Stochastic Differential Equations

A Solution Manual of Selected Exercise Problems

Yan Zeng

Stochastic Differential Equations, 6ed. A Solution Manual of Selected Exercise Problems

Yan Zeng

Version 1.2.1, last revised on 2024-10-20

This book is for sale at https://leanpub.com/oksendal_solman

This version was published on 2023-12-05.

This is a Leanpub book. Leanpub empowers authors and publishers with the Lean Publishing process. Lean Publishing is the act of publishing an in-progress ebook using lightweight tools and many iterations to get reader feedback, pivot until you have the right book and build traction once you do.

Abstract

This is a solution manual of selected exercise problems for the text book *Stochastic Differential Equations (6th Edition)*, by Bernt Øksendal. If you find any typos/errors, please email me at quantsummaries@gmail.com.

I dedicate this solution manual to the teacher of ORIE 768: Selected Topics in Applied Probability, Fall 2003, Cornell University.

It was an eye opener. It has been an intellectual adventure.

Contents

Chapter 1

Introduction

This chapter introduced seven problems arising from six different situations in order to convince readers that stochastic differential equations are an important subject. This approach illustrates the author's method of learning a subject when he does not know anything about it:
1) *In what situations does the subject arise*?
2) *What are its essential features*?
3) *What are the applications and the connections to other fields*?

Stochastic differential equations. The notion of a *stochastic differential equation* arises naturally by introducing randomness into the coefficients of an ordinary differential equation. The question is, *how do we define rigorously and then solve a stochastic differential equation*?

Filtering problems. Suppose $Q(t)$ is a solution of a stochastic differential equation and we do not observe it directly. Instead, we observe a disturbed version of it:

$$Z(s) = Q(s) + \text{“noise”}.$$

This leads to the *filtering problem*: What is the best estimate of $Q(t)$ based on the observations $(Z_s)_{s\leq t}$? In other words, we need to find a procedure for estimating the state of a system which satisfies a “noisy" linear differential equation, based on a series of “noisy" observations.

Boundary value problems. For a large class of semielliptic second order partial differential equations, the corresponding Dirichlet boundary value problem (i.e. finding a continuous function that is harmonic within a region and is equal to a prescribed function on the boundary of the region) can be solved using the solution of a stochastic differential equation.

Optimal stopping problems. The example used to illustrate the *optimal stopping problem* is a method to maximize the expected profit of selling an asset with stochastic price. The solution can be expressed in terms of the solution of a corresponding (free) boundary value problem. It can also be expressed in terms of a set of *variational inequalities*.

Stochastic control problems. This situation is illustrated via an optimal portfolio problem, in which an asset allocation process (the “control variable") needs to be determined to maximize the expected utility of the terminal fortune of a portfolio consisting of a safe asset and a risky asset.

Mathematical finance. The Black-Scholes option pricing formula and the underlying no-arbitrage pricing theory illustrate the applications of stochastic differential equations to finance.

Chapter 2

Some Mathematical Preliminaries

Chapter Summary. The core of the mathematical preliminaries in this chapter consists of answers to two fundamental questions:

• How to ensure the existence of a stochastic process with desired distributional properties?

• How to ensure such a stochastic process, if it exists, enjoys nice path-wise properties (like continuity)?

Theorem 2.1 (Komogorov's extension theorem). *For all* $t_1, \cdots, t_k \in T$, $k \in \mathbb{N}$, *let* $\nu_{t_1,\cdots,t_2}$ *be probability measures on* $\mathbb{R}^{nk}$ *s.t.*

$$v_{t_{\sigma(1)},\cdots,t_{\sigma(k)}}(F_1 \times \cdots \times F_k) = v_{t_1,\cdots,t_k}(F_{\sigma^{-1}(1)} \times \cdots \times F_{\sigma^{-1}(k)})$$

for all permutations σ *on* $\{1, 2, \cdots, k\}$ *and*

$$\nu_{t_1,\cdots,t_k}(F_1 \times \cdots \times F_k) = \nu_{t_1,\cdots,t_k,t_{k+1},\cdots,t_{k+m}}(F_1 \times \cdots \times F_k \times \mathbb{R}^n \times \mathbb{R}^n)$$

for all $m \in \mathbb{N}$, *where the set on the right hand side has a total of* $k + m$ *factors.*

Then there exists a probability space $(\Omega, \mathcal{F}, P)$ *and a stochastic process* $\{X_t\}$ *on* Ω, $X_t : \Omega \to \mathbb{R}^n$, *s.t.*

$$\nu_{t_1,\cdots,t_k}(F_1 \times \cdots \times F_k) = P(X_{t_1} \in F_1, \cdots, X_{t_k} \in F_k),$$

for all $t_i \in T$, $k \in \mathbb{N}$ *and all Borel sets* F_i.

A thorough exposition of the above theorem can be found in Shiryaev [10, Chapter II, §9].

Theorem 2.2 (Kolmogorov's continuity theorem). *Suppose that the process* $X = \{X_t\}_{t\geq 0}$ *satisfies the following condition: For all* $T > 0$ *there exist positive constants* α, β, D *such that*

$$E\left[|X_t - X_s|^\alpha\right] \leq D \cdot |t - s|^{1+\beta};\ 0 \leq s, t \leq T.$$

Then there exists a continuous version of X.

