i^n)\right| \sum_i \left|f(t_{i+1}^n) - f(t_i^n)\right| \\ &\leq \max_{|t-s|\leq\|\Pi\|} \left|g(t) - g(s)\right| V_f([a,b]) \to 0. \end{aligned}$$

From this theorem, it follows that

$$\left[\int_0^s f(u)du, \int_0^s X(u)dB(u)\right](t) = 0.$$

This can be expressed as

$$\big(f(t)dt\big)\big(X(t)dB(t)\big) = 0.$$

Especially,

$$dtdB(t) = 0.$$

So, we arrive at

$$(dB(t))^2 = dt, \quad dtdB(t) = 0, \quad dtdt = 0. \tag{2.3.3}$$

2.4 Itô formula

2.4.1 Itô formula for Brownian motion

Suppose f is a real function. We want calculate $df(B(t))$. Recall the Taylor's formula,

$$\Delta f(x) = f'(x)\Delta x + \frac{1}{2}f''(x)(\Delta x)^2 + \frac{1}{3!}f'''(x)(\Delta x)^3 + \cdots .$$

That is

$$df(x) = f'(x)dx + \frac{1}{2}f''(x)(dx)^2 + \frac{1}{3!}f'''(x)(dx)^3 + \cdots .$$

If B is a real variable or a real deferential function x, we know that

$$df(x) = f'(x)dx,$$

because $(dx)^2$, $(dx)^3$ etc are zeros. Now,

$$df(B(t)) = f'(B(t))dB(t) + \frac{1}{2}f''(B(t))\big(dB(t)\big)^2 + \frac{1}{3!}f'''(B(t))\big(dB(t)\big)^3 + \cdots$$

Notice $\big(dB(t)\big)^2 = dt$, $\big(dB(t)\big)^3 = \big(dB(t)\big)^2 dB(t) = dtdB(t) = 0$, $\big(dB(t)\big)^k = 0$, $k = 3, \ldots$. It follows that

$$df(B(t)) = f'(B(t))dB(t) + \frac{1}{2}f''(B(t))dt$$

or in integral notations

$$f(B(t)) - f(B(0)) = \int_0^t f'(B(u))dB(u) + \frac{1}{2}\int_0^t f''(B(u))du.$$

This is the Itô formula. The first integral is the stochastic integral and the second is the usual integral.

Theorem 2.4.1 (Itô formula) *If $f(x)$ is twice continuous differential function, then for any t,*

$$f(B(t)) = f(B(0)) + \int_0^t f'(B(u))dB(u) + \frac{1}{2}\int_0^t f''(B(u))du.$$

The Itô formula can also be written in the stochastic differential form:

$$df(B(t)) = f'(B(t))dB(t) + \frac{1}{2}f''(B(t))dt.$$

Proof. Let $\Pi = \{t_i^n\}$ be a partition of $[0,t]$. Clearly,

$$f(B(t)) = f(B(0)) + \sum_{i=0}^{n-1}\left(f\big(B(t_{i+1}^n)\big) - f\big(B(t_i^n)\big)\right).$$

Apply now Taylor's formula to $f\big(B(t_{i+1}^n)\big) - f\big(B(t_i^n)\big)$ to obtain

$$\begin{aligned}&f\big(B(t_{i+1}^n)\big) - f\big(B(t_i^n)\big)\\ =&f'\big(B(t_i^n)\big)\big(B(t_{i+1}^n) - B(t_i^n)\big) + \frac{1}{2}f''(\theta_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big)^2,\end{aligned}$$

where $\theta_i^n \in \big(B(t_{i+1}^n), B(t_i^n)\big)$. Thus,

$$\begin{aligned}f(B(t)) =&f(B(0)) + \sum_{i=0}^{n-1} f'\big(B(t_i^n)\big)\big(B(t_{i+1}^n) - B(t_i^n)\big)\\ &+\frac{1}{2}\sum_{i=0}^{n-1} f''(\theta_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big)^2.\end{aligned}$$

Taking limits as $\|\Pi\| = \max_i(t_{i+1}^n - t_i^n) \to 0$, the first sum converges to the Itô integral $\int_0^t f'\big(B(u)\big)dB(u)$; the second converges to $\int_0^t f''\big(B(u)\big)du$ by the following lemma.

Lemma 2.4.1 *If g is a continuous function and $\Pi = \{t_i\}$ represents partitions of $[0,t]$, then for any $\theta_i^n \in \big(B(t_i^n), B(t_{i+1}^n)\big)$,*

$$\lim_{\|\Pi\|\to 0}\sum_{i=0}^{n-1} g(\theta_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big)^2 = \int_0^t g\big(B(s)\big)du \;\; \textit{in probability.}$$

Proof. Notice

$$\begin{aligned}&\left|\sum_{i=0}^{n-1}\Big(g(\theta_i^n) - g(B(t_i^n))\Big)\big(B(t_{i+1}^n) - B(t_i^n)\big)^2\right|\\ \le&\max_i\left|g(\theta_i^n) - g(B(t_i^n))\right| \cdot \sum_{i=0}^{n-1}\big(B(t_{i+1}^n) - B(t_i^n)\big)^2\\ \to& 0 \quad \text{in probability },\end{aligned}$$

due to the fact that the first term converges to zero almost surely by the continuity of g and B, and the second converges in probability to the quadratic variation Brownian motion t. Now, we show that

$$\sum_{i=0}^{n-1} g(B(t_i^n))\big(B(t_{i+1}^n)-B(t_i^n)\big)^2 \to \int_0^t g(B(s))ds \text{ in probability.} \tag{2.4.1}$$

By continuity of $g(B(t))$ and the definition of the integral, it follows that

$$\sum_{i=0}^{n-1} g(B(t_i^n))\big(t_{i+1}^n - t_i^n\big) \to \int_0^t g(B(s))ds \;\; a.s.$$

Next we show that the difference between sums converges to zero in probability,

$$\sum_{i=0}^{n-1} g(B(t_i^n))\left[\big(B(t_{i+1}^n)-B(t_i^n)\big)^2 - \big(t_{i+1}^n - t_i^n\big)\right] \to 0 \text{ in probability.} \tag{2.4.2}$$

Also, for every given D,

$$\begin{aligned}&\left|\sum_{i=0}^{n-1} g(B(t_i^n))I\{\big|g(B(t_i^n))\big|>D\}\left[\big(B(t_{i+1}^n)-B(t_i^n)\big)^2 - \big(t_{i+1}^n - t_i^n\big)\right]\right\}\\ &\le \max_{s\le t}\big|g(B(s))I\{\big|g(s)\big|>D\}\cdot\left[\sum_{i=0}^{n-1}\big(B(t_{i+1}^n)-B(t_i^n)\big)^2 + t\right]\\ &\quad \to 0 \quad \text{in probability as } n\to\infty \text{ and } D\to\infty,\end{aligned}$$

because, the first term converges to zero as $D\to\infty$ almost surely by continuity of g and B, and the sum in the second term converges to t in probability.

So in showing (2.4.2), without loss of generality, we may assume that g is bounded such that $|g(x)|\le D$. Write $\Delta B_i = B(t_{i+1}^n)-B(t_i^n)$, $\Delta t_i = t_{i+1}^n - t_i^n$ and $g_i = g(B(t_i^n))$. It is easily seen that $\{g_i\big((\Delta B_i)^2-\Delta t_i\big), \mathscr{F}_{t_{i+1}}, i=0,\ldots,n-1\}$ is a sequence of martingale differences with

$$\mathsf{E}\left[\Big(g_i\big((\Delta B_i)^2-\Delta t_i\big)\Big)^2\Big|\mathscr{F}_{t_i}\right] = g_i^2\mathsf{Var}\big[(\Delta B_i)^2\big] = 2g_i^2(\Delta t_i)^2 \le 2D^2(\Delta t_i)^2.$$

So

$$\mathsf{E}\left[\Big(g_i\big((\Delta B_i)^2-\Delta t_i\big)\Big)^2\right] \le 2D^2(\Delta t_i)^2.$$

It follows that

$$\begin{aligned} & \mathsf{E}\left[\Big(\sum_{i=0}^{n-1} g_i\big((\Delta B_i)^2 - \Delta t_i\big)\Big)^2\right] \\ =& \sum_{i=0}^{n-1} \mathsf{E}\left[\Big(g_i\big((\Delta B_i)^2 - \Delta t_i\big)\Big)^2\right] \\ \le& 2D^2 \sum_{i=0}^{n-1} (\Delta t_i)^2 \le 2D^2 t \|\Pi\| \to 0. \end{aligned}$$

(2.4.2) is now proved.

Remark 2.4.1 *The main step of proving the Itô formula is*

$$\begin{aligned} & \sum_{i=0}^{n-1} f''(\theta_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big)^2 \\ \approx & \sum_{i=0}^{n-1} f''\big(B(t_i^n)\big)\big(B(t_{i+1}^n) - B(t_i^n)\big)^2 \\ \approx & \sum_{i=0}^{n-1} f''\big(B(t_i^n)\big)\big(t_{i+1}^n - t_i^n\big)^2 \approx \int_0^t f''\big(B(u)\big)du, \end{aligned}$$

where in the second approximation we used the important property that the quadratic variation Brownian motion is t, that is, $\big(B(t_{i+1}^n) - B(t_i^n)\big)^2 \approx [B,B](t_{i+1}^n) - [B,B](t_i^n) = t_{i+1}^n - t_i^n$.

