Time Series Analysis

Nonstationary and Noninvertible Distribution Theory

Katsuto Tanaka

Chapter 2   Stochastic Calculus in Mean Square

As was partly presented in Chapter 1, stochastic integrals often appear as limits in the sense of convergence in distribution, even if we deal with discrete-time processes. This chapter discusses three types of stochastic integrals defined in the mean square (m.s.) sense; one is the m.s. Riemann integral, another the m.s. Riemann-Stieltjes integral and the other the Ito integral. We naturally require that stochastic processes belong to the space L_2 – the space of random variables with finite second moment.

Among important stochastic processes are the (integrated) Brownian motion, Brownian bridge and Ornstein-Uhlenbeck processes; we examine various properties of their integrals. The so-called Ito calculus is also introduced to deal with stochastic differential equations.

The space L_2, however, is not always suitable when we discuss weak convergence of statistics. For that purpose we shall normally need to define another space and give another definition of stochastic integrals, which will be discussed in the next chapter.

2.1. The space L_2 of random variables

Let us denote as L_2 the space of scalar random variables defined on a common probability space with finite second moment, where we do not distinguish between any two random variables X and Y for which $P(X = Y) = 1$. Let $\{X_n\}$ be a sequence in L_2 and suppose that $E[(X_n - X)^2] \to 0$ as $n \to \infty$. Since the convergence to 0 is possible only if $E[(X_n - X)^2]$ is finite from some value of n on, X also belongs to L_2.

Definition of m.s. convergence. *A sequence $\{X_n\}$ in L_2 is said to converge in mean square (m.s.) to X in L_2 if $E[(X_n - X)^2] \to 0$ as $n \to \infty$, which we denote as $\underset{n\to\infty}{\text{l.i.m.}}\, X_n = X$.*

Note that the limit X is unique with probability 1 (Problem 1.1). The following is an essential property of the L_2 space.

L_2-Completeness Theorem *(Loève (1977, p.163)). The space L_2 is complete in the sense that $\underset{n\to\infty}{\text{l.i.m.}}\, X_n = X$ for some $X \in L_2$ if and only if $E[(X_m - X_n)^2] \to 0$ as $m, n \to \infty$ in any manner.*

As an application of the L_2-completeness theorem it is easy to see (Problem 1.2) that the independent sequence $\{X_n\}$ defined by $P(X_n = \sqrt{n}) = \dfrac{1}{n}$ and $P(X_n = 0) = 1 - \dfrac{1}{n}$ does not converge in the m.s. sense, though it converges in probability to 0.

We now consider $\{Y_n(t)\}$, $t \in [a, b]$, which is a sequence of q-dimensional stochastic processes, where every element of $Y_n(t)$ belongs to L_2. If $\lim_{n\to\infty} E[(Y_n(t) - Y(t))'(Y_n(t) - Y(t))] = 0$ for each $t \in [a, b]$, then $\{Y_n(t)\}$ is said to converge in m.s. to $\{Y(t)\}$, which we denote as $\underset{n\to\infty}{\text{l.i.m.}}\, \{Y_n(t)\} = \{Y(t)\}$. If a particular point t is our concern, the m.s. convergence at t is denoted as $\underset{n\to\infty}{\text{l.i.m.}}\, Y_n(t) = Y(t)$. It holds as before that every element of $E(Y(t)Y'(t))$ is finite and the limit $Y(t)$ is unique with probability 1. Occasionally we shall need to deal with a matrix-valued stochastic process, $\{Z_n(t)\}$, say. In that case we assume that $E[\text{tr}(Z_n'(t)Z_n(t))] < \infty$, and $\underset{n\to\infty}{\text{l.i.m.}}\, \{Z_n(t)\} = \{Z(t)\}$ means that $\lim_{n\to\infty} E[\text{tr}\{(Z_n(t) - Z(t))'(Z_n(t) - Z(t))\}] = 0$ for each t.

The following theorem describes an operational property of the m.s. convergence, whose proof is left as Problem 1.3.

Theorem 2.1. *Suppose that $\{X_n(t)\}$ and $\{Y_n(t)\}$, $t \in [a,b]$, are sequences of q-dimensional stochastic processes for which $E(X_n'(t)\,X_n(t)) < \infty$, $E(Y_n'(t)\,Y_n(t)) < \infty$, $\underset{n\to\infty}{l.i.m.}\, X_n(t) = X(t)$ and $\underset{n\to\infty}{l.i.m.}\, Y_n(t) = Y(t)$. Then it holds that*

$$
\begin{aligned}
(2.1) \qquad & \lim_{n\to\infty} E(aX_n(t) + bY_n(t)) = aE(X(t)) + bE(Y(t))\,, \\
& \lim_{n\to\infty} E(X_n'(t)Y_n(t)) = E(X'(t)Y(t))\,,
\end{aligned}
$$

where a and b are scalar constants.

The above theorem tells us that "l.i.m." and "E" commute. We further point out that "l.i.m." and "Gaussianity" commute. This is because m.s. convergence implies convergence in distribution, but the following is a proof for the present statement. Suppose that $\{X_n(t)\}$ is a q-dimensional Gaussian sequence, that is, the finite-dimensional distributions of $X_n(t_1), \cdots, X_n(t_k)$ for each finite k and each collection $t_1 < \cdots < t_k$ on $[a,b]$ are normal for all n. Then, if $\underset{n\to\infty}{\text{l.i.m.}}\,\{X_n(t)\} = \{X(t)\}$, $\{X(t)\}$ is also Gaussian. In fact, putting $X_n = (X_n'(t_1), \cdots, X_n'(t_k))'$ and $X = (X'(t_1), \cdots, X'(t_k))'$, we have

$$
\begin{aligned}
\phi_n(\theta) &= E\{\exp(i\theta' X_n)\} \\
&= \exp\{i\theta' E(X_n) - \frac{1}{2}\theta' V(X_n)\theta\}\,,
\end{aligned}
$$

where $E(X_n) \to E(X)$ and $V(X_n) \to V(X)$ by Theorem 2.1. Then the c.f. of X must be $\phi(\theta) = \exp\left\{i\theta' E(X) - \dfrac{1}{2}\theta' V(X)\theta\right\}$ since

$$
\begin{aligned}
|\phi_n(\theta) - \phi(\theta)|^2 &\le [\,E\{\,|\exp(i\theta' X_n) - \exp(i\theta' X)|\,\}\,]^2 \\
&\le 2E\{\,(1 - \cos\theta'(X_n - X))\,\} \\
&\le \theta' E\{\,(X_n - X)(X_n - X)'\,\}\,\theta \longrightarrow 0\,.
\end{aligned}
$$

The interchangeability of "l.i.m." and "Gaussianity" will be carried over to derivatives and integrals defined subsequently.

The next theorem relates the existence of a limit in m.s. with an operational moment condition, the proof of which can be found in Loève (1978, p.135). See also Problem 1.4.

Theorem 2.2. *Let $\{Y_n(t)\}$, $t \in [a,b]$, be a sequence of q-dimensional stochastic processes for which $E(Y_n'(t)Y_n(t)) < \infty$ for all t. Then $\underset{n\to\infty}{\text{l.i.m.}}\{Y_n(t)\}$ exists if and only if $E(Y_m'(t)Y_n(t))$ converges to a finite function on $[a,b]$ as $m,\ n \to \infty$ in any manner.*

We now introduce some notions associated with the m.s. convergence described in Loève(1978) extending to vector - and matrix - valued stochastic processes.