As a result of the two above theorems, Brownian motion can be constructed and can have a version with continuous sample paths.

► **2.1.**
a)

Proof. Necessity is obvious. For sufficiency, we note for any open set $U \in \mathbb{R}$,

$$X^{-1}(U) = \cup_{a_k \in U} X^{-1}(a_k)$$

is a countable union of elements of $\mathcal{F}$ and is therefore in $\mathcal{F}$. □

b)
c)
d)

Proof. Properties (b)-(d) are special cases of a series of theorems in real analysis (Monotone Convergence Theorem, Dominated Convergence Theorem, and Bounded Convergence Theorem, respectively). See Durrett [3] for details. □

▶ **2.2.**
a)

Proof. $F = P(X \leq x) \in [0, 1]$ is obvious. Using Bounded Convergence Theorem, we have

$$\lim_{x\to\infty} F(x) = E\left[\lim_{x\to\infty} 1_{\{X\leq x\}}\right] = 1, \quad \lim_{x\to-\infty} F(x) = E\left[\lim_{x\to-\infty} 1_{\{X\leq x\}}\right] = 0.$$

For monotonicity, we note for $a \leq b$,

$$F(b) - F(a) = P(a < X \leq b) \geq 0.$$

And for right continuity, we note for $\varepsilon > 0$,

$$0 \leq F(x+\varepsilon) - F(x) = P(x < X \leq x+\varepsilon) = E\left[1_{\{x<X\leq x+\varepsilon\}}\right] \to 0$$

as $\varepsilon \to 0$ due to the Bounded Convergence Theorem. □

b)

Proof. When g is a simple function in the form of $\sum_{i=1}^n g_i 1_{(a_i,b_i]}(x)$,

$$E[g(X)] = \sum_{i=1}^n g_i[F(b_i) - F(a_i)] = \int_{-\infty}^{\infty} g(x)dF(x).$$

Then using the property that any Lebesgue integrable function can be approximated by a series of simple functions and the Dominated Convergence Theorem, we can show

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)dF(x)$$

holds for general g. □

c)

Solution. For any $x < 0$, clearly $P(B_t^2 \leq x) = 0$. For any $x \geq 0$, we have

$$P\left(B_t^2 \leq x\right) = P\left(-\sqrt{x} \leq B_t \leq \sqrt{x}\right) = F\left(\sqrt{x}\right) - F\left(-\sqrt{x}\right).$$

so that

$$\frac{d}{dx}P\left(B_t^2 \leq x\right) = p\left(\sqrt{x}\right)\frac{1}{2\sqrt{x}} + p\left(-\sqrt{x}\right)\frac{1}{2\sqrt{x}} = \frac{p\left(\sqrt{x}\right)}{\sqrt{x}}$$

In conclusion, the density function $f(x)$ of B_t^2 is

$$f(x) = 1_{\{x\geq 0\}}\frac{1}{\sqrt{2\pi t x}}\exp\left(-\frac{x}{2t}\right).$$

□

▶ 2.3.

Proof. First, since $\emptyset \in \mathcal{H}_i$ $(\forall i \in I)$, $\emptyset \in \mathcal{H}$. Second, $F \in \mathcal{H} \Rightarrow F \in \mathcal{H}_i$, $\forall i \in I \Rightarrow F^C \in \mathcal{H}_i$, $\forall i \in I \Rightarrow F^C \in \mathcal{H}$. Finally,

$$A_1, A_2, \cdots \in \mathcal{H} \Rightarrow A_1, A_2, \cdots \in \mathcal{H}_i,\ \forall i \in I \Rightarrow A := \bigcup_{k=1}^{\infty} A_k \in \mathcal{H}_i,\ \forall i \in I \Rightarrow A \in \mathcal{H}$$

□

▶ 2.4. a)

Proof. Let $A = \{\omega : |X| \geq \lambda\}$, then

$$E[|X|^p] \geq E[1_A|X|^p] \geq E[1_A\lambda^p] = \lambda^p P(A) = \lambda^p P(|X| \geq \lambda).$$

□

b)

Proof. Let $A = \{\omega : |X| \geq \lambda\}$, then

$$M \geq E[1_A \exp(k|X|)] \geq e^{k\lambda}P(A)$$

so that

$$P(|X| \geq \lambda) \leq Me^{-k\lambda} \text{ for all } \lambda \geq 0.$$

□

▶ 2.5.

Proof. This is a straightforward application of Bounded Convergence Theorem. Just follow the logic of the hint. □

▶ 2.6.