Example 2.4.1 *Take $f(x) = x^2$, we have*

$$B^2(t) = 2\int_0^t B(u)dB(u) + \int_0^t du = 2\int_0^t B(u)dB(u) + t.$$

In general, take $f(x) = x^m$, $m \ge 2$, we have

$$B^m(t) = m\int_0^t B^{m-1}(u)dB(u) + \frac{m(m-1)}{2}\int_0^t B^{m-2}(u)du.$$

Example 2.4.2 *Find $de^{B(t)}$.*

By using Itô formula with $f(x) = e^x$, we have $f'(x) = e^x$, $f''(x) = e^x$ and

$$\begin{aligned} de^{B(t)} =& df(B(t)) = f'\big(B(t)\big)dB(t) + \frac{1}{2}f''\big(B(t)\big)dt \\ =& e^{B(t)}dB(t) + \frac{1}{2}e^{B(t)}dt. \end{aligned}$$

Thus $X(t) = e^{B(t)}$ has stochastic differential

$$dX(t) = X(t)dB(t) + \frac{1}{2}X(t)dt.$$

2.4.2 Itô Process

Definition

Process $Y(t)$ is called an Itô process if it can be represented as

$$Y(t) = Y(0) + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dB(s), \quad 0 \le t \le T, \tag{2.4.3}$$

or equivalently, it has a stochastic differential as

$$dY(t) = \mu(t)dt + \sigma(t)dB(t),$$

where processes $\mu(t)$ and $\sigma(t)$ satisfy conditions:

1. $\mu(t)$ is adapted and $\int_0^T |\mu(t)|dt < \infty$ a.s.

2. $\sigma(t)$ is predictable and $\int_0^T \sigma^2(s)ds < \infty$ a.s.

Function μ is often called the drift coefficient and function σ the diffusion coefficient. Notice that μ and σ can depend (and often do) on $Y(t)$ and $B(t)$.

A important case is when dependence of μ and σ on t only through $Y(t)$:

$$dY(t) = \mu(Y(t))dt + \sigma(Y(t))dB(t), \quad 0 \le t \le T.$$

Quadratic variation

Recall that the quadratic variation of a stochastic process Y is defined as

$$[Y,Y](t) = [Y,Y]([0,t]) = \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \big(Y(t_{i+1}^n) - Y(t_i^n)\big)^2 \quad \text{in probability,}$$

and the quadratic covariation of two stochastic processes X and Y is defined as

$$\begin{aligned}[X,Y](t) =& \frac{1}{2}\Big([X+Y,X+Y](t) - [X,X](t) - [Y,Y](t)\Big) \\ =& \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \big(X(t_{i+1}^n) - X(t_i^n)\big)\big(Y(t_{i+1}^n) - Y(t_i^n)\big) \quad \text{in probability,}\end{aligned}$$

where the limits are taken over all partitions $\Pi = \{t_i^n\}$ of $[0,t]$ with $\|\Pi\| = \max_i(t_{i+1}^n - t_i^n) \to 0$.

Let Y be a Itô process as

$$Y(t) = Y(0) + \int_0^t \mu(u)du + \int_0^t \sigma(u)dB(u).$$

Notice the process of the first is a adapted process of finite variation, and the second is an Itô integral process. It follows that

$$\begin{aligned}[Y,Y](t) =& \left[\int_0^s \mu(u)du, \int_0^s \mu(u)du\right](t) \\ &+ 2\left[\int_0^s \mu(u)du, \int_0^s \sigma(u)dB(u)\right](t) \\ &+ \left[\int_0^s \sigma(u)dB(u), \int_0^s \sigma(u)dB(u)\right](t) \\ =& 0 + 0 + \int_0^t \sigma^2(s)ds = \int_0^t \sigma^2(s)ds.\end{aligned}$$

And similarly, if $X(t) = X(0) + \int_0^t \overline{\mu}(u)du + \int_0^t \overline{\sigma}(u)dB(u)$ is another Itô process, then

$$[X,Y](t) = [X,Y]([0,t]) = \int_0^t \overline{\sigma}(s)\sigma(s)ds.$$

Introduce a convention

$$dX(t)dY(t) = d[X,Y](t) \text{ and in particular } (dY(t))^2 = d[Y,Y](t).$$

The following rules follow

$$(dt)^2 = d[s,s](t) = 0, \;\; dB(t)dt = d[B(s),s](t) = 0, \;\; \big(dB(t)\big)^2 = d[B,B](t) = dt.$$

Then in the stochastic differential notations,

$$\begin{aligned}d[X,Y](t) =& dX(t)dY(t) = \big(\overline{\mu}(t)dt + \overline{\sigma}(t)dB(t)\big)\big(\mu(t)dt + \sigma(t)dB(t)\big) \\ =& \overline{\mu}(t)\mu(t)(dt)^2 + \Big(\overline{\sigma}(t)\mu(t) + \overline{\mu}(t)\sigma(t)\Big)dB(t)dt + \overline{\sigma}(t)\sigma(t)\big(dB(t)\big)^2 \\ =& 0 + 0 + \overline{\sigma}(t)\sigma(t)dt\end{aligned}$$

2.4.3 Itô formula for Itô processes

Integrals with respect to stochastic differential

Suppose that X has a stochastic differential with respect to B,

$$dX(t) = \mu(t)dt + \sigma(t)dB(t),$$

and $H(t)$ is predictable and satisfies

$$\int_0^t H(s)^2\sigma(s)ds < \infty, \;\; \int_0^t |H(s)\mu(s)|ds < \infty,$$

then both $\int_0^t H(s)\mu(s)ds$ and $\int_0^t H(s)\sigma(s)dB(s)$ are well defined. The stochastic integral $Z(t) = \int_0^t H(s)dX(s)$ is defined as

$$Z(t) = \int_0^t H(s)dX(s) := \int_0^t H(s)\mu(s)ds + \int_0^t H(s)\sigma(s)dB(s)$$

or

$$dZ(t) = H(t)dY(t) = H(t)\mu(t)dt + H(t)\sigma(t)dB(t).$$

Itô formula for $f(X(t))$

Let $X(t)$ have a stochastic differential

$$X(t) = \mu(t)dt + \sigma(t)dB(t).$$

Then $(dX(t))^2 = d[X, X](t) = \sigma^2(t)dt$, $(dX(t))^3 = d[X, X](t) \cdot dX(t) = \sigma^2(t)dt \cdot \mu(t)dt + \sigma^2(t)dt \cdot \sigma^2(t)dB(t) = 0$, $(dX(t))^k = 0$, $k = 1, 2, \ldots$. So

$$\begin{aligned} df(X(t)) =& f'(X(t))dX(t) + \frac{1}{2}f''(X(t))(dX(t))^2 + \frac{1}{3!}f'''(X(t))(dX(t))^3 + \ldots \\ =& f'(X(t))dX(t) + \frac{1}{2}f''(X(t))(dX(t))^2 \\ =& f'(X(t))dX(t) + \frac{1}{2}f''(X(t))d[X, X](t) \\ =& f'(X(t))\mu(t)dt + f'(X(t))\sigma(t)dB(t) + \frac{1}{2}f''(X(t))\sigma^2(t)dt. \end{aligned}$$

Theorem 2.4.2 *Let $X(t)$ have a stochastic differential*

$$dX(t) = \mu(t)dt + \sigma(t)dB(t).$$

If $f(x)$ is twice continuously differentiable, then the stochastic differential of the process $Y(t) = f(X(t))$ exists and is given by

$$\begin{aligned} df(X(t)) =& f'(X(t))dX(t) + \frac{1}{2}f''(X(t))d[X, X](t) \\ =& \left(f'(X(t))\mu(t) + \frac{1}{2}f''(X(t))\sigma^2(t)\right)dt + f'(X(t))\sigma(t)dB(t). \end{aligned}$$

In integral notations

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s))dX(s) + \frac{1}{2}f''(X(s))\sigma^2(s)ds.$$

Example 2.4.3 *Find a process X having the stochastic differential*

$$dX(t) = X(t)dB(t) + \frac{1}{2}X(t)dt. \tag{2.4.4}$$

Solution. Let's look for a positive process X. Let $f(x) = \log x$. Then $f'(x) = 1/x$ and $f''(x) = -1/x^2$. So

$$\begin{aligned} d\log X(t) =& \frac{1}{X(t)}dX(t) - \frac{1}{2}\frac{1}{X^2(t)}(dX(t))^2 \\ =& dB(t) + \frac{1}{2}dt - \frac{1}{2}\frac{1}{X^2(t)}X^2(t)dt \\ =& dB(t). \end{aligned}$$

So that $\log X(t) = \log X(0) + B(t)$, and we find that

$$X(t) = X(0)e^{B(t)}.$$

Using the Itô formula for $X(t) = X(0)e^{B(t)}$ we verify that this X indeed satisfied (2.4.4).

2.4.4 Itô formula for functions of two-variables

If $X(t)$ and $Y(t)$ have stochastic differentials,

$$\begin{aligned} dX(t) =& \mu_X(t)dt + \sigma_X(t)dB(t). \\ dY(t) =& \mu_Y(t)dt + \sigma_Y(t)dB(t). \end{aligned}$$

Then

$$dX(t)dY(t) = d[X, Y](t) = \sigma_X(t)\sigma_Y(t)dt.$$

It follows that

$$\begin{aligned} d\big(X(t)Y(t)\big) =& X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t) \\ =& X(t)dY(t) + Y(t)dX(t) + \sigma_X(t)\sigma_Y(t)dt. \end{aligned}$$

So

$$\begin{aligned} & X(t)Y(t) - X(0)Y(0) \\ =& \int_0^t X(s)dY(s) + \int_0^t Y(s)dX(s) + \int_0^t d[X, Y](s) \\ =& \int_0^t X(s)dY(s) + \int_0^t Y(s)dX(s) + \int_0^t \sigma_X(s)\sigma_Y(s)ds. \end{aligned}$$

This is the formula for integration by parts.