Definition of m.s. continuity. $\{Y(t)\}$ *is m.s. continuous at* $t \in [a,b]$ *if*

$$\underset{h\to 0}{\text{l.i.m.}}\, Y(t+h) = Y(t) \qquad \text{for} \qquad t+h \in [a,b]\,.$$

Here the left side is interpreted as follows. Let $\{h_n\}$ be any sequence which converges to 0 as $n \to \infty$ maintaining $t + h_n \in [a,b]$ and put $Y_n(t) = Y(t+h_n)$. Then the above definition means that $\underset{n\to\infty}{\text{l.i.m.}}\, Y_n(t) = Y(t)$. The process $\{Y(t)\}$ is m.s. continuous at t if and only if $E(Y'(s)Y(t))$ is continuous at (t,t) (Loève(1978, p.136) and Problem 1.5). Moreover, $\{Y(t)\}$ is m.s. continuous at every $t \in [a,b]$ if and only if $E(Y'(s)Y(t))$ is continuous at every $(t,t) \in [a,b] \times [a,b]$. In this case $E(Y'(s)Y(t))$ is necessarily continuous on $[a,b] \times [a,b]$ (Problem 1.6).

The m.s. continuity does not necessarily imply the sample path continuity or the m.s. differentiability defined below. An example is the sample paths of Poisson processes, though such processes are not dealt with in this book (Problem 1.7).

Definition of m.s. differentiability. $\{Y(t)\}$ *is m.s. differentiable at* $t \in [a,b]$ *if*

$$\underset{h\to 0}{\text{l.i.m.}}\, \frac{Y(t+h) - Y(t)}{h} = \dot{Y}(t) \qquad \text{for} \quad t+h \in [a,b]\,.$$

The left side is interpreted in the same way as above putting $Y_n(t) = (Y(t+h_n) - Y(t))/h_n$. Because of Theorem 2.2 a necessary and sufficient condition for the m.s. differentiability of $\{Y(t)\}$ at t is that $E[(Y(t+h_1) - Y(t))'(Y(t+h_2) - Y(t))]/(h_1h_2)$ converges to a finite number as $h_1, h_2 \to 0$ in any manner. If $\{Y(t)\}$ is m.s. differentiable and Gaussian, the derivative process $\{\dot{Y}(t)\}$ is also Gaussian since $Y_n(t) = (Y(t+h_n) - Y(t))/h_n$ is Gaussian and $\underset{n\to\infty}{\text{l.i.m.}}\{Y_n(t)\} = \{\dot{Y}(t)\}$. If $\{Y(t)\}$ is m.s.

differentiable at s and t, then it can be shown (Problem 1.8) that

$$E(\dot{Y}(t)) = \frac{d}{dt}\, E(Y(t))\,, \tag{2.2}$$

$$E(\dot{Y}'(s)\dot{Y}(t)) = \frac{\partial^2}{\partial s \partial t}\, E(Y'(s)Y(t))\,. \tag{2.3}$$

Of course, the m.s. differentiability implies the m.s. continuity.

We also need to define the m.s. integral, but it is more involved and thus is treated separately. Since we need to consider the so-called Ito integral as well, which requires the knowledge of the Brownian motion, we next discuss the Brownian motion prior to integration.

Problems

1.1 Prove that, if a scalar sequence $\{X_n\}$ converges in m.s. to X, then X is unique with probability 1.

1.2 Show that the independent sequence $\{X_n\}$ defined by $P(X_n = \sqrt{n}) = \dfrac{1}{n}$ and $P(X_n = 0) = 1 - \dfrac{1}{n}$ does not converge in the m.s. sense, though it converges in probability to 0.

1.3 Prove Theorem 2.1.

1.4 Using Theorem 2.2, show that the one-dimensional independent stochastic process $\{Y_n(t)\}$ defined by

$$Y_n(t) = \begin{cases} n & P(Y_n(t) = n) = \dfrac{1}{n^2} \\ 1 & P(Y_n(t) = 1) = 1 - \dfrac{1}{n^2} \end{cases}$$

does not converge in the m.s. sense.

1.5 Prove that a q-dimensional stochastic process $\{Y(t)\}$ for which $E(Y'(t)Y(t)) < \infty$ is m.s. continuous at t if and only if $E(Y'(s)Y(t))$ is continuous at (t, t).

1.6 Prove that, if $\{Y(t)\}$ is m.s. continuous at every $t \in [a,b]$, then $E(Y'(s)Y(t))$ is continuous on $[a,b] \times [a,b]$.

1.7 The scalar Poisson process $\{X(t)\}$ defined on $[0,\infty)$ has independent increments and is characterized by

$$P(X(t)=k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}, \quad (k=0,1,2,\cdots).$$

Show that $\{X(t)\}$ is m.s. continuous at all t, but is nowhere m.s. differentiable.

1.8 Show that (2.2) and (2.3) hold if $\{Y(t)\}$ is m.s. differentiable.

1.9 Show that the one-dimensional stochastic process $\{Y(t)\}$ defined by $Y(t) = \cos(\omega t + U)$ is m.s. differentiable, where ω is a constant and U is uniformly distributed on $[0,\pi]$.

2.2. The standard Brownian motion and the Brownian bridge

In this section we become more specific about the stochastic process $\{Y(t)\}$ in L_2 and introduce two important processes frequently used throughout this book.

Definition of the standard Brownian motion. *We call a q-dimensional stochastic process $\{w(t)\}$ defined on [0,1] the q-dimensional standard Brownian motion if*

i) $P(w(0)=0)=1$ *;*

ii) $w(t_1)-w(t_0),\ w(t_2)-w(t_1), \cdots, w(t_n)-w(t_{n-1})$ *are independent for any positive integer n and time points* $0 \le t_0 < t_1 < \cdots < t_n \le 1$ *;*

iii) $w(t)-w(s) \sim N(0,(t-s)I_q)$ *for* $0 \le s < t \le 1$.

It follows that $\{w(t)\}$ is a zero-mean nonstationary Gaussian process with $\mathrm{Cov}(w(s), w(t)) = \min(s,t)I_q$ and has stationary independent increments. It is known that the sample path of $\{w(t)\}$ is continuous with probability 1, while it is nowhere differentiable (see, for example, Billingsley(1986, Section 37)) and of unbounded variation on any interval subset of [0,1] (Hida(1980, p.57)). In terms of the m.s. calculus it can be shown (Problem 2.1) that

(a) $\{w(t)\}$ is m.s. continuous ;

(b) $\{w(t)\}$ is nowhere m.s. differentiable ;

(c) $\underset{\substack{n\to\infty \\ \Delta_n\to 0}}{\text{l.i.m.}} \sum_{i=1}^{n}(w(t_i)-w(t_{i-1}))'(w(t_i)-w(t_{i-1})) = (b-a)q$;
where $0 \le a = t_0 < t_1 < \cdots < t_n = b \le 1$ and $\Delta_n = \max_i (t_i - t_{i-1})$.

The property (c) implies that $w(t)$ is of unbounded variation in the m.s. sense.

An example of the q-dimensional standard Brownian motion is (see Chan and Wei (1988) and Problem 2.2)

$$w(t) = \sum_{n=1}^{\infty} \frac{\sqrt{2}\sin\left(n-\frac{1}{2}\right)\pi t}{\left(n-\frac{1}{2}\right)\pi} Z_n\,, \tag{2.4}$$

where $\{Z_n\} \sim \text{NID}(0, I_q)$. Note that the right hand variable exists in the m.s. sense because of the L_2-completeness theorem and $\sum_{n=1}^{\infty} 1 \bigg/ \left(\left(n-\frac{1}{2}\right)\pi\right)^2 = \frac{1}{2} < \infty$.

Definition of the Brownian bridge. *We call a q-dimensional stochastic process $\{\bar{w}(t)\}$ defined on [0,1] the q-dimensional Brownian bridge if*

i) $\{\bar{w}(t)\}$ is Gaussian ;

ii) $E(\bar{w}(t)) = 0$ and $Cov(\bar{w}(s), \bar{w}(t)) = (\min(s,t) - st)I_q$ for $0 \le s,\, t \le 1$.

It follows that $\{\bar{w}(t)\}$ is the q-dimensional standard Brownian motion $\{w(t)\}$ conditioned on $P(w(1) = 0) = 1$ (Problem 2.3). The processes $\{w(t) - tw(1)\}$ and

$$\bar{w}(t) = \sum_{n=1}^{\infty} \frac{\sqrt{2}\sin n\pi t}{n\pi} Z_n\,, \tag{2.5}$$

where $\{Z_n\} \sim \text{NID}(0, I_q)$, are examples of the q-dimensional Brownian bridge (Problem 2.4).