Proof. See Durrett [3, p.65] for details. □

▶ 2.8.
a)

Proof. B_t has mean 0 and variance t. Applying equation (2.2.3) gives the above formula. $\square$

b)

Proof.

$$E[e^{iuB_t}] = \sum_{k=0}^{\infty} \frac{i^k}{k!} E[B_t^k] u^k = \exp\left(-\frac{1}{2}u^2 t\right) = \sum_{k=0}^{\infty} \frac{1}{k!}\left(-\frac{t}{2}\right)^k u^{2k}.$$

Matching the coefficients of even terms, we have

$$\frac{i^{2k}}{(2k)!} E[B_t^{2k}] u^{2k} = \frac{1}{k!}\left(-\frac{t}{2}\right)^k u^{2k} \Longrightarrow E[B_t^{2k}] = \frac{\frac{1}{k!}(-\frac{t}{2})^k}{\frac{(-1)^k}{(2k)!}} = \frac{(2k)!}{2^k \cdot k!} t^k.$$

$\square$

c)

Proof. It can be made rigorous. See, for example, Durrett [3], Appendix A.5. $\square$

d)

Proof.

$$\begin{aligned} E^x[|B_t - B_s|^4] &= \sum_{i=1}^{n} E^x\left[\left(B_t^{(i)} - B_s^{(i)}\right)^4\right] + \sum_{i \neq j} E^x\left[\left(B_t^{(i)} - B_s^{(i)}\right)^2 \left(B_t^{(j)} - B_s^{(j)}\right)^2\right] \\ &= n \cdot \frac{4!}{2! \cdot 4} \cdot (t-s)^2 + n(n-1)(t-s)^2 \\ &= n(n+2)(t-s)^2. \end{aligned}$$

$\square$

▶ 2.15.

Proof. Since $B_t - B_s \perp \mathcal{F}_s := \sigma(B_u : u \leq s)$, $U(B_t - B_s) \perp \mathcal{F}_s$. Note $U(B_t - B_s) \overset{d}{=} N(0, t-s)$, as seen by the characteristic functions. $\square$

Chapter 3

Itô Integrals

Chapter Summary. The first key component of this chapter is to use Itô isometry to extend the definition of stochastic integral from elementary functions to functions in $\mathcal{V} = \mathcal{V}(S,T) = \{f(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}, f(t,\omega)$ is $\mathcal{B} \times \mathcal{F}$-measurable, f is $\mathcal{F}_t$-adapted, $E\left[\int_S^T f(t,\omega)^2 dt\right] < \infty\}$. The three-step approximation is

Step 1. Let $g \in \mathcal{V}$ be bounded and $g(\cdot,\omega)$ continuous for each ω. Then there exists elementary functions $\phi_n \in \mathcal{V}$ such that

$$E\left[\int_S^T (g-\phi_n)^2 dt\right] \to 0, \text{as } n \to \infty.$$

Here ϕ_n is defined as $\phi_n(t,\omega) = \sum_j g(t_j,\omega) \cdot 1_{[t_j,t_{j+1})}(t)$.

Step 2. Let $h \in \mathcal{V}$ be bounded. Then there exist bounded functions $g_n \in \mathcal{V}$ such that $g_n(\cdot,\omega)$ is continuous for all ω and n, and

$$E\left[\int_S^T (h-g_n)^2 dt\right] \to 0.$$

Here g_n is constructed via convolution with a mollifier ψ_n:

$$g_n(t,\omega) = \int_0^t \psi_n(s-t)h(s,\omega)ds,$$

where each ψ_n is a non-negative, continuous function on $\mathbb{R}$ such that $\psi_n = 0$ for $x \in (-\infty, -\frac{1}{n}] \bigcup [0,\infty)$ and $\int_{-\infty}^{\infty} \psi_n(x) = 1$.

Step 3. Let $f \in \mathcal{V}$, Then there exists a sequence $\{h_n\} \subset \mathcal{V}$ such that h_n is bounded for each n and

$$E\left[\int_S^T (f-h_n)^2 dt\right] \to 0 \text{as } n \to \infty.$$

Here

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

Long story short, the insight is that $\mathcal{V}$ can be obtained by the closure of simpler subspaces, where the closure is in the L^2-norm:

$$\begin{aligned}
\mathcal{V} &= \text{closure(bounded elements of } \mathcal{V}) \\
&= \text{closure(bounded continuous elements of } \mathcal{V}) \\
&= \text{closure(elementary functions in } \mathcal{V})
\end{aligned}$$

The second key component of this chapter is about the sample path properties of stochastic integral. By Doob's martingale inequality and the Borel-Cantelli lemma, we have the following result on path continuity and uniform estimate:

Theorem 3.1. *Let $f \in \mathcal{V}(0,T)$. Then there exists a t-continuous version of*

$$M_t = \int_0^t f(s,\omega)dB_s(\omega);\ 0 \le t \le T.$$

Moreover, M_t is a martingale w.r.t. $\mathcal{F}_t$ and

$$P\left[\sup_{0\le t\le T} |M_t| \ge \lambda\right] \le \frac{1}{\lambda^2} \cdot E\left[\int_0^T f(s,\omega)^2 ds\right];\ \lambda, T > 0.$$

The final component of this chapter is about extensions of the Itô integral. The extensions include

a) measurability of the integrand and martingale property of the integrator with respect to a generic filtration, not just the filtration generated by Brownian motion. This will allow us to define the multi-dimensional Itô integral.

b) integrability of the integrand in a.e. sense, not in $L^2(\Omega)$ sense. The resulted integral is no longer a martingale, but a local martingale.

c) a comparison of the Stratonovich integral vs. the Itô integral.

• The Stratonovich integral has the advantage of leading to ordinary chain rule formulas under a transformation (change of variable), making it natural to use in connection with SDEs on manifolds. It also makes the solutions of a family of SDEs whose integrators converge uniformly to Brownian motion converge uniformly to the solution of a SDE whose integrator is Brownian motion.