Example 2.4.4 *Let f and g be C^2 functions and $B(t)$ the Brownian motion. Find d $f(B)g(B)$.*

Solution Using Itô formula

$$\begin{aligned} df(B) =& f'(B)dB + \frac{1}{2}f''(B)dt, \\ dg(B) =& g'(B)dB + \frac{1}{2}g''(B)dt., \end{aligned}$$

So

$$df(B)dg(B) = f'(B)g'(B)(dB)^2 = f'(B)g'(B)dt.$$

$$\begin{aligned} d(f(B)g(B)) =& f(B)dg(B) + g(B)df(B) + df(B)dg(B) \\ =& \left[f(B)g'(B) + f'(B)g(B)\right] dB \\ &+ \frac{1}{2}\left[f''(B) + 2f'(B)g'(B) + g''(B)\right] dt \end{aligned}$$

In general, if $f(x, y)$ has continuous partial derivatives up to order two. Then

$$\begin{aligned} df(x,y) =& \frac{\partial}{\partial x}f(x,y)dx + \frac{\partial}{\partial y}f(x,y)dy \\ &+ \frac{1}{2}\left(\frac{\partial^2}{\partial x^2}f(x,y)(dx)^2 + \frac{\partial^2}{\partial y^2}f(x,y)(dy)^2 + 2\frac{\partial^2}{\partial y^2}f(x,y)dxdy\right). \end{aligned}$$

Now,

$$\begin{aligned} &(dX(t))^2 = d[X,X](t) = \sigma_X^2(t)dt, \\ &(dY(t))^2 = d[Y,Y](t) = \sigma_Y^2(t)dt, \\ &dX(t)dY(t) = d[X,Y](t) = \sigma_X(t)\sigma_Y dt. \end{aligned}$$

Theorem 2.4.3 *Let $f(x, y)$ have continuous partial derivatives up to order two (a*

C^2 function) and X, Y be Itô process, then

$$\begin{aligned} df(X(t),Y(t)) =& \frac{\partial}{\partial x} f(X(t),Y(t))dX(t) + \frac{\partial}{\partial y} f(X(t),Y(t))dY(t) \\ &+ \frac{1}{2}\left(\frac{\partial^2}{\partial x^2} f(X(t),Y(t))d[X,X](t) + \frac{\partial^2}{\partial y^2} f((X(t),Y(t))d[Y,Y](t) \right. \\ &\left. +2\frac{\partial^2}{\partial y^2} f(X(t),Y(t))d[X,Y](t)\right) \\ =& \frac{\partial}{\partial x} f(X(t),Y(t))dX(t) + \frac{\partial}{\partial y} f(X(t),Y(t))dY(t) \\ &+ \frac{1}{2}\left(\frac{\partial^2}{\partial x^2} f(X(t),Y(t))\sigma_X^2(t) + \frac{\partial^2}{\partial y^2} f((X(t),Y(t))\sigma_Y^2(t) \right. \\ &\left. +2\frac{\partial^2}{\partial y^2} f(X(t),Y(t))\sigma_X(t)\sigma_Y(t)\right) dt. \end{aligned}$$

Example 2.4.5 *Let $X(t) = e^{B(t)-t/2}$. Find $dX(t)$.*

Solution. $f(x,t) = e^{x-t/2}$. Here $Y(t) = t$. Notice $(dt)^2 = 0$ and $dB(t)dt = 0$. We obtain

$$\begin{aligned} df(B(t),t) =& \frac{\partial f}{\partial x} dB + \frac{\partial f}{\partial t} dt + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dB)^2 \\ =& e^{B-t/2} dB - \frac{1}{2} e^{B-t/2} dt + \frac{1}{2} e^{B-t/2} dt \\ =& e^{B(t)-t/2} dB(t) = X(t)dB(t). \end{aligned}$$

So that

$$dX(t) = X(t)dB(t).$$

2.5 Stochastic differential equation with examples

2.5.1 Stochastic exponential

Let X have a stochastic differential. Find U satisfy

$$dU(t) = U(t)dX(t), \quad U(0) = 1. \tag{2.5.5}$$

If $X(t)$ is of finite variation, then the equation is the ordinary differential equation (ODE) and the solution is $U(t) = e^{X(t)}$ and $\log U(t) = X(t)$. Now, let

$f(x)=\log x$, then $f'(x)=\frac{1}{x}$, $f''(x)=-\frac{1}{x^2}$. From (2.5.5), it follows that $d[U,U](t)=U^2(t)d[X,X](t)$. So

$$\begin{aligned} d\log U(t) =& \frac{1}{U(t)}dU(t)-\frac{1}{2}\frac{1}{U^2(t)}d[U,U](t) \\ =& dX(t)-\frac{1}{2}d[X,X](t)=d\left(X(t)-\frac{1}{2}[X,X](t)\right). \end{aligned}$$

Hence

$$\log U(t)-\log U(0)=X(t)-\frac{1}{2}[X,X](t)-\left(X(0)-\frac{1}{2}[X,X](0)\right).$$

That is

$$\log U(t)=X(t)-\frac{1}{2}[X,X](t),$$

$$U(t)=\exp\left\{X(t)-\frac{1}{2}[X,X](t)\right\}.$$

On the other hand, if $U(t)=\exp\left\{X(t)-\frac{1}{2}[X,X](t)\right\}$, then by Itô formula with $f(x)=e^x$,

$$\begin{aligned} dU(t) =& \exp\left\{X(t)-\frac{1}{2}[X,X](t)\right\}d\big(X(t)-\frac{1}{2}[X,X](t)\big) \\ &+\frac{1}{2}\exp\left\{X(t)-\frac{1}{2}[X,X](t)\right\}d\big[X-[X,X]\big](t) \\ =& U(t)dX(t)-\frac{1}{2}U(t)d[X,X](t)+\frac{1}{2}U(t)d[X,X](t)= \qquad U(t)dX(t). \end{aligned}$$

Indeed, $U(t)=\exp\left\{X(t)-\frac{1}{2}[X,X](t)\right\}$ is a solution of (2.5.5).

Finally, if $V(t)$ is also a solution. Then by Itô formula with $f(x,y)=x/y$,

$$\begin{aligned} d(U(t)/V(t)) =& \frac{1}{V(t)}dU(t)-\frac{U(t)}{V^2(t)}dV(t) \\ &+\frac{1}{2}\cdot 0d[U,U](t)+\frac{1}{2}2\frac{U(t)}{V^3(t)}d[V,V](t)-\frac{1}{V^2(t)}d[U,V](t) \\ =& \frac{1}{V(t)}U(t)dX(t)-\frac{U(t)}{V^2(t)}V(t)dX(t) \\ &+\frac{1}{2}2\frac{U(t)}{V^3(t)}V^2(t)d[X,X](t)-\frac{1}{V^2(t)}U(t)V(t)d[X,X](t) \\ \equiv& 0. \end{aligned}$$

Hence the solution is unique.

Theorem 2.5.1 *Let $X(t)$ be an Itô process. The equation*

$$dU(t) = U(t)dX(t), \quad U(0) = 1$$

has an unique solution

$$U(t) =: \mathscr{E}(X)(t) = \exp\left\{X(t) - \frac{1}{2}[X,X](t)\right\},$$

this process is called the stochastic exponential of X.

Example 2.5.1 *The stochastic exponential of $\alpha B(t)$ is $e^{\alpha B(t) - \frac{\alpha^2}{2}t}$.*

Example 2.5.2 *(Stock processes and its return process.) Let $S(t)$ denotes the price of stock and assume that it has a stochastic differential. Return on stock $R(t)$ is defined by the relation*

$$dR(t) = \frac{dS(t)}{S(t)}.$$

Thus

$$dS(t) = S(t)dR(t).$$

The return is the value of \$1 invested at time t after a unit time.

If return is a constant rate r, then $dS(t) = S(t)rdt$ is the ordinary differential equation and the solution is $S(t) = S(0)e^{rt}$.

In Black-Scholes model, the return is uncertain and assumed to be a constant rate r plus a white noise, that is

$$R(t) = r + \sigma \dot{B}(t),$$

which means that

$$R(t)\Delta t \approx r\Delta t + \sigma(B(t + \Delta t) - B(t)).$$

So

$$dR(t) = rdt + \sigma dB(t).$$

The stock price $S(t)$ is the stochastic exponential of the return $R(t)$,

$$\begin{aligned} S(t) =& S(0)\exp\{R(t) - \frac{1}{2}[R,R](t)\} \\ =& S(0)\exp\left\{B(t) + \left(r - \frac{1}{2}\sigma^2\right)t\right\}. \end{aligned}$$

2.5.2 Definition of the stochastic differential equations

Definition 2.5.1 *A equation of the form*

$$dX(t) = \mu(X(t), t)dt + \sigma(X(t), t)dB(t), \tag{2.5.6}$$

where $\mu(x,t)$ and $\sigma(x,t)$ are given and $X(t)$ is the unknown process, is called a s-tochastic differential equation (SDE) driven by Brownian motion.

$X(t)$ is called a strong solution of the SDE (2.5.6) with initial value $X(0)$ if for all $t > 0$, $X(t)$ is a function $F(t, X(0), (B(s), s \leq t))$ of the given Brownian motion $B(t)$ and $X(0)$, integrals $\int_0^t \mu(X(s), s)ds$ and $\int_0^t \sigma(X(s), s)dB(s)$ exist, and the integral equation is satisfied

$$X(t) = X(0) + \int_0^t \mu\big(X(s), s\big)ds + \int_0^t \sigma\big(X(s), s\big)dB(s).$$

Example 2.5.3 *(Kabgevin equation and Ornstein-Uhlenbeck process). Find the solution of the SDE*

$$dX(t) = -\alpha X(t)dt + \sigma dB(t).$$

Solution. If $\alpha = 0$, then the solution is $X(t) = X(0) + \sigma \int_0^t dB(s) = X(0) + \sigma B(t)$. If $\sigma = 0$, then the solution is $X(t) = X(0)e^{-\alpha t}$. Now, let $Y(t) = X(t)e^{\alpha t}$. Then

$$\begin{aligned} dY(t) =& e^{\alpha t}dX(t) + X(t)\alpha e^{\alpha t}dt + dX(t)de^{\alpha t} \\ =& - e^{\alpha t}\alpha X(t)dt + e^{\alpha t}\sigma dB(t) + X(t)\alpha e^{\alpha t}dt + 0 \\ =& \sigma e^{\alpha t}dB(t). \end{aligned}$$

This gives

$$Y(t) = Y(0) + \int_0^t \sigma e^{\alpha s}dB(s).$$

Hence

$$X(t) = e^{-\alpha t}\left(X(0) + \int_0^t \sigma e^{\alpha s}dB(s)\right).$$

Integration by parts yields

$$\int_0^t \sigma e^{\alpha s}dB(s) = \sigma e^{\alpha t}B(s) - \int_0^t B(s)d\left(\sigma e^{\alpha t}\right).$$

Hence $X(t)$ is a function of $B(s), s \leq t$ and so a strong solution of the equation.