Figure 2.1 shows a sample path of the one-dimensional standard Brownian motion simulated from (2.4), while Figure 2.2 that of the one-dimensional Brownian bridge simulated from (2.5). We observe an erratic nature of both processes because of nondifferentiability.

Figure 2.1 Figure 2.2

Problems

2.1 When $\{w(t)\}$ is the q-dimensional standard Brownian motion, show that (a) $\{w(t)\}$ is m.s. continuous; (b) $\{w(t)\}$ is nowhere m.s. differentiable; and

$$\text{(c)}\quad \underset{\substack{n\to\infty \\ \Delta_n\to 0}}{\text{l.i.m.}} \sum_{i=1}^{n}(w(t_i)-w(t_{i-1}))'(w(t_i)-w(t_{i-1})) = (b-a)q\ ;$$

where $0 \le a = t_0 < t_1 < \cdots < t_n = b \le 1$ and $\Delta_n = \max_i(t_i - t_{i-1})$.

2.2 Show that $\{w(t)\}$ defined in (2.4) is the q-dimensional standard Brownian motion, using the formula :

$$\sum_{n=1}^{\infty} \frac{\cos\left(n-\frac{1}{2}\right)\pi x}{\left(\left(n-\frac{1}{2}\right)\pi\right)^2} = \frac{1}{2}(1-|x|) \qquad \text{for} \qquad |x| \le 2\,.$$

2.3 Show that $\text{Cov}(w(s), w(t)|w(1) = 0) = (\min(s,t) - st)I_q$, where $\{w(t)\}$ is the q-dimensional standard Brownian motion.

2.4 Show that $\{\bar{w}(t)\}$ defined in (2.5) is the q-dimensional Brownian bridge, using the formula :

$$\sum_{n=1}^{\infty} \frac{\cos n\pi x}{n^2\pi^2} = \frac{1}{4}(x-1)^2 - \frac{1}{12} \qquad \text{for} \qquad 0 \le x \le 2\,.$$

2.3. Mean square integration

Here we assume that $\{Y(t)\}$ is a q-dimensional stochastic process defined on $[a, b]$ for which $E(Y'(t)Y(t)) < \infty$. We can think of $Y(t)$ as the q-dimensional standard Brownian motion $w(t)$, but, of course, can think of any $Y(t)$ with $E(Y'(t)Y(t)) < \infty$.

Soong (1973) gives a good exposition for m.s. integrals, which we follow extending to the vector case. Consider a collection of all finite partitions $\{p_m\}$ of an interval $[a, b]$. The partition p_m is defined by

$$(2.6) \qquad p_m\ :\quad a = s_0 < s_1 < \cdots < s_m = b\,.$$

Put $\Delta_m = \max_i(s_i - s_{i-1})$ and let s'_i be an arbitrary point in the interval $[s_{i-1}, s_i)$. Suppose that $f(s,t)$ is an ordinary real-valued function defined on $[a,b] \times [a,b]$ and Riemann integrable with respect to s for every $t \in [a,b]$.

2.3.1. The mean square Riemann integral

We define

$$V(t) = \int_a^b f(s,t)Y(s)ds \tag{2.7}$$

if $\underset{\substack{m\to\infty \\ \Delta_m \to 0}}{\text{l.i.m.}}\ V_m(t)$ exists for any sequence of subdivisions p_m and for any $s'_i \in [s_{i-1}, s_i)$, $i = 1, \cdots, m$, where

$$V_m(t) = \sum_{i=1}^{m} f(s'_i, t)Y(s'_i)(s_i - s_{i-1})\,.$$

The m.s. integral (2.7) is independent of the sequence of subdivisions p_m as well as the positions of $s'_i \in [s_{i-1}, s_i)$. The existence condition for (2.7) can be related with the usual Riemann integral and is given by the following theorem (Soong (1973, Theorem 4.5.1) and Problem 3.1).

Theorem 2.3. *The m.s. integral (2.7) exists if and only if the ordinary double Riemann integral*

$$\int_a^b \int_a^b f(r,t)f(s,t)E(Y'(r)Y(s))drds \tag{2.8}$$

exists and is finite.

It is clear that, if the m.s. integral is well defined, "expectation" up to the second order and "integration" commute, i.e.,

$$E\left(\int_a^b f(s,t)Y(s)ds\right) = \int_a^b f(s,t)E(Y(s))ds\,,$$

$$E\left(\int_a^b \int_a^b f(r,t)f(s,t)Y'(r)Y(s)drds\right) = \int_a^b \int_a^b f(r,t)f(s,t)E(Y'(r)Y(s))drds\,.$$

If $\{Y(t)\}$ is Gaussian and $f(s,t) \equiv 1$, then the integral $V(t)$ in (2.7) is evidently Gaussian. Moreover we have (Soong (1973, Theorem 4.6.4))

Theorem 2.4. *If the m.s. integral*

$$V(t) = \int_a^b f(s,t)Y(s)ds$$

of a normal process $\{Y(s)\}$ *exists, then* $\{V(t)\}$ *is also normal with*

$$\begin{aligned} E(V(t)) &= \int_a^b f(s,t)E(Y(s))ds\,, \\ E(V'(s)V(t)) &= \int_a^b \int_a^b f(r,s)f(u,t)E(Y'(r)Y(u))drdu\,. \end{aligned}$$

Some other properties of the m.s. Riemann integral follow (Soong (1973) and Problem 3.2).

(a) If $Y(t)$ is m.s. continuous on $[a,b]$, $Y(t)$ is m.s. integrable on $[a,b]$;

(b) The m.s. integral of $Y(t)$ on $[a,b]$, if it exists, is unique ;

(c) If $Y(t)$ is m.s. continuous on $[a,b]$, then $X(t) = \int_a^t Y(s)ds$ $(t \in [a,b])$ is m.s. differentiable on $[a,b]$ with $\dot{X}(t) = Y(t)$.

(d) If $\dot{Y}(t)$ is m.s. integrable on $[a,b]$, then

$$Y(t) - Y(a) = \int_a^t \dot{Y}(s)ds\,, \qquad [a,t] \subset [a,b]\,.$$

We have already presented one-dimensional m.s. integrals in (1.11) and (1.18) :

$$V = \int_0^1 w^2(t)dt\,, \quad W = \int_0^1 (w(t) - tw(1))^2 dt\,. \tag{2.9}$$

It is evident that V and W are well defined and it holds (Problem 3.3) that $E(V) = \frac{1}{2}$, $E(V^2) = \frac{7}{12}$, $E(W) = \frac{1}{6}$ and $E(W^2) = \frac{1}{20}$. Moments of all orders are finite in these cases since the c.f.'s of V and W were found in Chapter 1 to be $(\cos\sqrt{2i\theta})^{-\frac{1}{2}}$ and $\left(\sin\sqrt{2i\theta}\big/\sqrt{2i\theta}\right)^{-\frac{1}{2}}$, respectively. As another example consider the q-dimensional integrated Brownian motion :

$$V(t) = \int_0^t w(s)ds = \int_0^1 I_{[0,t]}(s)w(s)ds\,, \tag{2.10}$$

where $I_A(s)$ is the indicator function of the set A. It can be checked (Problem 3.4) that the integral $V(t)$ is well defined and $V(t) \sim \mathrm{N}(0, t^3 I_q/3)$.