• On the other hand, the specific feature of the Itô model of "not looking into the future" seems to be a reason for choosing the Itô interpretation in many cases. And the martingale property of the Itô integral gives an important computational advantage.

Below are the definitions and notations used frequently in the book.

Definition 3.1. *Let $B_t(\omega)$ be n-dimensional Brownian motion. We define $\mathcal{F}_t = \mathcal{F}_t^{(n)}$ to be the σ-algebra generated by the random variables $\{B_i(s)\}_{1\le i\le n, 0\le s\le t}$.*

Definition 3.2. *Let $\mathcal{V} = \mathcal{V}(S,T)$ be the class of functions*

$$f(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}$$

such that

(i) $(t,\omega) \to f(t,\omega)$ is $\mathcal{B} \times \mathcal{F}$-measurable, where $\mathcal{B}$ denotes the Borel σ-algebra on $[0,\infty)$.

(ii) $f(t,\omega)$ is $\mathcal{F}_t$-adapted.

(iii) $E\left[\int_S^T f(t,\omega)^2 dt\right] < \infty$.

Definition 3.3. *Let B be n-dimensional Brownian motion. Then $\mathcal{V}_{\mathcal{H}}^{m\times n}(S,T)$ denotes the set of $m\times n$ matrices $v=[v_{ij}(t,\omega)]$ where each entry $v_{ij}(t,\omega)$ satisfies (i) and (iii) of Definition 3.2 and the following condition (ii)':*

(ii)' There exists an increasing family of σ-algebra $\mathcal{H}_t : t\geq 0$ such that

a) B_t is a martingale with respect to $\mathcal{H}_t$ and

b) f_t is $\mathcal{H}_t$-adapted.

If $\mathcal{H}=\mathcal{F}^{(n)}=\{\mathcal{F}_t^{(n)}\}_{t\geq 0}$, we write $\mathcal{V}^{m\times n}(S,T)$ and define $\mathcal{V}^{m\times n}=\bigcap_{T>0}\mathcal{V}^{m\times n}(0,T)$.

The motivation of the extension from $\mathcal{V}(S,T)$ to $\mathcal{V}_H^{m\times n}(S,T)$ is to generalize the Itô integral so that f_t can depend on more than $\mathcal{F}_t$ as long as B_t remains a martingale with respect to the "history" of f_s; $s\leq t$.

Definition 3.4. *$\mathcal{W}_H(S,T)$ denotes the class of processes $f(t,\omega)\in\mathbb{R}$ satisfying (i) of Definition 3.2, (ii)' of Definition 3.3, and the following condition (iii)':*

(iii)' $P\left(\int_S^T f(s,\omega)^2ds<\infty\right)=1$.

Similarly to the notation for $\mathcal{V}$ we put $\mathcal{W}_H=\bigcap_{T>0}\mathcal{W}_H(0,T)$ and in the matrix case we write $\mathcal{W}_H^{m\times n}(S,T)$ etc. If $\mathcal{H}=\mathcal{F}^{(n)}$ we write $\mathcal{W}(S,T)$ instead of $\mathcal{W}_{\mathcal{F}^{(n)}}(S,T)$ etc.

Itô integral can be defined on $\mathcal{W}_H$ via convergence in probability (not via the Itô isometry). However, the resulted integral is not in general a martingale, but a *local* martingale.

▶ 3.2.

Proof. WLOG, we assume $t=1$, then

$$\begin{aligned}
B_1^3 &= \sum_{j=1}^n\left(B_{j/n}^3-B_{(j-1)/n}^3\right)\\
&= \sum_{j=1}^n\left[\left(B_{j/n}-B_{(j-1)/n}\right)^3+3B_{(j-1)/n}B_{j/n}\left(B_{j/n}-B_{(j-1)/n}\right)\right]\\
&= \sum_{j=1}^n\left(B_{j/n}-B_{(j-1)/n}\right)^3+\sum_{j=1}^n 3B_{(j-1)/n}^2\left(B_{j/n}-B_{(j-1)/n}\right)\\
&\quad +\sum_{j=1}^n 3B_{(j-1)/n}\left(B_{j/n}-B_{(j-1)/n}\right)^2\\
&:= I+II+III
\end{aligned}$$

By Problem EP1-1 and the continuity of Brownian motion.

$$I\leq\left[\sum_{j=1}^n\left(B_{j/n}-B_{(j-1)/n}\right)^2\right]\max_{1\leq j\leq n}\left|B_{j/n}-B_{(j-1)/n}\right|\to 0\quad a.s.$$

To argue $II\to 3\int_0^1 B_t^2dB_t$ as $n\to\infty$, it suffices to show $E\left[\int_0^1(B_t^2-B_t^{(n)})^2dt\right]\to 0$, where $B_t^{(n)}=\sum_{j=1}^n B_{(j-1)/n}^2 1_{\{(j-1)/n<t\leq j/n\}}$. Indeed,

$$E\left[\int_0^1\left|B_t^2-B_t^{(n)}\right|^2dt\right]=\sum_{j=1}^n\int_{(j-1)/n}^{j/n}E\left[\left(B_t^2-B_{(j-1)/n}^2\right)^2\right]dt$$