Suppose $X(0) = 0$. Then $X(t)$ is a mean zero Gaussian process with

$$\begin{aligned}
\mathsf{Cov}\{X(t), X(s)\} =& \sigma^2 e^{-\alpha t} e^{-\alpha s} \mathsf{E}\left[\int_0^t e^{\alpha u} dB(u) \int_0^s e^{\alpha v} dB(v)\right] \\
=& \sigma^2 e^{-\alpha(t+s)} \int_0^{s\wedge t} e^{2\alpha u} du \\
=& \frac{\sigma^2}{2\alpha} e^{-\alpha(t+s)} e^{2\alpha(s\wedge t)} \\
=& \frac{\sigma^2}{2\alpha} e^{-\alpha|t-s|}.
\end{aligned}$$

So, $X(t)$ is a stationary Gaussian process.

2.5.3 Existence and uniqueness of strong solution

Theorem 2.5.2 *(Existence and Uniqueness) For the SDE*

$$dX(t) = \mu\big(X(t), t\big)dt + \sigma\big(X(t), t\big)dB(t), \tag{2.5.7}$$

if the following conditions are satisfied

1. *Coefficients are locally Lipschitz in x uniformly in t, that is for every T and N, there is a constant K depending only on T and N such that for all $|x|, |y| \le N$ and all $0 \le t \le T$,*

$$|\mu(x,t) - \mu(y,t)| + |\sigma(x,t) - \sigma(y,t)| \le K|x - y|,$$

2. *Coefficients satisfy the linear growth condition*

$$|\mu(x,t)| + |\sigma(x,t)| \le K(1 + |x|),$$

3. *$X(0)$ is independent of $(B(t), 0 \le t \le T)$, and $\mathsf{E}X^2(0) < \infty$,*

then there exists a unique strong solution $X(t)$ of the SDE. $X(t)$ has continuous paths, moreover

$$\mathsf{E}\left[\sup_{0\le t\le T} X^2(t)\right] \le C\left(1 + \mathsf{E}X^2(0)\right), \tag{2.5.8}$$

where constant C depends only on K and T.

If the coefficients depend on x only, the conditions can be weakened.

Theorem 2.5.3 *(Yamada-Watanabe) Suppose that $\mu(x)$ satisfies Lipschitz condition and $\sigma(x)$ satisfies a Hölder condition of order α, $\alpha \geq 1/2$, that is, there is constant K such that*

$$|\sigma(x) - \sigma(y)| \leq K|x-y|^{\alpha}.$$

Then the strong solution exists and is unique.

The proof of the above theorems can be found in L.C.G. Rogers and D. Williams (1990), *Diffusion, Markov Processes, and Martingale. Volume 2 Itô Calculus*, Wiley. Here we give the proof of Theorem 2.5.2.

Lemma 2.5.1 *Let μ and σ satisfy the locally Lipschitz condition. Let X and Y be adapted processes, and define $\widetilde{X}$ and $\widetilde{Y}$ as follows:*

$$\widetilde{X}(t) = \xi + \int_0^t \mu\big(X(u),u\big)du + \int_0^t \sigma\big(X(u),u\big)dB(u),$$

$$\widetilde{Y}(t) = \eta + \int_0^t \mu\big(Y(u),u\big)du + \int_0^t \sigma\big(Y(u),u\big)dB(u)$$

Then there is a constant C such that

$$E[(\widetilde{X}-\widetilde{Y})_t^{*2}] \leq C\left\{E[|\xi-\eta|^2] + E\left(\int_0^t (X-Y)_u^{*2}du\right)\right\}, \quad 0 \leq t \leq T,$$

where $f_t^* = \sup\{|f(s)| : s \leq t\}$.

Proof. Let $a(u) = \mu\big(X(u),u\big) - \mu\big(Y(u),u\big)$, $b(u) = \sigma\big(X(u),u\big) - \sigma\big(Y(u),u\big)$. Then

$$\widetilde{X}(t) - \widetilde{Y}(t) = \xi - \eta + \int_0^t a(u)du + \int_0^t b(u)dB(u).$$

So

$$\begin{aligned}
&\mathsf{E}[(\widetilde{X}-\widetilde{Y})_t^{*2}] \\
\leq& 3\mathsf{E}[|\xi-\eta|^2] + 3\mathsf{E}\left[\int_0^t |a(u)|du\right]^2 + 3\mathsf{E}\left[\sup_{s\leq t}\left(\int_0^s b(u)dB(u)\right)^2\right] \\
\leq& 3\mathsf{E}[|\xi-\eta|^2] + 3t\mathsf{E}\left[\int_0^t a^2(u)du\right] + 6\mathsf{E}\left[\int_0^t b^2(u)du\right] \\
\leq& 3\mathsf{E}[|\xi-\eta|^2] + 3TK^2\mathsf{E}\left(\int_0^t (X-Y)_u^{*2}du\right) + 3K^2\mathsf{E}\left(\int_0^t (X-Y)_u^{*2}du\right).
\end{aligned}$$

The proof is completed. □

Proof of Theorem 2.5.2. Define

$$\begin{aligned} X_0(t) =& X(0), \quad 0 \le t \le T, \\ X_{n+1}(t) =& (\mathscr{R}X_n)(t) \\ =& X(0) + \int_0^t \mu\big(X_n(u), u\big)du + \int_0^t \sigma\big(X_n(u), u\big)dB(u), \quad 0 \le t \le T. \end{aligned}$$

By the condition in the theorem, X_0 and X_1 are in $L^2_{[0,T]}$. Now, define

$$\Delta_{n+1}(t) = \mathsf{E}\left[(X_{n+1} - X_n)_t^{*2}\right] = \mathsf{E}\left[(\mathscr{R}X_n - \mathscr{R}X_{n-1})_t^{*2}\right].$$

By Lemma 2.5.1, we find that

$$\Delta_{n+1}(t) \le C \int_0^t \Delta_n(u)du,$$

and, by induction, it follows that, for all $0 \le t \le T$,

$$\Delta_n(t) \le \eta C^n t^n / n!,$$

where $\eta = \Delta_0(T)$ with

$$\begin{aligned} \Delta_0(t) \le& 2K^2t^2\big(1 + \mathsf{E}[X^2(0)]\big) + 2K^2t\big(1 + \mathsf{E}[X^2(0)]\big) \\ =& 2K^2(t^2 + t)\big(1 + \mathsf{E}[X^2(0)]\big) < \infty \end{aligned}$$

by Condition 2. Hence, for $m > n \ge 0$,

$$\begin{aligned} &\big(\mathsf{E}[(X_m - X_n)_T^{*2}]\big)^{1/2} \le \left(\mathsf{E}\left[(\sum_{k=n}^{m-1}(X_{k+1} - X_k)_T^{*})^2\right]\right)^{1/2} \\ \le& \sum_{k=n}^{m-1}\big(\Delta_{k+1}(T)\big)^{1/2} \le \sum_{k=n}^{\infty}\left(\eta\frac{C^{k+1}T^{k+1}}{(k+1)!}\right)^{1/2} \to 0 \text{ as } n \to \infty. \end{aligned}$$

Hence, on $[0, T]$, the sequence X_n converges uniformly in L_2 to a adapted process X. Since

$$\mathsf{E}[(X - X_n)_T^{*2}] \le \eta \left(\sum_{k=n}^{\infty}\left(\frac{C^{k+1}T^{k+1}}{(k+1)!}\right)^{1/2}\right)^2 \to 0 \text{ as } n \to \infty, \tag{2.5.9}$$

it follows by Lemma 2.5.1 that

$$\mathsf{E}[(\mathscr{R}X - X_{n+1})_T^{*2}] = \mathsf{E}[(\mathscr{R}X - \mathscr{R}X_n)_T^{*2}] \le C\mathsf{E}[(X - X_n)_T^{*2}] \to 0 \text{ as } n \to \infty.$$

Hence, $X = \mathscr{R}X$, and X is a solution of the SDE. Since the integrals are continuous, so X is continuous. (2.5.8) follows from (2.5.9) with $n = 0$.

Next, we show the uniqueness of the solution. Let X and X' be two solutions. By Lemma 2.5.1 it follows that

$$\mathsf{E}[(X - X')_t^{*2}] \le C \int_0^t \mathsf{E}[(X - X')_u^{*2}]du \ \ 0 \le t \le T,$$

which implies that $\mathsf{E}[(X - X')_t^{*2}] = 0$ for all $0 \le t \le T$, and hence that $X = X'$ on $[0, T]$.

In fact, let $w(t) \int_0^t \mathsf{E}[(X - X')_u^{*2}]du$ and $f(t) = w(t)e^{-Ct}$. Then $0 \le \frac{d}{dt}w(t) \le Cw(t)$ and $\frac{d}{dt}f(t) \le 0$. It follows that $f(t) \le f(0)$, $w(t) \le w(0)e^{Ct} = 0$, $\frac{d}{dt}w(t) = 0$. □

Example 2.5.4 *(Gisanov's SDE)*

$$dX(t) = |X(t)|^r dB(t), \quad X(0) = 0, 1/2 \le r < 1.$$

By Yamada-Watanabe theorem, this SDE has a unique solution. Obviously, $X(t) \equiv 0$ is the solution.