2.3.2. The mean square Riemann-Stieltjes integral

We first define integrals of the following types :

$$(2.11) \qquad V_1 = \int_a^b f(s)dY(s)\,,$$

$$(2.12) \qquad V_2 = \int_a^b Y(s)df(s)\,,$$

when $\underset{\substack{m\to\infty \\ \Delta_m \to 0}}{\text{l.i.m.}}\ V_{1m}$ and $\underset{\substack{m\to\infty \\ \Delta_m \to 0}}{\text{l.i.m.}}\ V_{2m}$ exist, respectively, for any sequence of subdivisions p_m, where $f(t)$ is an ordinary real-valued function and

$$\begin{aligned} V_{1m} &= \sum_{i=1}^{m} f(s_i')[Y(s_i) - Y(s_{i-1})]\,, \\ V_{2m} &= \sum_{i=1}^{m} Y(s_i')[f(s_i) - f(s_{i-1})]\,. \end{aligned}$$

The integrals in (2.11) and (2.12) are called the m.s. Riemann-Stieltjes integrals and are independent of the sequence of subdivisions as well as the positions of $s_i' \in [s_{i-1}, s_i)$.

Normality is retained in the integrals (2.11) and (2.12). Namely V_1 and V_2 are both normal if $\{Y(t)\}$ is normal. The existence conditions for (2.11) and (2.12) are derived from Theorem 2.2 and stated as follows.

Theorem 2.5. *The m.s. Riemann-Stieltjes integrals in (2.11) and (2.12) exist if and only if the ordinary double Riemann-Stieltjes integrals*

$$(2.13) \qquad \int_a^b \int_a^b f(s)f(t)E(dY'(s)dY(t)) = E(V_1'V_1)\,,$$

$$(2.14) \qquad \int_a^b \int_a^b E(Y'(s)Y(t))df(s)df(t) = E(V_2'V_2)\,,$$

exist and are finite, respectively, where the integrals are defined in a self-evident manner as limits of approximating sums.

A sufficient condition for the existence of (2.11) is that $E(Y'(s)Y(t))$ is of bounded variation and $f(t)$ is continuous, while (2.12) exists if $E(Y'(s)Y(t))$ is continuous and $f(t)$ is of bounded variation. If these are the cases, (2.13) and (2.14) can be computed as

$$E(V_1'V_1) = \lim_{\substack{m\to\infty \\ \Delta_m\to 0}} \sum_{i=1}^{m}\sum_{j=1}^{m} f(s_i)f(s_j)E[(Y(s_i)-Y(s_{i-1}))'(Y(s_j)-Y(s_{j-1}))]\,, \tag{2.15}$$

$$E(V_2'V_2) = \lim_{\substack{m\to\infty \\ \Delta_m\to 0}} \sum_{i=1}^{m}\sum_{j=1}^{m} E(Y'(s_i)Y(s_j))[f(s_i)-f(s_{i-1})][f(s_j)-f(s_{j-1})]\,. \tag{2.16}$$

As an example, take $[a,b]=[0,1]$ and consider

$$V_1 = \int_0^1 (1-s)dw(s)\,, \quad (f(s)=1-s,\ \ Y(s)=w(s))\,. \tag{2.17}$$

Then V_1 is well defined (Problem 3.5) and is normal with $E(V_1)=0$ and (2.15) leads us to

$$\begin{aligned} E(V_1'V_1) &= \lim_{\substack{m\to\infty \\ \Delta_m\to 0}} \sum_{i=1}^{m}(1-s_i)^2(s_i-s_{i-1})q \\ &= q\int_0^1 (1-s)^2 ds = \frac{q}{3}\,. \end{aligned}$$

Note that we also have $E(V_1V_1')=I_q/3$ and that, to compute $E(V_1'V_1)$ and $E(V_1V_1')$, we may use the following relation :

$$E(dw'(s)dw(t)) = \begin{cases} q\,ds & (s=t) \\ 0 & (s\neq t)\,. \end{cases}$$

The two integrals (2.11) and (2.12) can be combined together in the following theorem for *integration by parts* (Soong (1973, Theorem 4.5.3)).

Theorem 2.6. *If either V_1 in (2.11) or V_2 in (2.12) exists, then both integrals exist, and*

$$\int_a^b f(s)dY(s) = [f(s)Y(s)]_a^b - \int_a^b Y(s)df(s)\,. \tag{2.18}$$

For example we have

$$\int_0^1 dw(s) = w(1), \qquad \int_0^1 (1-s)dw(s) = \int_0^1 w(s)ds\,.$$

For later discussions we also consider the following m.s. double Riemann-Stieltjes integral defined on $[0,1] \times [0,1]$:

$$X = \int_0^1 \int_0^1 K(s,t)dw'(s)Hdw(t)\,, \tag{2.19}$$

where $K(s,t)$ is symmetric on $[0,1] \times [0,1]$, while H is a $q \times q$ symmetric, constant matrix. Define the partition $p_{m,n}$ of $[0,1] \times [0,1]$ by

$$p_{m,n}\ :\ 0 = s_0 < s_1 < \cdots < s_m = 1\,;\ 0 = t_0 < t_1 < \cdots < t_n = 1 \tag{2.20}$$

and put

$$\Delta_{m,n} = \max(s_1 - s_0, \cdots, s_m - s_{m-1}, t_1 - t_0, \cdots, t_n - t_{n-1})\,.$$

Then the m.s. double integral X in (2.19) is well defined if

$$X_{m,n} = \sum_{i=1}^{m}\sum_{j=1}^{n} K(s_i', t_j')[w(s_i) - w(s_{i-1})]'H[w(t_j) - w(t_{j-1})] \tag{2.21}$$

converges in m.s. as $m,\, n \to \infty$ and $\Delta_{m,n} \to 0$ for any sequence of subdivisions $p_{m,n}$. The integral X is independent of the choice of $p_{m,n}$ as well as the choice of $s_i' \in [s_{i-1}, s_i)$ and $t_j' \in [t_{j-1}, t_j)$.

The existence condition for the integral (2.19) is similarly given as in Theorem 2.5 using the quadruplex integral. A sufficient condition is that $K(s,t)$ is continuous on $[0,1] \times [0,1]$ (Problem 3.6). If this is the case, it holds (Problem 3.7) that

$$E(X) = \int_0^1 K(s,s)ds \times \mathrm{tr}(H)\,, \tag{2.22}$$

$$E(X^2) = 2\int_0^1 \int_0^1 K^2(s,t)dsdt \times \mathrm{tr}(H^2) + (\int_0^1 K(s,s)ds \times \mathrm{tr}(H))^2\,. \tag{2.23}$$

Throughout this book it will be assumed that $K(s,t)$ is symmetric and continuous on $[0,1] \times [0,1]$, although this assumption can be relaxed (see Anderson and Darling (1952)) for the existence of the integral (2.19).

Some examples of (2.19) follow. Suppose first that $K(s,t) = g(s)g(t)$, where $g(t)$ is a continuous function on [0,1]. Then the double integral can be reduced to the product of single integrals :

$$\int_0^1 \int_0^1 K(s,t)dw'(s)Hdw(t) = \int_0^1 g(s)dw'(s)H \int_0^1 g(t)dw(t)\,, \tag{2.24}$$

the distribution of which is equal to that of $\int_0^1 g^2(t)dt\sum_{i=1}^{q} \lambda_i Z_i^2$, where λ_i's are the eigenvalues of H and $\{Z_i\} \sim \mathrm{NID}(0,1)$ (Problem 3.8). Thus the distribution of (2.24) is that of a finite sum of weighted $\chi^2(1)$ random variables. In general the symmetric and continuous function $K(s,t)$ is said to be *degenerate* if it can be expressed as the sum of a finite number of terms, each of which is the product of a function of s only and a function of t only, that is,

$$K(s,t) = \sum_{i=1}^{n} g_i(s) g_i(t)\,.$$

If this is not the case, $K(s,t)$ is said to be *nondegenerate*.