We note $\left(B_t^2 - B_{(j-1)/n}^2\right)^2$ is equal to

$$\left(B_t - B_{\frac{j-1}{n}}\right)^4 + 4\left(B_t - B_{\frac{j-1}{n}}\right)^3 B_{\frac{j-1}{n}} + 4\left(B_t - B_{\frac{j-1}{n}}\right)^2 B_{\frac{j-1}{n}}^2$$

so $E\left[\left(B_{(j-1)/n}^2 - B_t^2\right)^2\right] = 3(t-(j-1)/n)^2 + 4(t-(j-1)/n)(j-1)/n$, and

$$\int_{\frac{j-1}{n}}^{\frac{j}{n}} E\left[\left(B_{\frac{j-1}{n}}^2 - B_t^2\right)^2\right] dt = \frac{2j+1}{n^3}$$

Hence $E\left[\int_0^1 \left(B_t - B_t^{(n)}\right)^2 dt\right] = \sum_{j=1}^n \frac{2j-1}{n^3} \to 0$ as $n \to \infty$.

To argue $III \to 3\int_0^1 B_t dt$ as $n \to \infty$, it suffices to prove

$$\sum_{j=1}^n B_{(j-1)/n}\left(B_{j/n} - B_{(j-1)/n}\right)^2 - \sum_{j=1}^n B_{(j-1)/n}\left(\frac{j}{n} - \frac{j-1}{n}\right) \to 0 \quad a.s.$$

By looking at a subsequence, we only need to prove the L^2-convergence. Indeed,

$$\begin{aligned}
& E\left(\sum_{j=1}^n B_{(j-1)/n}\left[\left(B_{j/n} - B_{(j-1)/n}\right)^2 - \frac{1}{n}\right]\right)^2 \\
= & \sum_{j=1}^n E\left(B_{(j-1)/n}^2\left[\left(B_{j/n} - B_{(j-1)/n}\right)^2 - \frac{1}{n}\right]^2\right) \\
= & \sum_{j=1}^n \frac{j-1}{n} E\left[\left(B_{j/n} - B_{(j-1)/n}\right)^4 - \frac{2}{n}\left(B_{j/n} - B_{(j-1)/n}\right)^2 + \frac{1}{n^2}\right] \\
= & \sum_{j=1}^n \frac{j-1}{n}\left(3\frac{1}{n^2} - 2\frac{1}{n^2} + \frac{1}{n^2}\right) \\
= & \sum_{j=1}^n \frac{2(j-1)}{n^3} \to 0
\end{aligned}$$

as $n \to \infty$. This completes our proof.

□

▶ 3.9.

Proof. We first note that

$$\begin{aligned}
& \sum_j B_{\frac{t_j+t_{j+1}}{2}}\left(B_{t_{j+1}} - B_{t_j}\right) \\
= & \sum_j \left[B_{\frac{t_j+t_{j+1}}{2}}\left(B_{t_{j+1}} - B_{\frac{t_j+t_{j+1}}{2}}\right) + B_{t_j}\left(B_{\frac{t_j+t_{j+1}}{2}} - B_{t_j}\right)\right] + \sum_j \left(B_{\frac{t_j+t_{j+1}}{2}} - B_{t_j}\right)^2.
\end{aligned}$$

The first term converges in $L^2(P)$ to $\int_0^T B_t dB_t$. For the second term, we note

$$\begin{aligned}
& E\left[\left(\sum_j \left(B_{\frac{t_j+t_{j+1}}{2}} - B_{t_j}\right)^2 - \frac{t}{2}\right)^2\right] \\
= \; & E\left[\left(\sum_j \left(B_{\frac{t_j+t_{j+1}}{2}} - B_{t_j}\right)^2 - \sum_j \frac{t_{j+1}-t_j}{2}\right)^2\right] \\
= \; & \sum_{j,k} E\left[\left(\left(B_{\frac{t_j+t_{j+1}}{2}} - B_{t_j}\right)^2 - \frac{t_{j+1}-t_j}{2}\right)\left(\left(B_{\frac{t_k+t_{k+1}}{2}} - B_{t_k}\right)^2 - \frac{t_{k+1}-t_k}{2}\right)\right] \\
= \; & \sum_j E\left[\left(B^2_{\frac{t_{j+1}-t_j}{2}} - \frac{t_{j+1}-t_j}{2}\right)^2\right] \\
= \; & \sum_j 2 \cdot \left(\frac{t_{j+1}-t_j}{2}\right)^2 \\
\leq \; & \frac{T}{2} \max_{1 \leq j \leq n} |t_{j+1} - t_j| \to 0,
\end{aligned}$$

since $E\left[(B_t^2 - t)^2\right] = E\left[B_t^4 - 2tB_t^2 + t^2\right] = 3E\left[B_t^2\right]^2 - 2t^2 + t^2 = 2t^2$. So

$$\sum_j B_{\frac{t_j+t_{j+1}}{2}}\left(B_{t_{j+1}} - B_{t_j}\right) \to \int_0^T B_t dB_t + \frac{T}{2} = \frac{1}{2}B_T^2 \quad \text{in } L^2(P).$$

□

► 3.10.