Example 2.5.5 *(Tanaka's SDE)*

$$dX(t) = sign\big(X(t)\big)dB(t), \tag{2.5.10}$$

where

$$sign(x) = \begin{cases} 1 & if\ x > 0, \\ -1 & if\ x \le 0. \end{cases}.$$

Since $\sigma(x) = \text{sign}(x)$ is discontinuous, it is not Lipschitz, and conditions for strong existence fail. It can be shown that a strong solution to Tanaks's SDE does not exist, for example, one can refer to Gihman I. I. and Skorohod A. V. (1982), *Stochastic Differential Equations*, Springer-Verlag.

Consider the case $X(0) = 0$. Here we show that if the Tanaka's SDE has a strong solution, then the solution is not unique.

Suppose $X(t)$ is an unique strong solution, then

$$X(t) = \int_0^t \text{sign}\big(X(s)\big)dB(s)$$

is a function of $B(s), s \le t$. Also the solution is a continuous martingale with $[X,X](t) = \int_0^t \big(\text{sign}\big(X(s)\big)\big)^2 ds = t$. So by Levy's theorem $X(t)$ is a Brownian motion. Let $Y(t) = -X(t)$. Then

$$\begin{aligned} Y(t) &= \int_0^t -\text{sign}\big(X(s)\big) dB(s) \\ &= \int_0^t \text{sign}\big(-X(s)\big) dB(s) + 2\int_0^t I_{\{0\}}\big(X(s)\big) dB(s). \end{aligned}$$

Notice

$$\begin{aligned} &\mathsf{E} \sup_{t \le T} \left(\int_0^t I_{\{0\}}\big(X(s)\big) dB(s) \right)^2 \\ \le & 4\mathsf{E} \left(\int_0^T I_{\{0\}}\big(X(s)\big) dB(s) \right)^2 \\ = & 4\mathsf{E} \int_0^T I^2_{\{0\}}\big(X(s)\big) ds \\ = & 4\mathsf{E}\lambda(\{s \in [0,T] : X(s) = 0\}) = 0. \end{aligned}$$

Here λ is the Lebegue measure. It follows that

$$\mathsf{P}\Big(\int_0^t I_{\{0\}}\big(X(s)\big) dB(s) = 0 \ \ \forall t\Big) = 1.$$

Hence

$$Y(t) = \int_0^t \text{sign}\big(Y(s)\big) dB(s),$$

which means that $Y(t) = -X(t)$ is also a solution. By the uniqueness, we must have $\mathsf{P}(X(t) = -X(t)) = 1$ which is impossible because $X(t)$ is a Brownian motion.

Markov property of the solution

Theorem 2.5.4 *Under the theorem for the existence and uniqueness, the strong solution of the SDE*

$$dX(t) = \mu\big(X(t), t\big) dt + \sigma\big(X(t), t\big) dB(t), \quad X(0) = X_0,$$

is a Markov process.

Proof. Let $f(x)$ be a bounded function, $s, t \ge 0$. It is sufficient to show that

$$\mathsf{E}\left[f(X(s+t)) \big| \mathscr{F}_s\right] = \mathsf{E}\left[f(X(s+t)) \big| (X(s)\right].$$

Let $\widetilde{B}(u) = B(s+u) - B(s)$. Then $\widetilde{B}$ is independent of $\mathcal{F}_s$. Denote $X^{s,x}(t)$ be the unique solution of the SDE

$$dX(t) = \mu\big(X(t), s+t\big)dt + \sigma\big(X(t), s+t\big)d\widetilde{B}(t), \quad X(0) = x,$$

i.e.,

$$X^{s,x}(t) = x + \int_0^t \mu\big(X^{s,x}(u), s+u\big)du + \int_0^t \sigma\big(X^{s,x}(u), s+u\big)d\widetilde{B}(u).$$

Then $X^{s,x}(t)$ is independent of $\mathcal{F}_s$. Denote $F(x,s,t) = X^{s,x}(t)$. Note

$$\begin{aligned} & X(s+t) - X(s) \\ =& \int_s^{s+t} \mu\big(X(u), u\big)du + \int_s^{s+t} \sigma\big(X(u), u\big)dB(u) \\ =& \int_0^t \mu\big(X(s+u), s+u\big)du + \int_0^t \sigma\big(X(s+u), s+u\big)d\widetilde{B}(u). \end{aligned}$$

So

$$X(s+t) = X^{s,X(s)}(t) = F(X(s), s, t).$$

Let $G(x) = f\left(F(x,s,t)\right)$. Then $G(x)$ is independent of $\mathscr{F}_s$.

It follows that

$$\begin{aligned} & \mathsf{E}\left[f(X(s+t))\big|\mathscr{F}_s\right] = \mathsf{E}\left[f\big(F(X(s),s,t)\big)\big|\mathscr{F}_s\right] \\ =& \mathsf{E}\left[G(X(s))\big|\mathscr{F}_s\right] \\ =& \Big(\mathsf{E}\left[G(y)\big|\mathscr{F}_s\right]\Big)\Big|_{y=X(s)} \\ =& \Big(\mathsf{E}\left[G(y)\right]\Big)\Big|_{y=X(s)} \\ =& \mathsf{E}\left[G(X(s))\big|X(s)\right] = \mathsf{E}\left[f(X(s+t))\big|X(s)\right]. \end{aligned}$$

The proof is completed. □

Strong Markov property of the solution

Theorem 2.5.5 *Under the theorem for the existence and uniqueness, the strong solution of the SDE*

$$dX(t) = \mu\big(X(t)\big)dt + \sigma\big(X(t)\big)dB(t), \quad X(0) = X_0,$$

is a strong Markov process.

Proof. Let $f(x)$ be a bounded function, τ is a stopping time, $t \geq 0$. It is sufficient to show that

$$\mathsf{E}\left[f(X(\tau+t))\middle|\mathscr{F}_\tau\right] = \mathsf{E}\left[f(X(\tau+t))\middle|(X(\tau)\right].$$

Let $\widetilde{B}(u) = B(\tau+u) - B(\tau)$. Then $\widetilde{B}$ is independent of $\mathcal{F}_\tau$. Denote $X^x(t)$ be the unique solution of the SDE

$$dX(t) = \mu\big(X(t)\big)dt + \sigma\big(X(t)\big)d\widetilde{B}(t), \quad X(0) = x,$$

i.e.,

$$X^x(t) = x + \int_0^t \mu\big(X^x(u)\big)du + \int_0^t \sigma\big(X^x(u)\big)d\widetilde{B}(u).$$

Then $X^x(t)$ is independent of $\mathcal{F}_\tau$. Denote $F(x,t) = X^x(t)$. Note

$$\begin{aligned} & X(\tau+t) - X(\tau) \\ = & \int_\tau^{\tau+t} \mu\big(X(u)\big)du + \int_\tau^{\tau+t} \sigma\big(X(u)\big)dB(u) \\ = & \int_0^t \mu\big(X(\tau+u)\big)du + \int_0^t \sigma\big(X(\tau+u)\big)d\widetilde{B}(u). \end{aligned}$$

So

$$X(\tau+t) = X^{X(\tau)}(t) = F(X(\tau), t).$$

Let $G(x) = f\left(F(x,t)\right)$. Then $G(x)$ is independent of $\mathscr{F}_\tau$. It follows that

$$\begin{aligned} & \mathsf{E}\left[f(X(\tau+t))\middle|\mathscr{F}_\tau\right] = \mathsf{E}\left[f\big(F(X(\tau),t)\big)\middle|\mathscr{F}_\tau\right] \\ = & \mathsf{E}\left[G(X(\tau))\middle|\mathscr{F}_\tau\right] \\ = & \Big(\mathsf{E}\left[G(y)\middle|\mathscr{F}_\tau\right]\Big)\Big|_{y=X(\tau)} \\ = & \Big(\mathsf{E}\left[G(y)\right]\Big)\Big|_{y=X(s)} \\ = & \mathsf{E}\left[G(X(\tau))\middle|X(\tau)\right] = \mathsf{E}\left[f(X(\tau+t))\middle|X(\tau)\right]. \end{aligned}$$

The proof is completed. □

2.5.4 Solutions to linear stochastic differential equations

Consider the general linear SDE in one dimension:

$$dX(t) = \big(\alpha(t) + \beta(t)X(t)\big)dt + \big(\gamma(t) + \delta(t)X(t)\big)dB(t), \tag{2.5.11}$$

where functions $\alpha(\cdot)$, $\beta(\cdot)$, $\gamma(\cdot)$ and $\delta(t)$ are given adapted processes, and are continuous functions of t.

Stochastic exponential SDE's

First, consider the case when $\alpha(t) \equiv 0$ and $\gamma(t) \equiv 0$. The SDE becomes

$$dU(t) = \beta(t)U(t)dt + \delta(t)U(t)dB(t). \tag{2.5.12}$$

This SDE is of the form

$$dU(t) = U(t)dY(t),$$

where the Itô process $Y(t)$ is defined by

$$dY(t) = \beta(t)dt + \delta(t)dB(t).$$

So, the uniqueness solution of (2.5.12) is the stochastic exponential of $Y(t)$ given by

$$\begin{aligned} U(t) =& \mathcal{E}(Y)(t) \\ =& U(0)\exp\left\{Y(t) - Y(0) - \frac{1}{2}[Y,Y](t)\right\} \\ =& U(0)\exp\left\{\int_0^t \beta(s)ds + \int_0^t \delta(s)dB(s) - \frac{1}{2}\int_0^t \delta^2(s)ds\right\} \\ =& U(0)\exp\left\{\int_0^t \Big(\beta(s) - \frac{1}{2}\delta^2(s)\Big)ds + \int_0^t \delta(s)dB(s)\right\}. \end{aligned} \tag{2.5.13}$$