Let us present a few nondegenerate cases of $K(s,t)$ relating to the m.s. Riemann integral. For this purpose we first take up $K(s,t) = 1 - \max(s,t)$. Then, using integration by parts, we obtain

$$\begin{aligned}
(2.25) \qquad & \int_0^1 \int_0^1 [1 - \max(s,t)]dw'(s)dw(t) \\
&= \int_0^1 \left[(1-t)\int_0^t dw'(s) + \int_t^1 (1-s)dw'(s)\right] dw(t) \\
&= \int_0^1 \left[\int_t^1 w'(s)ds\right] dw(t) = \int_0^1 w'(t)w(t)dt\,.
\end{aligned}$$

A more general relation may be derived in a reversed direction as follows. Let $g(t)$ be continuous on [0,1]. Then, putting $g_j = g(j/n)$ and $\Delta w_j = w(j/n) - w((j-1)/n)$, we have

$$\begin{aligned}
\int_0^1 g(t)w'(t)w(t)dt &= \underset{n\to\infty}{\mathrm{l.i.m.}}\ \frac{1}{n}\sum_{l=1}^{n}\left[g_l \sum_{j=1}^{l}\Delta w_j' \sum_{k=1}^{l}\Delta w_k\right] \\
&= \underset{n\to\infty}{\mathrm{l.i.m.}} \sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{1}{n}\sum_{l=\max(j,k)}^{n} g_l\right)\Delta w_j' \Delta w_k \\
&= \underset{n\to\infty}{\mathrm{l.i.m.}} \sum_{j=1}^{n}\sum_{k=1}^{n}\left(\int_{\max(j,k)/n}^{1} g(r)dr\right)\Delta w_j' \Delta w_k \\
&= \int_0^1\int_0^1 \left(\int_{\max(s,t)}^{1} g(r)dr\right) dw'(s)dw(t)\,.
\end{aligned}$$

This implies that we can change the order of integration to get

$$\int_0^1 g(t)w'(t)w(t)dt = \int_0^1 g(t)\left[\int_0^t\int_0^t dw'(u)dw(v)\right]dt$$

$$= \int_0^1 \int_0^1 \left[\int_{\max(u,v)}^1 g(t)dt \right] dw'(u)dw(v) \, .$$

Note that we have encountered in (1.11) the distributional equivalence of the first and last expressions in (2.25) for the scalar case. We have just seen that these are, in fact, the same in the m.s. sense. The first expression in (1.11) indicates that $K(s,t) = 1 - \max(s,t)$ is nondegenerate, which will be further studied in Chapter 5.

The following relations can also be obtained similarly (Problem 3.9).

$$\int_0^1 \int_0^1 [\min(s,t) - st] dw'(s) H dw(t) = \int_0^1 \tilde{w}'(t) H \tilde{w}(t) dt \, , \tag{2.26}$$

$$\int_0^1 \int_0^1 \left[\frac{1}{3} - \max(s,t) + \frac{s^2+t^2}{2} \right] dw'(s) H dw(t) = \int_0^1 \bar{w}'(t) H \bar{w}(t) dt \, , \tag{2.27}$$

where $\tilde{w}(t) = w(t) - \int_0^1 w(t)dt$ is the demeaned Brownian motion and $\bar{w}(t) = w(t) - tw(1)$ is the Brownian bridge. We note in passing that (1.18) has given the distributional equivalence of the left side in (2.26) and the right side in (2.27). This, however, cannot be carried over to the equivalence in the m.s. sense. The functions $\min(s,t) - st$ and $\frac{1}{3} - \max(s,t) + \frac{s^2+t^2}{2}$ are also nondegenerate. The proofs of these will also be deferred until Chapter 5.

Problems

3.1 Prove Theorem 2.3.

3.2 Prove that, if $\{Y(t)\}$ is m.s. continuous on $[a,b]$, then $\{Y(t)\}$ is m.s. integrable on $[a,b]$.

3.3 Show that $E(V) = \frac{1}{2}$, $E(V^2) = \frac{7}{12}$, $E(W) = \frac{1}{6}$ and $E(W^2) = \frac{1}{20}$ in (2.9).

3.4 Prove that the integral $V(t)$ in (2.10) is well defined and that $V(t) \sim \mathrm{N}(0, t^3 I_q/3)$.

3.5 Prove that the integral V_1 in (2.17) is well defined.

3.6 Show that the integral in (2.19) exists and is finite if $K(s,t)$ is symmetric and continuous on $[0,1] \times [0,1]$.

3.7 Derive (2.22) and (2.23).

3.8 Show that

$$\mathcal{L}\left(\int_0^1 g(s)dw'(s)H\int_0^1 g(t)dw(t)\right) = \mathcal{L}\left(\int_0^1 g^2(t)dt\sum_{i=1}^{q}\lambda_i Z_i^2\right),$$

where λ_i's are the eigenvalues of H and $\{Z_i\} \sim \mathrm{NID}(0,1)$.

3.9 Prove the relations in (2.26) and (2.27).

2.4. The integrated Brownian motion

As was noted in the previous section, the Brownian motion itself is quite erratic in the sense that the sample path is nowhere differentiable, though it is continuous. The integral of the Brownian motion, however, becomes smooth. Here we consider a special case of such integrals defined by

$$F_g(t) = \int_0^t F_{g-1}(s)ds, \quad (g = 1, 2, \cdots), \quad F_0(t) = w(t), \tag{2.28}$$

where $\{w(t)\}$ is the q-dimensional standard Brownian motion. The process $\{F_g(t)\}$ defined recursively in this way may be called the *g-fold integrated Brownian motion* (Chan and Wei (1988)). It is clear that the m.s. integral in (2.28) is well defined for any positive integer g. Thus $F_g(t)$ with $g \geq 1$ is g-times m.s. continuously differentiable, though $F_0(t) = w(t)$ is not differentiable.

Figure 2.3 shows a sample path of $\{F_1(t)\}$ for $q = 1$, which was simulated, on the basis of (2.4), from

$$F_1(t) = \int_0^t w(s)ds = \sum_{n=1}^{\infty}\frac{\sqrt{2}}{\left(n-\frac{1}{2}\right)^2\pi^2}\left\{1-\cos\left(n-\frac{1}{2}\right)\pi t\right\}Z_n,$$

while Figure 2.4 that of $\{F_2(t)\}$ for $q = 1$, which was simulated similarly from

$$F_2(t) = \int_0^t F_1(s)ds = \sum_{n=1}^{\infty}\frac{\sqrt{2}}{\left(n-\frac{1}{2}\right)^3\pi^3}\left\{\left(n-\frac{1}{2}\right)\pi t-\sin\left(n-\frac{1}{2}\right)\pi t\right\}Z_n.$$

In comparison with the sample path of $\{F_0(t)\} = \{w(t)\}$ shown in Figure 2.1, we recognize that the sample paths of $\{F_g(t)\}$ become smoother as g gets large because of g-times differentiability. We also observe a decrease in the variation of $\{F_g(t)\}$ with g. In fact it holds that, when $q = 1$, $V(F_0(t)) = t$, $V(F_1(t)) = t^3/3$ and $V(F_2(t)) = t^5/20$. In general we have $V(F_g(t)) = t^{2g+1}/\left((2g+1)(g!)^2\right)$, which can be proved easily by using Theorem 2.7 described below.

Figure 2.3 Figure 2.4

We also define

$$\text{(2.29)} \qquad \bar{F}_g(t) = F_g(t) - tF_g(1) \qquad : g\text{-fold integrated Brownian bridge},$$

$$\text{(2.30)} \qquad \tilde{F}_g(t) = F_g(t) - \int_0^1 F_g(s)ds \qquad : g\text{-fold integrated demeaned Brownian motion}.$$

As for $\{F_g(t)\}$ we have the following equivalent expression, the proof of which is left as Problem 4.1.

Theorem 2.7. *The stochastic process $\{F_g(t)\}$ in (2.28) can be expressed by the Riemann-Stieltjes integral as*

$$\text{(2.31)} \qquad F_g(t) = \int_0^t \frac{(t-s)^g}{g!}\, dw(s), \quad (g = 0, 1, \cdots).$$

The superiority of the expression (2.31) over (2.28) is that the former may be defined for real g replacing $g!$ by $\Gamma(g+1)$. In fact the expression (2.31) with $g!$ replaced by $\Gamma(g+1)$ was earlier introduced as the Holmgren-Riemann-Liouville fractional integral (Mandelbrot and Van Ness (1968)), where g may be any value greater than $-\frac{1}{2}$. The fractional case is certainly interesting, but the analysis needs a separate treatment. We maintain in this book that g is a nonnegative integer.