Proof. According to the result of Exercise 3.9., it suffices to show

$$E\left[\left|\sum_j f(t_j,\omega)\Delta B_j - \sum_j f(t'_j,\omega)\Delta B_j\right|\right] \to 0.$$

Indeed, note

$$\begin{aligned}
& E\left[\left|\sum_j f(t_j,\omega)\Delta B_j - \sum_j f(t'_j,\omega)\Delta B_j\right|\right] \\
\leq \; & \sum_j E\left[\left|f(t_j) - f(t'_j)\right| |\Delta B_j|\right] \\
\leq \; & \sum_j \sqrt{E\left[\left|f(t_j) - f(t'_j)\right|^2\right] E\left[|\Delta B_j|^2\right]} \\
\leq \; & \sum_j \sqrt{K}\left|t_j - t'_j\right|^{\frac{1+\epsilon}{2}} \left|t_j - t'_j\right|^{\frac{1}{2}} \\
= \; & \sqrt{K}\sum_j \left|t_j - t'_j\right|^{1+\frac{\epsilon}{2}} \\
\leq \; & T\sqrt{K} \max_{1 \leq j \leq n} \left|t_j - t'_j\right|^{\frac{\epsilon}{2}} \\
\to \; & 0.
\end{aligned}$$

□

▶ 3.11.

Proof. Assume W is continuous, then by the Bounded Convergence Theorem,

$$\lim_{s \to t} E\left[\left(W_t^{(N)} - W_s^{(N)}\right)^2\right] = 0.$$

Since W_s and W_t are independent and identically distributed, so are $W_s^{(N)}$ and $W_t^{(N)}$. Hence

$$\begin{aligned} E\left[\left(W_t^{(N)} - W_s^{(N)}\right)^2\right] &= E\left[\left(W_t^{(N)}\right)^2\right] - 2E\left[W_t^{(N)}\right]E\left[W_s^{(N)}\right] + E\left[\left(W_s^{(N)}\right)^2\right] \\ &= 2E\left[\left(W_t^{(N)}\right)^2\right] - 2E\left[W_t^{(N)}\right]^2. \end{aligned}$$

Since the RHS=2Var$\left(W_t^{(N)}\right)$ is independent of s, we must have RHS=0, i.e. $W_t^{(N)} = E\left[W_t^{(N)}\right]$ a.s.. Let $N \to \infty$ and apply the Dominated Convergence Theorem to $E\left[W_t^{(N)}\right]$, we get $W_t = 0$. Therefore $W_\cdot \equiv 0$. □

▶ 3.18.

Proof. If $t > s$, then

$$E\left[\frac{M_t}{M_s}\middle| \mathcal{F}_s\right] = E\left[e^{\sigma(B_t - B_s) - \frac{1}{2}\sigma^2(t-s)}\middle| \mathcal{F}_s\right] = \frac{E[e^{\sigma B_{t-s}}]}{e^{\frac{1}{2}\sigma^2(t-s)}} = 1$$

The second equality is due to the fact $B_t - B_s$ is independent of $\mathcal{F}_s$. □

Chapter 4

The Itô Formula and the Martingale Representation Theorem

Chapter Summary. To derive the 1-dimensional **Itô formula**, i.e. the SDE for $Y_t = g(t, X_t)$ where X_t is a 1-dimensional Itô process with $dX_t = u_t dt + v_t dB_t$ and $g(t,x) \in C^2([0,\infty) \times \mathbb{R})$, the intuition is to consider the Taylor expansion of Y and study the convergence property of each term in its expansion.

More specifically, let $t_0 = 0$ and $t_n = t$, the Taylor expansion of Y_t is

$$\begin{aligned} Y_t - Y_0 &= g(t, X_t) - g(0, X_0) \\ &= \sum_{j=0}^{n-1} \Big[g(t_{j+1}, X_{t_{j+1}}) - g(t_j, X_{t_j}) \Big] \\ &= \sum_j \frac{\partial g}{\partial t}(t_j, X_{t_j})\Delta t_j + \sum_j \frac{\partial g}{\partial x}(t_j, X_{t_j})\Delta X_j + \frac{1}{2}\sum_j \frac{\partial^2 g}{\partial t^2}(t_j, X_{t_j})(\Delta t_j)^2 \\ &\quad + \sum_j \frac{\partial^2 g}{\partial t \partial x}(t_j, X_{t_j})(\Delta t_j)(\Delta X_j) + \frac{1}{2}\sum_j \frac{\partial^2 g}{\partial x^2}(t_j, X_{t_j})(\Delta X_j)^2 + \sum_j R_j, \end{aligned}$$

where $\Delta t_j = t_{j+1} - t_j$, $\Delta X_j = X_{t_{j+1}} - X_{t_j}$, and $R_j = o(|\Delta t_j|^2 + |\Delta X_j|^2)$. Then assuming g is nice enough, we have

$$\begin{aligned} &\sum_j \frac{\partial g}{\partial t}(t_j, X_{t_j})\Delta t_j \to \int_0^t \frac{\partial g}{\partial t}(s, X_s) ds \\ &\sum_j \frac{\partial g}{\partial x}(t_j, X_{t_j})\Delta X_j \to \int_0^t \frac{\partial g}{\partial x}(s, X_s) dX_s \\ &\frac{1}{2}\sum_j \frac{\partial^2 g}{\partial x^2}(t_j, X_{t_j})(\Delta X_j)^2 \to \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(s, X_s) v_s^2 ds \\ &\text{other terms} \to 0 \end{aligned}$$

This gives the 1-dimensional Itô formula:

$$dY_t = dg(t, X_t) = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t)v_t^2 dt.$$

Remark 4.1. *Note that it is enough that $g(t,x)$ is C^2 on $[0,\infty)\times U$, if $U \subset \mathbb{R}$ is an open set such that $X_t(\omega) \in U$ for all $t \geq 0$, $\omega \in \Omega$. Moreover, it is sufficient that $g(t,x)$ is C^1 w.r.t. t and C^2 w.r.t. x.*

When Itô formula is extended to multi-dimension, this condition on g will allow Itô formula to be applied to $g(t,x) = |x|$ for $n \geq 2$ since B_t never hits the origin a.s. when $n \geq 2$.