General linear SDE's

To find a solution in the general case with nonzero $\alpha(t)$ and $\gamma(t)$, look for a solution of the form

$$X(t) = U(t)V(t), \tag{2.5.14}$$

where

$$dU(t) = \beta(t)U(t)dt + \delta(t)U(t)dB(t) \tag{2.5.15}$$

and

$$dV(t) = a(t)dt + b(t)dB(t). \tag{2.5.16}$$

Set $U(0) = 1$ and $V(0) = X(0)$. Note that U is given by (2.5.13). Taking the differential of the product it is easy seen that

$$\begin{aligned} dX(t) =& U(t)dV(t) + V(t)dU(t) + dU(t)dV(t) \\ =& U(t)\big(a(t)dt + b(t)dB(t)\big) + V(t)U(t)\big(\beta(t)dt + \delta(t)dB(t)\big) \\ &+ \delta(t)U(t)b(t)dB(t) \\ =& \Big(a(t)U(t) + \delta(t)b(t)U(t) + X(t)\beta(t)\Big)dt + \Big(U(t)b(t) + X(t)\delta(t)\Big)dB(t). \end{aligned}$$

By comparing the coefficients in the above equation and the equation (2.5.11), it follows that

$$b(t)U(t) = \gamma(t) \text{ and } a(t)U(t) = \alpha(t) - \delta(t)\gamma(t).$$

Hence

$$\begin{aligned} V(t) - V(0) =& \int_0^t a(s)ds + \int_0^t b(s)dB(s) \\ =& \int_0^t \frac{\alpha(s) - \delta(s)\gamma(s)}{U(s)}ds + \int_0^t \frac{\gamma(s)}{U(s)}dB(s). \end{aligned}$$

Thus, $X(t)$ is found to be

$$X(t) = U(t)\left(X(0) + \int_0^t \frac{\alpha(s) - \delta(s)\gamma(s)}{U(s)}ds + \int_0^t \frac{\gamma(s)}{U(s)}dB(s)\right),$$

where

$$U(t) = \exp\left\{\int_0^t \Big(\beta(s) - \frac{1}{2}\delta^2(s)\Big)ds + \int_0^t \delta(s)dB(s)\right\}.$$

To show the uniqueness, suppose X_1 and X_2 are solutions of (2.5.11) with $X_1(0) = X_2(0)$,

$$\begin{aligned} dX_1(t) =& \big(\alpha(t) + \beta(t)X_1(t)\big)dt + \big(\gamma(t) + \delta(t)X_1(t)\big)dB(t), \\ dX_2(t) =& \big(\alpha(t) + \beta(t)X_2(t)\big)dt + \big(\gamma(t) + \delta(t)X_2(t)\big)dB(t). \end{aligned}$$

Then $Z(t) = X_1(t) - X_2(t)$ satisfies

$$dZ(t) = \beta(t)Z(t)dt + \delta(t)Z(t)dB(t), \quad Z(0) = 0.$$

So $Z(t)$ satisfies the equation (2.5.16). $Z(t) \equiv 0$ is the unique solution. In fact,

$$\begin{aligned} dZ(t) =& Z(t)dY(t), \quad Z(0) = 0, \\ dU(t) =& U(t)dY(t), \quad U(0) = 1, \\ U(t) =& \exp\left\{Y(t) - \frac{1}{2}[U, U](t)\right\}. \end{aligned}$$

Then

$$d\frac{Z(t)}{U(t)} \equiv 0.$$

So

$$Z(t) = CU(t).$$

Letting $t = 0$ yields $C = 0$. It follows that

$$Z(t) \equiv 0.$$

2.5.5 Examples

I. Interest models

Example 2.5.6 *(Vasicek interest rate model) The vasicek model for the interest rate process $R(t)$ is*

$$dR(t) = \Big(\alpha - \beta R(t)\Big)dt + \sigma dB(t), \tag{2.5.17}$$

where α, β and σ are positive constants.

To solve the equation, we first solute $dR(t) = -\beta R(t)dt$. Obviously, the solution is $R(t) = R(0)e^{-\beta t}$. Then write

$$R(t) = e^{-\beta t}V(t), \quad \text{with } V(0) = R(0).$$

Then

$$dR(t) = -\beta e^{-\beta t}V(t)dt + e^{-\beta t}dV(t) = -\beta R(t)dt + e^{-\beta t}dV(t).$$

So

$$e^{-\beta t}dV(t) = \alpha dt + \sigma dB(t).$$

Hence

$$\begin{aligned} V(t) =& V(0) + \alpha \int_0^t e^{\beta s}ds + \sigma \int_0^t e^{\beta s}dB(s) \\ =& R(0) + \frac{\alpha}{\beta}\big(e^{\beta t} - 1\big) + \sigma \int_0^t e^{\beta s}dB(s). \end{aligned}$$

It follows that

$$R(t) = e^{-\beta t}R(0) + \frac{\alpha}{\beta}\big(1 - e^{-\beta t}\big) + \sigma e^{-\beta t}\int_0^t e^{\beta s}dB(s).$$

The term $\int_0^t e^{\beta s} dB(s)$ is a normal random variable with mean zero and variance

$$\int_0^s e^{2\beta s} ds = \frac{1}{2\beta}\big(e^{2\beta t} - 1\big).$$

Therefore, $R(t)$ is normally distributed with mean

$$e^{-\beta t} R(0) + \frac{\alpha}{\beta}\big(1 - e^{-\beta t}\big)$$

and variance

$$\frac{\sigma^2}{2\beta}\big(1 - e^{-2\beta t}\big).$$

Desirable property The Vasicek model has the desirable property that the interest rate is *mean-reverting*.

1. When $R(t) = \frac{\alpha}{\beta}$, the drift term in (2.5.17) is zero.
2. When $R(t) > \frac{\alpha}{\beta}$, this term is negative, which pushes $R(t)$ back toward $\frac{\alpha}{\beta}$.
3. When $R(t) < \frac{\alpha}{\beta}$, this term is positive, which again pushes $R(t)$ back toward $\frac{\alpha}{\beta}$.

If $R(0) = \frac{\alpha}{\beta}$, then $\mathsf{E}R(t) = \frac{\alpha}{\beta}$ for all t. If $R(0) \neq \frac{\alpha}{\beta}$, then $\lim_{t\to\infty} \mathsf{E}R(t) = \frac{\alpha}{\beta}$.

Undesirable property Note that $R(t)$ is normal distributed, no matter how the parameter $\alpha > 0$, $\beta > 0$ and $\sigma > 0$ are chosen, there is positive probability that $R(t)$ is negative, an undesirable property for an interest rate model.

Example 2.5.7 *(Cox-Ingesoll-Ross (CIR) interest rate model). The Cox-Ingesoll-Ross model the interest rate $R(t)$ is*

$$dR(t) = \Big(\alpha - \beta R(t)\Big)dt + \sigma\sqrt{R(t)}dB(t), \tag{2.5.18}$$

where α, β and σ are positive constants.

Like the Vasicek model, the CIR model is mean-reverting. The advantage of the CIR model over the Vasicek model is that the interest rate in the CIR model does not become negative. If $R(t)$ reaches zero, the term multiplying $dB(t)$ vanishes and the positive drift term αdt in the equation (2.5.18) drives the interest rate back into positive territory.

Unlike the Vasicek equation (2.5.17), the CIR equation (2.5.18) does not have a closed-form solution. However, the distribution of $R(t)$ for each positive t can be determined. That computation would take us too far afield. We derive the mean and variance of $R(t)$ instead.

To do so, we also write

$$R(t) = e^{-\beta t}V(t), \quad \text{with } V(0) = R(0).$$

Then

$$dR(t) = -\beta e^{-\beta t}V(t)dt + e^{-\beta t}dV(t) = -\beta R(t)dt + e^{-\beta t}dV(t).$$

So

$$e^{-\beta t}dV(t) = \alpha dt + \sigma\sqrt{R(t)}dB(t).$$

Hence

$$\begin{aligned} V(t) =& V(0) + \alpha\int_0^t e^{\beta s}ds + \sigma\int_0^t e^{\beta s}\sqrt{R(s)}dB(s) \\ =& R(0) + \frac{\alpha}{\beta}\big(e^{\beta t} - 1\big) + \sigma\int_0^t e^{\beta s}\sqrt{R(s)}dB(s). \end{aligned}$$

It follows that

$$R(t) = e^{-\beta t}R(0) + \frac{\alpha}{\beta}\big(1 - e^{-\beta t}\big) + \sigma e^{-\beta t}\int_0^t e^{\beta s}\sqrt{R(s)}dB(s).$$

Notice the expectation of an Itô integral is zero, we obtain

$$\mathsf{E}R(t) = e^{-\beta t}R(0) + \frac{\alpha}{\beta}\big(1 - e^{-\beta t}\big).$$

This is the same expectation as in the Vasicek model.

Also, by the Isometry property of the Itô integral, we obtain

$$\begin{aligned} \mathsf{Var}\{R(t)\} =& \sigma^2 e^{-2\beta t}\mathsf{Var}\left\{\int_0^t e^{\beta s}\sqrt{R(s)}dB(s)\right\} \\ =& \sigma^2 e^{-2\beta t}\int_0^t e^{2\beta s}\mathsf{E}[R(s)]ds \\ =& \sigma^2 e^{-2\beta t}\int_0^t e^{2\beta s}\left[e^{-\beta s}R(0) + \frac{\alpha}{\beta}\big(1 - e^{-\beta s}\big)\right]ds \\ =& \frac{\sigma^2}{\beta}R(0)\left(e^{-\beta t} - e^{-2\beta t}\right) + \frac{\alpha\sigma^2}{2\beta^2}\left(1 - 2e^{-\beta t} + e^{-2\beta t}\right). \end{aligned}$$

In particular,

$$\lim_{t\to\infty}\mathsf{Var}\{R(t)\} = \frac{\alpha\sigma^2}{2\beta^2}.$$

Let $m(u,t) = Ee^{uR(t)}$ be the moment generating function of $R(t)$. Then, it seem that we shall have

$$\frac{\partial m}{\partial t} = E\frac{\partial}{\partial t}e^{uR(t)}.$$

However, $\frac{\partial}{\partial t}e^{uR(t)}$ shall be understood as a stochastic derivative. By Itô's formula,

$$\begin{aligned} de^{uR(t)} &= e^{uR(t)}udR(t) + \frac{1}{2}e^{uR(t)}u^2(dR(t))^2 \\ =&e^{uR(t)}u\left[(\alpha-\beta R(t))dt + \sigma\sqrt{R(t)}dB(t)\right] + \frac{1}{2}e^{uR(t)}u^2\sigma^2R(t)dt \\ =&\alpha ue^{uR(t)}dt + \left(\frac{1}{2}u^2\sigma^2 - u\beta\right)e^{uR(t)}R(t)dt \\ &+ e^{uR(t)}\sigma\sqrt{R(t)}dB(t). \end{aligned}$$