Using (2.31) we obtain the following integral relations concerning $\{F_g(t)\}$, $\{\bar{F}_g(t)\}$ and $\{\tilde{F}_g(t)\}$ (Problem 4.2).

$$\int_0^1 F_g(t)dt = \int_0^1 \frac{(1-s)^{g+1}}{(g+1)!}\, dw(s),$$

$$\int_0^1 \bar{F}_g(t)dt = \int_0^1 \left[\frac{(1-s)^{g+1}}{(g+1)!} - \frac{(1-s)^g}{2(g!)}\right] dw(s)\,,$$

$$\int_0^1 \tilde{F}_g(t)dt = 0\,,$$

$$\int_0^1 F_g'(t)F_g(t)dt = \int_0^1\int_0^1 K_g(s,t)dw'(s)dw(t)\,,$$

$$\int_0^1 \bar{F}_g'(t)\bar{F}_g(t)dt = \int_0^1\int_0^1 [K_g(s,t)+L_g(s,t)]dw'(s)dw(t)\,,$$

$$\int_0^1 \tilde{F}_g'(t)\tilde{F}_g(t)dt = \int_0^1\int_0^1 \left[K_g(s,t) - \frac{((1-s)(1-t))^{g+1}}{((g+1)!)^2}\right] dw'(s)dw(t)\,,$$

where

$$K_g(s,t) = \int_{\max(s,t)}^1 \frac{((u-s)(u-t))^g}{(g!)^2}\,du\,, \tag{2.32}$$

$$L_g(s,t) = \frac{(1-s)^{g+2}+(1-t)^{g+2}}{g!(g+2)!} + \frac{((1-s)(1-t))^g}{g!}\left(\frac{1}{3(g!)} - \frac{2-s-t}{(g+1)!}\right).$$

It can be checked easily that, when $g=0$, the above formulas reduce to those obtained in the previous section. The expressions on the right sides above may be more useful for computing moments and deriving their distributions, which will be discussed in Chapter 5.

We note in passing that the integrated Brownian motion naturally appears from a simple I(d) process :

$$(1-L)^d\, y_t = \varepsilon_t\,, \quad (t=1,\cdots,T)\,,$$

where L is the lag operator, d is a positive integer, $\{\varepsilon_t\} \sim$ i.i.d.$(0, I_q)$, and the initial values y_t $(t=-(d-1), -(d-2), \cdots, 0)$ are all set at 0. We shall show in Chapter 3 that

$$\mathcal{L}\left(\frac{1}{T^{d-\frac{1}{2}}}\, y_T\right) \longrightarrow \mathcal{L}\left(F_{d-1}(1)\right)\,,$$

$$\mathcal{L}\left(\frac{1}{T^{d+\frac{1}{2}}}\sum_{t=1}^T y_t\right) \longrightarrow \mathcal{L}\left(\int_0^1 F_{d-1}(t)dt\right)\,,$$

$$\mathcal{L}\left(\frac{1}{T^{2d}}\sum_{t=1}^{T} y_t' y_t\right) \longrightarrow \mathcal{L}\left(\int_0^1 F_{d-1}'(t) F_{d-1}(t)dt\right) ,$$

and so on. Thus a simple I(d) process $\{y_t\}$ divided by $T^{d-\frac{1}{2}}$ is essentially the $(d-1)$-fold integrated Brownian motion.

We remark finally that, for $g \geq 1$, $F_g(t)$ is of bounded variation and

$$\underset{\substack{m\to\infty \\ \Delta_m\to 0}}{\text{l.i.m.}} \sum_{i=1}^{m}(F_g(t_i) - F_g(t_{i-1}))'(F_g(t_i) - F_g(t_{i-1})) = 0 .$$

Thus we can define the m.s. integral $\int_0^t F_g(s)dF_g'(s)$. Then we can use integration by parts to obtain (Problem 4.3)

Theorem 2.8. *For any positive integer g it holds that*

$$\text{(2.33)} \qquad \int_0^t F_g(s)dF_g'(s) + \left(\int_0^t F_g(s)dF_g'(s)\right)' = F_g(t)F_g'(t) , \quad (g = 1, 2, \cdots) .$$

Note that the formula (2.33) does not apply to the case $g = 0$, that is, the case where $F_g(t) = w(t)$. We need a separate analysis for this case, which is discussed in the next section.

Problems

4.1 Derive the formula (2.31) from (2.28)

4.2 Prove that

$$\int_0^1 \bar{F}_g(t)dt = \int_0^1 \left[\frac{(1-s)^{g+1}}{(g+1)!} - \frac{(1-s)^g}{2(g!)}\right] dw(s) ,$$

$$\int_0^1 \tilde{F}_g'(t)\tilde{F}_g(t)dt = \int_0^1\int_0^1 \left[K_g(s,t) - \frac{((1-s)(1-t))^{g+1}}{((g+1)!)^2}\right] dw'(s)dw(t) ,$$

where $\bar{F}_g(t)$ and $\tilde{F}_g(t)$ are defined in (2.29) and (2.30), respectively, while $K_g(s,t)$ is defined in (2.32).

4.3 Prove that the relation (2.33) holds.

2.5. The mean square Ito integral : the scalar case

In this section we deal with the integral of the form

$$U(t) = \int_0^t X(s)dw(s)\,, \tag{2.34}$$

where $0 \le t \le 1$, $\{w(t)\}$ is the one-dimensional standard Brownian motion, and $\{X(t)\}$ is a scalar stochastic process in L_2.

We define the *Ito integral* for (2.34) as follows. Let the partition p_m of $[0,t]$ be $0 = s_0 < s_1 < \cdots < s_m = t$ and put $\Delta_m = \max_i(s_i - s_{i-1})$. Then, form the random variable

$$U_m(t) = \sum_{i=1}^{m} X(s_{i-1})(w(s_i) - w(s_{i-1}))\,. \tag{2.35}$$

The Ito integral for (2.34) is said to be well defined if $\underset{\substack{m\to\infty \\ \Delta_m\to 0}}{\text{l.i.m.}}\ U_m(t)$ exists. The following is the existence and uniqueness theorem for the Ito integral. For the proof see Jazwinski(1970, Theorem 4.2) and Soong(1973, Theorem 5.2.1).

Existence and Uniqueness Theorem for the Ito Integral. *Suppose that*

i) $X(t)$ is m.s. continuous on [0,1];

ii) $X(t)$ is independent of $\{w(t_j) - w(t_i);\ 0 \le t \le t_i \le t_j \le 1\}$ for all $t \in [0,1]$.

Then the Ito integral defined above exists and is unique.

We notice in (2.35) that the value of $X(s)$ taken in the interval $[s_{i-1}, s_i)$ is $X(s_{i-1})$, unlike the m.s. Riemann or Riemann-Stieltjes integral. If we take $s_i'\ (\ne s_{i-1})$ in $[s_{i-1}, s_i)$, then the m.s. limit will be different, as is exemplified later. It is clear that $E(U(t)) = 0$ and, for $h > 0$,

$$\begin{aligned} E[(U(t+h) - U(t))^2] &= \int_t^{t+h} E(X^2(s))ds \\ &\le h \max_{t\le s\le t+h} E(X^2(s)) \end{aligned}$$

so that $U(t)$ is m.s. continuous. The m.s. differentiability of $U(t)$, however, is not ensured in the usual sense. Nonetheless we formally write (2.34) as $dU(t) = X(t)dw(t)$.