The intuition of **Martingale Representation Theorem** can be incrementally built in the following steps:

Step 1. For a dense subset $\{t_i\}_{i=1}^{\infty}$ of $[0,T]$, the set of random variables $\{\phi(B_{t_1},\cdots,B_{t_n}) : t_i \in [0,T], \phi \in C_0^{\infty}(\mathbb{R}^n), n = 1,2,\cdots\}$ is dense in $L^2(\mathcal{F}_T, P)$. This insight is based on a non-trivial result from the martingale convergence theorem:

$$g = E[g|\mathcal{F}_T] = \lim_{n\to\infty} E[g|\sigma(B_{t_1},\cdots,B_{t_n})] \text{ a.s. and in } L^2(\mathcal{F}_T, P)$$

Step 2. The linear span of random variables of the type

$$\exp\left\{\int_0^T h(t)dB_t(\omega) - \frac{1}{2}\int_0^T h^2(t)dt\right\};\ h \in L^2[0,T]$$

is dense in $L^2(\mathcal{F}_T, P)$. Here we have relied on the inverse Fourier transform theorem

$$\phi(x) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} \widehat{\phi}(y)e^{ix\cdot y}dy$$

to convert the orthogonality to functions of the form $\exp\{\sum_i \lambda_i B_{t_i}\}$ to functions of the form $\phi(B_{t_1},\cdots,B_{t_n})$.

Step 3. The Itô representation theorem: for any $F \in L^2\left(\mathcal{F}_T^{(n)}, P\right)$, there exists a unique $f(t,\omega) \in \mathcal{V}^n(0,T)$ such that

$$F(\omega) = E[F] + \int_0^T f(t,\omega)dB(t).$$

We can see this is a direct consequence of Step 2 (modulo some technicalities), since

$$Z_t = \exp\left\{\int_0^T h(t)dB_t(\omega) - \frac{1}{2}\int_0^T h^2(t)dt\right\}$$

satisfies the SDE $dZ_t = h(t)Z_t dB_t$.

Step 4. The martingale representation theorem. The connection with the Itô representation theorem is that for $0 \leq t_1 < t_2$, we have

$$\begin{aligned} & M_{t_1} = E[M_0] + \int_0^{t_1} f^{(t_1)}(s,\omega)dB_s(\omega) = E[M_{t_2}|\mathcal{F}_{t_1}] = E[M_0] + \int_0^{t_1} f^{(t_2)}(s,\omega)dB_s(\omega) \\ \Longrightarrow\quad & f^{(t_1)}(s,\omega) = f^{(t_2)}(s,\omega) \text{ for a.a. } (s,\omega) \in [0,t_1]\times\Omega \end{aligned}$$

So we can define $f(s,\omega)$ for a.s. $s \in [0,\infty)\times\Omega$ by setting

$$f(s,\omega) = f^{(N)}(s,\omega), \text{ if } s \in [0,N].$$

► **4.4.** (Exponential martingales)

Proof. For part a), use Theorem 4.1.2 by setting $g(t,x) = e^x$ and

$$X_t = \int_0^t \theta(s,\omega)dB(s) - \frac{1}{2}\int_0^t \theta^2(s,\omega)ds,$$

so that

$$\begin{aligned} dZ_t = dg(t,X_t) &= \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\cdot(dX_t)^2 \\ &= 0 + e^{X_t}dX_t + \frac{1}{2}e^{X_t}\cdot\theta^2(t,\omega)dt \\ &= Z_t\theta(t,\omega)dB(t). \end{aligned}$$

For part b), it comes from the fundamental property of Itô integral, i.e. Itô integral preserves martingale property for integrands in $\mathcal{V}$ (Corollary 3.2.6). □

▶ 4.5.

Proof.

$$B_t^k = \int_0^t kB_s^{k-1}dB_s + \frac{1}{2}k(k-1)\int_0^t B_s^{k-2}ds$$

Take expectation of both sides and use the fact that $\int_0^t kB_s^{k-1}dB_s$ is a martingale, we have

$$\beta_k(t) = \frac{k(k-1)}{2}\int_0^t \beta_{k-2}(s)ds.$$

This gives $E[B_t^4]$ and $E[B_t^6]$. For part b), prove by induction. □

▶ 4.6.