It follows that

$$e^{uR(t)} - e^{uR(0)} = \int_0^t (\cdot)(s)ds + \int_0^t e^{uR(s)}\sigma\sqrt{R(s)}dB(s).$$

Taking the expectation yields

$$m(u,t) - m(u,0) = \int_0^t [E(\cdot)(s)]\,ds.$$

So

$$\begin{aligned} \frac{\partial m}{\partial t} =&E(\cdot)(t) \\ =&\alpha uEe^{uR(t)} + \left(\frac{1}{2}u^2\sigma^2 - u\beta\right)E[e^{uR(t)}R(t)] \\ =&\alpha um + \left(\frac{1}{2}u^2\sigma^2 - u\beta\right)\frac{\partial m}{\partial u}. \end{aligned}$$

We arrive the equation that

$$\begin{cases} \frac{\partial m}{\partial t} = \alpha um + \left(\frac{1}{2}u^2\sigma^2 - u\beta\right)\frac{\partial m}{\partial u}, \\ m(0,t) = 1, \\ m(u,0) = e^{uR(0)}. \end{cases}$$

II. Black-Scholes-Merton Equation

Evolution of portfolio value

Consider an agent who at each time t has a portfolio valued at $X(t)$. This portfolio invests in a money market account paying a constant rate of interest r and in stock modeled by

$$dS(t) = \alpha S(t)dt + \sigma S(t)dB(t).$$

Suppose at each time t, the investor holds $\Delta(t)$ shares of stock. The position $\Delta(t)$ can be random but must be adapted to the filtration associated with the Brownian motion $B(t)$. The remainder of the portfolio value, $X(t) - \Delta(t)S(t)$, is invested in the money market account.

The differential $dX(t)$ is due to two factors,

1. the capital gain $\Delta(t)dS(t)$ on the stock position and
2. the interest earnings $r\big(X(t) - \Delta(t)S(t)\big)dt$ on the cash position.

In the other words

$$\begin{aligned} dX(t) =& \Delta(t)dS(t) + r\big(X(t) - \Delta(t)S(t)\big)dt \\ =& \Delta(t)\big(\alpha S(t)dt + \sigma S(t)dB(t)\big) + r\big(X(t) - \Delta(t)S(t)\big)dt \\ =& rX(t)dt + \Delta(t)\big(\alpha - r\big)S(t)dt + \Delta(t)\sigma S(t)dB(t). \end{aligned} \tag{2.5.19}$$

The three terms appearing in the last line above can be understood as follows:

1. any average underlying rate of return r on the portfolio, which is reflected by the term $rX(t)dt$,
2. a risk premium $\alpha - r$ for investing in the stock, which is reflected by the term $\Delta(t)(\alpha - r)S(t)dt$, and
3. a volatility term proportional to the size of stock investment, which is the term $\Delta(t)\sigma S(t)dB(t)$.

Discount. We shall often consider the discounted stock price $e^{-rt}S(t)$ and the discounted portfolio value of agent, $e^{-rt}X(t)$. According the Itô formula,

$$\begin{aligned} d\big(e^{-rt}S(t)\big) =& S(t)de^{-rt} + e^{-rt}dS(t) + d[e^{-rs}, S(s)](t) \\ =& -re^{-rt}S(t)dt + e^{-rt}dS(t) \\ =& (\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dB(t) \end{aligned}$$

and

$$\begin{aligned} d\big(e^{-rt}X(t)\big) =& X(t)de^{-rt} + e^{-rt}dX(t) + d[e^{-rs}, X(s)](t) \\ =& -re^{-rt}X(t)dt + e^{-rt}dX(t) \\ =& \Delta(t)(\alpha - r)e^{-rt}S(t)dt + \Delta(t)\sigma e^{-rt}S(t)dB(t) \\ =& \Delta(t)d\big(e^{-rt}S(t)\big). \end{aligned}$$

The last line shows that change in the discounted portfolio value is solely due to change in the discounted stock price.

Evolution of option value

Consider a European call that pays $(S(T) - K)^+$ at time T. Black, Scholes and Merton argued that the value of this call at any time should depend on the time (more precisely, on the time to expiration) and on the value of the stock price at that time. Following this reasoning, we let $V(t, x)$ denote the value of the call at time t if the stock price at that time is $S(t) = x$. So the value of the call is $V(t) = V(t, S(t))$. According to the Itô formula, the differential of $V(t)$ is

$$\begin{aligned} dV(t) =& V_t(t, S)dt + V_x(t, S)dS + \frac{1}{2}V_{xx}(t, S)(dS)^2 \\ =& V_t(t, S)dt + V_x(t, S)\big(\alpha S dt + \alpha S dB\big) + \frac{1}{2}V_{xx}(t, S)\sigma^2 S^2 dt \\ =& \big[V_t(t, S) + \alpha S V_x(t, S) + \frac{1}{2}\sigma^2 S^2 V_{xx}(t, S)\big]dt + \sigma S V_x(t, S)dB. \end{aligned} \tag{2.5.20}$$

Equating the Evolutions

A (short option) hedging portfolio starts with some initial capital $X(0)$ and invests in the stock and money market account so that the portfolio value $X(t)$ at each time $t \in [0, T]$ agrees with $V(t, S(t))$. This happens if and only if

$$dX(t) = dV(t, S(t)) \tag{2.5.21}$$

and $X(0) = V(0, S(0))$. Comparing (2.5.19) and (2.5.19), (2.5.21) holds if and only if

$$\begin{aligned} &\big[rV(t, S) + \Delta(t)\big(\alpha - r\big)S\big]dt + \Delta(t)\sigma S dB \\ =& \big[V_t(t, S) + \alpha S V_x(t, S) + \frac{1}{2}\sigma^2 S^2 V_{xx}(t, S)\big]dt + \sigma S V_x(t, S)dB. \end{aligned}$$

It follows that

$$\begin{cases} \Delta(t) = V_x(t,S), \\ rV(t,S) + \Delta(t)\big(\alpha - r\big)S = V_t(t,S) + \alpha S V_x(t,S) + \frac{1}{2}\sigma^2 S^2 V_{xx}(t,S). \end{cases}$$

So, $V(t,x)$ satisfies the *Black-Scholes-Merton* partial differential equation:

$$V_t + rxV_x + \frac{1}{2}\sigma^2 x^2 V_{xx} = rV \quad \text{for all } t \in [0,T), \quad x \geq 0, \tag{2.5.22}$$

and that satisfies the terminal equation

$$V(t,x) = (x-K)^+.$$

2.5.6 Weak solutions to the stochastic differential equations

Definition 2.5.2 *If there exist a probability space with filtration, Brownian motion $\widehat{B}(t)$ adapted to that filtration, a process $\widehat{X}(t)$ adapted to that filtration, such that $\widehat{X}(0)$ has distribution F_0, and for all t integrals below are defined, and $\widehat{X}(t)$ satisfies*

$$\widehat{X}(t) = \widehat{X}(0) + \int_0^t \mu\big(\widehat{X}(s), s\big)ds + \int_0^t \sigma\big(\widehat{X}(s), s\big)d\widehat{B}(s),$$

then $\widehat{X}(t)$ is called a weak solution to the SDE

$$dX(t) = \mu(X(t),t)dt + \sigma(X(t),t)dB(t), \tag{2.5.23}$$

with initial distribution F_0.

Definition 2.5.3 *Weak solution is called unique if whenever $X(t)$ and $X'(t)$ are two solutions (perhaps on different probability spaces) such that the distributions of $X(0)$ and $X'(0)$ are the same, then all finite dimensional distributions of $X(t)$ and $X'(t)$ are the same.*

Example 2.5.8 *(Tanaka's SDE)*

$$dX(t) = sign\big(X(t)\big)dB(t), \tag{2.5.24}$$

where

$$sign(x) = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x \leq 0. \end{cases}.$$

It can be shown that a strong solution to Tanaks's SDE does not exist, for example, one can refer to Gihman I. I. and Skorohod A. V. (1982), *Stochastic Differential Equations*, Springer-Verlag.

We show that Brownian motion is the unique weak solution of Tanaka's SDE. Let X_0 be the initial value and $\overline{B}(t)$ be some standard Brownian motion defined on the same probability space. Consider the Processes

$$\begin{aligned} X(t) =& X_0 + \overline{B}(t), \\ \widehat{B}(t) =& \int_0^t \frac{1}{\operatorname{sign}\big(X(s)\big)} d\overline{B}(s) = \int_0^t \operatorname{sign}\big(X(s)\big) d\overline{B}(s). \end{aligned}$$

Then $\widehat{B}(t)$ is a continuous martingale with

$$[\widehat{B}, \widehat{B}](t) = \int_0^t \big[\operatorname{sign}\big(X(s)\big)\big]^2 \, ds = t.$$

So by Levy's theorem $\widehat{B}(t)$ is a Brownian motion, different to the original one $\overline{B}(t)$. Then

$$dX(t) = d\overline{B}(t) = \operatorname{sign}\big(X(t)\big) d\widehat{B}(t), \quad X(0) = X_0.$$

Levy's Theorem implies also any weak solution is a Brownian motion.

Existence and uniqueness of weak solution

Theorem 2.5.6 *If for each $t > 0$, functions $\mu(x,t)$ and $\sigma(x,t)$ are bounded and continuous then the SDE (2.5.23) has at least one weak solution starting at time s and point x, for all s and x.*

If in addition their partial derivatives with respect to x up to order two are also bounded and continuous, then the SDE (2.5.23)has unique weak solution starting at time s and point x, for all s and x.