As an example of the Ito integral let us consider

$$U(t) = \int_0^t w(s)dw(s)\,. \tag{2.36}$$

Since $X(t) = w(t)$ satisfies the conditions i) and ii) in the above theorem, the sum

$$\begin{aligned} U_m(t) &= \sum_{i=1}^{m} w(s_{i-1})(w(s_i) - w(s_{i-1})) \\ &= -\frac{1}{2}\left[\sum_{i=1}^{m}(w(s_i) - w(s_{i-1}))^2 - \sum_{i=1}^{m} w^2(s_i) + \sum_{i=1}^{m} w^2(s_{i-1})\right] \\ &= \frac{1}{2}w^2(t) - \frac{1}{2}\sum_{i=1}^{m}(w(s_i) - w(s_{i-1}))^2 \end{aligned}$$

must converge in the m.s. sense. We have already shown (see Problem 2.1) that

$$\underset{\substack{m\to\infty \\ \Delta_m \to 0}}{\text{l.i.m.}} \sum_{i=1}^{m}(w(s_i) - w(s_{i-1}))^2 = t\,, \tag{2.37}$$

where $0 = s_0 < s_1 < \cdots < s_m = t$ and $\Delta_m = \max_i (s_i - s_{i-1})$. Therefore

$$\int_0^t w(s)dw(s) = \frac{1}{2}(w^2(t) - t)\,. \tag{2.38}$$

We now formally have $d\left(\frac{1}{2}(w^2(t) - t)\right) = w(t)dw(t)$ or

$$d(w^2(t)) = 2w(t)dw(t) + dt\,. \tag{2.39}$$

This is a simplified version of the Ito calculus to be discussed in Section 2.7.

We now show that the m.s. limit of the sum like (2.35) crucially depends on the choice of values of $X(s)$ in the intervals $[s_{i-1}, s_i)$. In fact we have

$$\underset{\substack{m\to\infty \\ \Delta_m \to 0}}{\text{l.i.m.}} \sum_{i=1}^{m} w((1-a)s_{i-1} + as_i)(w(s_i) - w(s_{i-1})) = \frac{1}{2}(w^2(t) - t) + at\,, \tag{2.40}$$

where $0 \le a \le 1$ (Problem 5.1). The case $a = 0$ corresponds to the Ito integral, while the integral with $a = \frac{1}{2}$ is called the Stratonovich integral. Simple and convenient properties of the Ito integral are that $U(t)$ in (2.34) has a zero mean and is a *martingale*, that is, for any $s \le t$, $E(U(t)|U(s)) = U(s)$ with probability 1 (Problem 5.2).

According to the definition of the Ito integral the following relations can also be established (Problem 5.3).

$$\int_0^t X(s)(dw(s))^2 = \int_0^t X(s)ds\,, \tag{2.41}$$

$$\int_0^t X(s)(dw(s))^3 = 0\,, \tag{2.42}$$

where $X(t)$ satisfies the conditions i) and ii) in the existence and uniqueness theorem, while

$$\int_0^t X(s)(dw(s))^j = \underset{\substack{m\to\infty \\ \Delta_m\to 0}}{\text{l.i.m.}} \sum_{i=1}^m X(s_{i-1})(w(s_i)-w(s_{i-1}))^j\,.$$

In particular, putting $X(s)\equiv 1$, we formally have $(dw(s))^2 = ds$ and $(dw(s))^3 = 0$.

In connection with the Ito Integral in (2.34) we also wish to define the integrals like

$$\int_0^t w(s)dX(s)\,, \qquad \int_0^t X(s)dX(s)$$

for some stochastic process $\{X(t)\}$. Of course these integrals are not always well defined. In Section 2.7 we define the above integrals using the Ito calculus.

We note in passing that the Ito integral in its simplest form naturally appears in

$$\mathcal{L}\left(\frac{1}{T}\sum_{t=1}^T y_{t-1}\varepsilon_t\right) \longrightarrow \mathcal{L}\left(\int_0^1 w(t)dw(t)\right) = \mathcal{L}\left(\frac{1}{2}\left(w^2(1)-1\right)\right)\,, \tag{2.43}$$

where $y_t = y_{t-1}+\varepsilon_t$, $y_0 = 0$ and $\{\varepsilon_t\}\sim$ i.i.d.(0,1). We already encountered a similar expression in (1.51). Although we discuss weak convergence like (2.43) in more general setting in Chapter 3, it is easy to prove (2.43) at this stage (Problem 5.4).

Problems

5.1 Show that the relation (2.40) holds.

5.2 Prove that $U(t)$ in (2.34) is an m.s. continuous martingale.

5.3 Derive the formulas in (2.41) and (2.42).

5.4 Prove the weak convergence in (2.43).

2.6. The mean square Ito integral : the vector case

In this section we extend the scalar m.s. Ito integral to the vector case. Although various cases may be possible, we concentrate here on the integral of the form

$$(2.44) \qquad V(t) = \int_0^t Y(s)dw'(s),$$

where $0 \le t \le 1$, $\{w(t)\}$ is the q-dimensional standard Brownian motion, and $\{Y(t)\}$ is a p-dimensional stochastic process each component of which belongs to L_2. Note that $V(t)$ in (2.44) is a $p \times q$ matrix. As in the scalar case we assume that i) $\{Y(t)\}$ is m.s. continuous on [0,1]; ii) $Y(t)$ is independent of $\{w(t_j) - w(t_i)\,;\, 0 \le t \le t_i \le t_j \le 1\}$ for all $t \in [0,1]$. Then the m.s. Ito integral for (2.44) exists and is unique.

As the simplest example let us put $Y(t) = w(t)$ so that

$$(2.45) \qquad V(t) = \int_0^t w(s)dw'(s)\,.$$

The (i,j)-element $V_{ij}(t)$ of $V(t)$ with $i \ne j$ and $t = 1$ has the following distributional relation (Problem 6.1) :

$$(2.46) \qquad \mathcal{L}(V_{ij}(1)) = \mathcal{L}\left(\frac{1}{2}\sum_{n=1}^{\infty} \frac{\xi_{1n}^2 - \xi_{2n}^2}{\left(n - \frac{1}{2}\right)\pi}\right), \quad (i \ne j)\,,$$

where $(\xi_{1n}, \xi_{2n})' \sim \text{NID}(0, I_2)$. This relation was earlier introduced in (1.49). Note that, unlike in the scalar case, we have

$$\int_0^t w(s)dw'(s) \ne \frac{1}{2}[w(t)w'(t) - tI_q]\,,$$

since the right side is symmetric while the left side not. We have (Problem 6.2)

$$(2.47) \qquad \int_0^t w(s)dw'(s) + \left(\int_0^t w(s)dw'(s)\right)' = w(t)w'(t) - tI_q\,.$$

The m.s. Ito integral for (2.45) with $t = 1$ appears in

$$(2.48) \qquad \mathcal{L}\left(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\varepsilon_t'\right) \longrightarrow \mathcal{L}\left(\int_0^1 w(s)dw'(s)\right),$$

where $y_t = y_{t-1} + \varepsilon_t$, $y_0 = 0$ and $\{\varepsilon_t\} \sim$ i.i.d.$(0, I_q)$. The proof for (2.48) is much involved, unlike in the scalar case. This will be discussed in Chapter 3.

As another example let us put $Y(t) = Aw(t)$ with A being any $q \times q$ constant matrix and consider $\mathrm{tr}(V(t))$ in (2.44), which is

$$\text{(2.49)} \qquad \mathrm{tr}\left(\int_0^t A\, w(s) dw'(s)\right) = \int_0^t w'(s) A' dw(s)\,.$$

We have already presented this integral in (1.58) and (1.59), where the former corresponds to A being symmetric, while the latter to A being skew symmetric ($A' = -A$). It can be shown (Problem 6.3) that, when A is symmetric with $\mathrm{tr}(A) = 0$, (2.49) with $t = 1$ follows the distribution of a finite sum of weighted, independent $\chi^2(1)$ random variables. When A is skew symmetric, the distribution of (2.49) is equal, in general, to that of an infinite sum of independent random variables, as was indicated in (1.59).

The m.s. Ito integral for (2.49) with $t = 1$ appears in

$$\text{(2.50)} \qquad \mathcal{L}\left(\frac{1}{T}\sum_{t=1}^{T} y'_{t-1} A' \varepsilon_t\right) \longrightarrow \mathcal{L}\left(\int_0^1 w'(s) A' dw(s)\right)\,,$$

where $y_t = y_{t-1} + \varepsilon_t$, $y_0 = 0$ and $\{\varepsilon_t\} \sim$ i.i.d.$(0, I_q)$. The weak convergence in (2.50) will also be discussed in Chapter 3.