Proof. For part a), use Theorem 4.1.2 by setting $g(t,x) = e^x$ and

$$Y_t = ct + \alpha B_t$$

so that

$$dX_t = de^{Y_t} = 0 + e^{Y_t}dY_t + \frac{1}{2}e^{Y_t}\cdot(dY_t)^2 = X_t(cdt + \alpha dB_t + \frac{1}{2}\alpha^2 dt) = \left(c + \frac{1}{2}\alpha^2\right)X_t dt + \alpha X_t dB_t.$$

For part b), apply Theorem 4.1.2 with $g(t,x) = e^x$ and $X_t = ct + \sum_{j=1}^n \alpha_j B_j(t)$. Note $\sum_{j=1}^n \alpha_j B_j$ is a BM, up to a constant coefficient. □

▶ 4.7.

Proof. For part a), $v \equiv I_{n\times n}$. For part b), use integration by parts formula (Exercise 4.3.), we have

$$X_t^2 = X_0^2 + 2\int_0^t X_s dX + \int_0^t |v_s|^2 ds = X_0^2 + 2\int_0^t X_s v_s dB_s + \int_0^t |v_s|^2 ds.$$

So $M_t := X_t^2 - \int_0^t |v_s|^2 ds = X_0^2 + 2\int_0^t X_s v_s dB_s$ is a local martingale. To show it is a martingale, let C be a bound for $|v|$, then

$$\begin{aligned} E\left[\int_0^t |X_s v_s|^2 ds\right] &\leq C^2 E\left[\int_0^t |X_s|^2 ds\right] \\ &= C^2 \int_0^t E\left[\left|\int_0^s v_u dB_u\right|^2\right] ds \\ &= C^2 \int_0^t E\left[\int_0^s |v_u|^2 du\right] ds \\ &\leq \frac{C^4 t^2}{2}. \end{aligned}$$

This shows $X_t v_t \in \mathcal{V}(0,T)$. So by Corollary 3.2.6, M_t is a martingale. □

▶ **4.12.**

Proof. Let $Y_t = \int_0^t u(s,\omega)ds = X_t - \int_0^t v(s,\omega)dB_s$. Then Y is a continuous $\{\mathcal{F}_t^{(n)}\}$-martingale with finite variation. On one hand,

$$\langle Y\rangle_t = \lim_{\Delta t_k \to 0} \sum_{t_k \leq t} |Y_{t_{k+1}} - Y_{t_k}|^2 \leq \lim_{\Delta t_k \to 0} (\text{total variation of } Y \text{ on } [0,t]) \cdot \max_{t_k} |Y_{t_{k+1}} - Y_{t_k}| = 0.$$

On the other hand, integration by parts formula yields

$$Y_t^2 = 2\int_0^t Y_s dY_s + \langle Y\rangle_t.$$

So Y_t^2 is a local martingale. If $(T_n)_n$ is a localizing sequence of stopping times, by Fatou's lemma,

$$E[Y_t^2] \leq \lim_n E[Y_{t\wedge T_n}^2] = E[Y_0^2] = 0.$$

So $Y_\cdot \equiv 0$. Take derivative, we conclude $u = 0$ for a.a. $(s,\omega) \in [0,\infty) \times \Omega$. □

▶ **4.16.**

Proof. For part a), use Jensen's inequality for conditional expectations.

For part b),

(i) Apply Itô's formula to B_t^2, we have

$$d(B_t^2) = 2B_t dB_t + dt.$$

So $Y = B^2(T) = 2\int_0^T B_t dB_t + T$. By the martingale property of $\int_0^t B_s dB_s$, we conclude $M_t = E[Y|\mathcal{F}_t] = T + 2\int_0^t B_s dB_s$. So $g(t,\omega) = 2B_t$.

(ii) By Itô's formula and integration-by-parts formula, we have

$$B_T^3 = \int_0^T 3B_s^2 dB_s + 3\int_0^T B_s ds = 3\int_0^T B_s^2 dB_s + 3\left(B_T T - \int_0^T s dB_s\right).$$

So

$$M_t = E[B_T^3|\mathcal{F}_t] = 3\int_0^t B_s^2 dB_s + 3TB_t - 3\int_0^t s dB_s = \int_0^t 3(B_s^2 + (T-s))dB_s$$

Therefore $g(t,\omega) = 3(B_t^2 + (T-t))$.

(iii) We follow the hint that $Z_t = \exp\left(\sigma B(t) - \frac{1}{2}\sigma^2 t\right)$ is a martingale, then

$$M_t = E\left[\exp(\sigma B_T)|\mathcal{F}_t\right] = E\left[\exp\left(\sigma B_T - \frac{1}{2}\sigma^2 T\right)\middle| \mathcal{F}_t\right]\exp\left(\frac{1}{2}\sigma^2 T\right) = Z_t \exp\left(\frac{1}{2}\sigma^2 T\right).$$

Since Z solves the SDE $dZ_t = Z_t \sigma dB_t$ with $Z_0 = 1$, we have

$$M_t = \left(1 + \int_0^t Z_s \sigma dB_s\right)\exp\left(\frac{1}{2}\sigma^2 T\right) = \exp\left(\frac{1}{2}\sigma^2 T\right) + \int_0^t \sigma \exp\left(\sigma B_s + \frac{1}{2}\sigma^2 (T-s)\right) dB_s.$$

This shows

$$g(t,\omega) = \sigma \exp\left(\sigma B_t + \frac{\sigma^2}{2}(T-t)\right).$$

□