Theorem 2.5.7 *If $\sigma(x,t)$ is positive and continuous and for any $T > 0$ there is K_T such that for all $x \in \mathscr{R}$ and $0 \leq t \leq T$,*

$$\big|\mu(x,t)\big| + \big|\sigma(x,t)\big| \leq K_T(1 + |x|),$$

then the SDE (2.5.23) has unique weak solution starting at time s and point x, for all $s \geq 0$ and x.

The proof of the above two theorems can be found in D. Stroock and S.R.S. Varadhan (1979), *Multidimensional Diffusion Processes*, Springer-Verlag.

2.5.7 Martingale Problem and Heat Equation

Let $X(t)$ solve the stochastic differential equation

$$dX(t) = \mu\big(X(t), t\big)dt + \sigma\big(X(t), t\big)dB(t), \quad t \geq 0. \tag{2.5.25}$$

Then $X(t)$ is a Markov process. Suppose $f(x,t) \in C^{2,1}$. Then

$$\begin{aligned}
df\big(X(t),t\big) =& \frac{\partial f\big(X(t),t\big)}{\partial x} dX(t) + \frac{\partial f\big(X(t),t\big)}{\partial t} dt + \frac{1}{2}\frac{\partial^2 f\big(X(t),t\big)}{\partial x^2}\big(dX(t)\big)^2 \\
=& \left(\frac{1}{2}\sigma^2\big(X(t),t\big)\frac{\partial^2 f\big(X(t),t\big)}{\partial x^2} + \mu\big(X(t),t\big)\frac{\partial f\big(X(t),t\big)}{\partial x} + \frac{\partial f\big(X(t),t\big)}{\partial t}\right) dt \\
&+ \sigma\big(X(t),t\big)\frac{\partial f\big(X(t),t\big)}{\partial x} dB(t) \\
=& \left(\mathscr{L}_t f\big(X(t),t\big) + \frac{\partial f}{\partial t}\big(X(t),t\big)\right) dt + \sigma\big(X(t),t\big)\frac{\partial f\big(X(t),t\big)}{\partial x} dB(t),
\end{aligned}$$

where

$$\mathscr{L}_t f(x,t) = \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 f(x,t)}{\partial x^2} + \mu(x,t)\frac{\partial f(x,t)}{\partial x}.$$

So,

$$\begin{aligned}
f\big(X(t),t\big) - f\big(X(0),0\big) =& \int_0^t \left(\mathscr{L}_u f + \frac{\partial f}{\partial t}\right)\big(X(u),u\big)du \\
&+ \int_0^t \sigma\big(X(u),u\big)\frac{\partial f\big(X(u),u\big)}{\partial x} dB(u).
\end{aligned}$$

Example 2.5.9 *If*

$$dX(t) = \alpha X(t)dt + \sigma X(t)dB(t),$$

then

$$\mathscr{L}_t f(x,t) = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2}(x,t) + \alpha x \frac{\partial f}{\partial x}(x,t).$$

Let

$$M_f(t) = f\big(X(t),t\big) - \int_0^t \left(\mathscr{L}_u f + \frac{\partial f}{\partial t}\right)\big(X(u),u\big)du.$$

Then

$$M_f(t) = f\big(X(0),0\big) + \int_0^t \sigma\big(X(u),u\big)\frac{\partial f\big(X(u),u\big)}{\partial x} dB(u)$$

is a local martingale.

If f solves the partial differential equation

$$\mathscr{L}_t f(x,t) + \frac{\partial f}{\partial t}(x,t) = 0,$$

Then $f\big(X(t),t\big)$ is a local martingale. If

$$\mathsf{E}\left[\int_0^T \sigma^2\big(X(u),u\big)\left(\frac{\partial f\big(X(u),u\big)}{\partial x}\right)^2 du\right] < \infty,$$

then $f\big(X(t),t\big)$ is a martingale. In such case,

$$f\big(X(t),t\big) = \mathsf{E}\left[f\big(X(T),T\big)\big|\mathscr{F}_t\right] = \mathsf{E}\left[f\big(X(T),T\big)\big|X(t)\right],$$

$$f\big(x,t\big) = \mathsf{E}\left[f\big(X(T),T\big)\big|X(t)=x\right].$$

Theorem 2.5.8 *Let $f(x,t)$ solve the backward equation*

$$\mathscr{L}_t f(x,t) + \frac{\partial f}{\partial t}(x,t) = 0, \quad \text{with } f(x,T) = g(x).$$

Suppose the following conditions are satisfied

1. *Coefficients are locally Lipschitz in x uniformly in t, that is for every T and N, there is a constant K depending only on T and N such that for all $|x|,|y| \leq N$ and all $0 \leq t \leq T$,*

$$|\mu(x,t) - \mu(y,t)| + |\sigma(x,t) - \sigma(y,t)| \leq K|x-y|,$$

2. *Coefficients satisfy the linear growth condition*

$$|\mu(x,t)| + |\sigma(x,t)| \leq K(1+|x|),$$

3. *$X(0)$ is independent of $(B(t), 0 \leq t \leq T)$, and $\mathsf{E}X^2(0) < \infty$,*

and $\frac{\partial f}{\partial x}(x,t)$ is bounded. Then

$$f\big(x,t\big) = E\left[g\big(X(T)\big)\big|X(t)=x\right].$$

Proof.

$$\begin{aligned}
&\mathsf{E}\left[\int_0^T \sigma^2\big(X(u),u\big)\left(\frac{\partial f\big(X(u),u\big)}{\partial x}\right)^2 du\right]\\
\leq &C\mathsf{E}\left[\int_0^T \sigma^2\big(X(u),u\big)du\right] \leq C\left[1+\int_0^T \mathsf{E}[X^2(u)]du\right]\\
\leq &C\big(1+T\mathsf{E}[X^2(0)]\big) < \infty.
\end{aligned}$$

So, $f\big(X(t),t\big)$ is a martingale.

Theorem 2.5.9 *(Feynman-Kac Formula) Let $X(t)$ be a solution of the SDE (2.5.25) Let $C(x,t)$ denote*

$$C(x,t) = E\left[e^{-\int_t^T r\left(X(u),u\right)du} g\big(X(T)\big)\Big|X(t)=x\right], \tag{2.5.26}$$

for given bounded functions $r(x,t)$ and $g(x)$. Assume that there is a solution to

$$\mathscr{L}_t f(x,t) + \frac{\partial f}{\partial t}(x,t) = r(x,t)f(x,t), \quad \textit{with } f(x,T) = g(x). \tag{2.5.27}$$

Then the solution is unique and $C(x,t)$ is that solution.

(2.5.26) is equivalent to

$$e^{-\int_0^t r\left(X(u),u\right)du} C(x,t) = \mathsf{E}\left[e^{-\int_0^T r\left(X(u),u\right)du} g\big(X(T)\big)\Big|X(t)=x\right].$$

Sketch of the proof. By Itô̂'s formula,

$$\begin{aligned}
df\big(X(t),t\big) =&\left(\mathscr{L}_t f\big(X(t),t\big) + \frac{\partial f}{\partial t}\big(X(t),t\big)\right)dt + \sigma\big(X(t),t\big)\frac{\partial f\big(X(t),t\big)}{\partial x}dB(t)\\
=&r\big(X(t),t\big)f\big(X(t),t\big)dt + \sigma\big(X(t),t\big)\frac{\partial f\big(X(t),t\big)}{\partial x}dB(t).
\end{aligned}$$

Write $h_t = e^{-\int_t^T r\left(X(u),u\right)du}$ and $r_t = r\big(X(t),t\big)$. Then

$$\begin{aligned}
d\left[h_t f\big(X(t),t\big)\right] =&h_t d\left[f\big(X(t),t\big)\right] + f\big(X(t),t\big)dh_t + d[h_t]d\left[f\big(X(t),t\big)\right]\\
=&h_t r\big(X(t),t\big)f\big(X(t),t\big)dt + h_t\sigma\big(X(t),t\big)\frac{\partial f\big(X(t),t\big)}{\partial x}dB(t)\\
&+ f\big(X(t),t\big)h_t\big(-r\big(X(t),t\big)\big)dt\\
=&h_t\sigma\big(X(t),t\big)\frac{\partial f\big(X(t),t\big)}{\partial x}dB(t).
\end{aligned}$$

So, $h_t f\big(X(t), t\big)$ is a martingale. It follows that

$$h_t f\big(X(t), t\big) = \mathsf{E}\left[h_T f\big(X(T), T\big)\big|\mathscr{F}_t\right].$$

So

$$\begin{aligned} f\big(X(t), t\big) =& \mathsf{E}\left[h_t^{-1} h_T f\big(X(T), T\big)\big|\mathscr{F}_t\right] \\ =& \mathsf{E}\left[e^{-\int_t^T r\left(X(u), u\right) du} g\big(X(T)\big)\Big|\mathscr{F}_t\right] \\ =& \mathsf{E}\left[e^{-\int_t^T r\left(X(u), u\right) du} g\big(X(T)\big)\Big|X(t)\right], \end{aligned}$$

by the Markov property. That is

$$f(x, t) = \mathsf{E}\left[e^{-\int_t^T r\left(X(u), u\right) du} g\big(X(T)\big)\Big|X(t) = x\right]. \quad \square$$

Example 2.5.10 *If*

$$dX(t) = \alpha X(t)dt + \sigma X(t)dB(t) = X(t)(\alpha dt + \sigma dB(t)),$$

then

$$\mathscr{L}_t f(x, t) = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2}(x, t) + \alpha x \frac{\partial f}{\partial x}(x, t),$$

and

$$C(x, t) = \mathsf{E}\left[e^{-r(T-t)} g\big(X(T)\big)\Big|X(t) = x\right],$$

is a solution to

$$\frac{\partial f}{\partial t} + \alpha x \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} = rf, \quad \text{with } f(x, T) = g(x).$$

$C(x, t)$ satisfies the Black-Scholes-Merton partial differential equation