Problems

6.1 Prove the distributional equivalence in (2.46).

6.2 Prove that the relation in (2.47) holds.

6.3 Prove that the distribution of (2.49) with $t = 1$ is equal to that of a finite sum of weighted independent $\chi^2(1)$ random variables when A is symmetric and $\mathrm{tr}(A)$=0.

2.7. The Ito calculus

This section applies the m.s. integrals defined in previous sections to consider the scalar integral equation of the form :

$$\text{(2.51)} \qquad X(t) = X(0) + \int_0^t \mu(X(s), s) ds + \int_0^t \sigma(X(s), s) dw(s)\,, \quad 0 \le t \le 1\,,$$

where we notice that there are two types of integrals. The one is the Riemann integral and the other the Ito integral. The possibility of whether these integrals can be defined in the m.s. sense, and, more importantly, the possibility of whether this integral equation has a unique m.s. solution $X(t)$, can be answered in the affirmative in the following theorem, whose proof is given in Jazwinski (1970, p.105).

Theorem 2.9. *Suppose that*

i) $X(0)$ *is any random variable with* $E(X^2(0)) < \infty$ *and is independent of* $\{w(t_j) - w(t_i)\,;\ 0 \le t_i \le t_j \le 1\}\,;$

ii) there is a positive constant K *such that*

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

$$|\mu(x,s) - \mu(x,t)| \le K|s-t|\,, \quad |\sigma(x,s) - \sigma(x,t)| \le K|s-t|\,,$$

$$|\mu(x,t)| \le K(1+x^2)^{\frac{1}{2}}\,, \quad |\sigma(x,t)| \le K(1+x^2)^{\frac{1}{2}}\,.$$

Then (2.51) has a unique m.s. continuous solution $X(t)$ *on* $[0,1]$ *such that* $X(t) - X(0)$ *is independent of* $\{w(t_j) - w(t_i)\,;\ 0 \le t \le t_i \le t_j\}$ *for every* $t \in [0,1]$.

This theorem ensures that the two integrals appearing in (2.51) are well defined in the m.s. sense. Note also that the solution process $\{X(t)\}$ is not m.s. differentiable in general. Nonetheless we write (2.51) as

$$dX(t) = \mu(X(t),t)dt + \sigma(X(t),t)dw(t)\,, \quad 0 \le t \le 1\,. \tag{2.52}$$

Here $dX(t)$ is called the *stochastic differential* of $X(t)$ and we call this equation the Ito *stochastic differential equation* (SDE), which is always understood in terms of the integral equation (2.51).

The idea of stochastic differentials can further be developed for functions of $X(t)$ and t. Namely we can consider another SDE that $f(X(t),t)$ satisfies on the basis of the following theorem (Jazwinski (1970, p.112)).

Ito's Theorem. *Suppose that $X(t)$ has the stochastic differential (2.52), where the conditions i) and ii) in Theorem 2.9 are satisfied. Let $f(x,t)$ denote a continuous function on $(-\infty,\infty)\times[0,1]$ with continuous partial derivatives $f_x = \partial f(x,t)/\partial x$, $f_{xx} = \partial^2 f(x,t)/\partial x^2$ and $f_t = \partial f(x,t)/\partial t$. Assume further that $f_t(X(t),t)$ and $f_{xx}(X(t),t)\sigma^2(X(t),t)$ are m.s. Riemann integrable. Then $f(X(t),t)$ has the stochastic differential :*

$$
(2.53)\qquad df(X(t),t) = f_x(X(t),t)dX(t) + \left(f_t(X(t),t) + \frac{1}{2}f_{xx}(X(t),t)\sigma^2(X(t),t)\right)dt\,.
$$

Ito's theorem tells us that, if $X(t)$ satisfies the SDE in (2.52), then $f(X(t),t)$ satisfies the SDE in (2.53). The SDE in (2.53) should be interpreted in terms of an integral equation like (2.51).

Some implications of Ito's theorem follow. If we expand $df(X(t),t)$ formally as

$$
df \sim f_x dX(t) + f_t dt + \frac{1}{2}(f_{xx}d^2X(t) + 2f_{xt}dX(t)dt + f_{tt}d^2t)\,,
$$

then (2.53) implies that we may put $d^2X(t) = \sigma^2(X(t),t)dt$, $dX(t)dt = 0$ and $d^2t = 0$. The corresponding integral equation to (2.53) may be written as

$$
(2.54)\qquad \int_a^b f_x(X(t),t)dX(t) = f(X(b),b) - f(X(a),a) - \int_a^b \left(f_t(X(t),t) + \frac{1}{2}f_{xx}(X(t),t)\sigma^2(X(t),t)\right)dt\,.
$$

The integral on the left side has never appeared before, but it can also be defined by using the Riemann integral as on the right side. This is an important message that Ito's theorem conveys to us. Ito's theorem is also useful for solving SDE's, as is exemplified shortly.

Three examples of stochastic differentials follow (Problem 7.1).

$$
(2.55)\qquad d(X^n(t)) = nX^{n-1}(t)dX(t) + \frac{n(n-1)}{2}X^{n-2}(t)\sigma^2(X(t),t)dt\,,
$$

$$
(2.56)\qquad d(w^n(t)) = nw^{n-1}(t)dw(t) + \frac{n(n-1)}{2}w^{n-2}(t)dt\,,
$$

$$
(2.57)\qquad d(e^{w(t)}) = e^{w(t)}dw(t) + \frac{1}{2}e^{w(t)}dt\,.
$$

Note that, when $n = 2$, (2.55) implies

$$\int_0^t X(s)dX(s) = \frac{1}{2}(X^2(t) - X^2(0)) - \frac{1}{2}\int_0^t \sigma^2(X(s), s)ds\,, \tag{2.58}$$

which reduces to (2.38) if $X(t) = w(t)$. The relation in (2.58) is quite important and will be frequently used in later chapters. We shall also use the following relation derived from (2.56):

$$\int_0^t w^n(s)dw(s) = \frac{1}{n+1}w^{n+1}(t) - \frac{n}{2}\int_0^t w^{n-1}(s)ds\,, \tag{2.59}$$

which also reduces to (2.38) if $n = 1$. The formulas (2.58) and (2.59) enable us to convert (extended) Ito integrals into the usual Riemann integrals.

Ito's theorem can be used to obtain a solution to the SDE (2.52). Two examples follow. The first is

$$dX(t) = X(t)dw(t)\,, \tag{2.60}$$

which has a solution $X(t) = X(0)\exp\left\{w(t) - \frac{t}{2}\right\}$ (Problem 7.2). The process $\{X(t)\}$ in this case is called the *geometric Brownian motion*. The second is

$$dX(t) = (\alpha X(t) + \beta)dt + \gamma dw(t)\,, \tag{2.61}$$

where α, β and γ are constants. We obtain (Problem 7.3)

$$X(t) = e^{\alpha t}X(0) + \frac{\beta(e^{\alpha t} - 1)}{\alpha} + \gamma e^{\alpha t}\int_0^t e^{-\alpha s}dw(s)\,. \tag{2.62}$$

When $\beta = 0$ and $\gamma = 1$, we have

Definition of the Ornstein-Uhlenbeck process. *The process* $\{X(t)\}$ *defined by*

$$dX(t) = \alpha X(t)dt + dw(t) \Longleftrightarrow X(t) = e^{\alpha t}X(0) + e^{\alpha t}\int_0^t e^{-\alpha s}dw(s) \tag{2.63}$$

is called the Ornstein-Uhlenbeck (O-U) process, where $X(0)$ *is independent of increments of* $\{w(t)\}$*.*

Note that $\{X(t)\}$ reduces to $\{w(t)\}$ when $\alpha = 0$ and $X(0) = 0$. Note also that, because of (2.63), the following integral can be well defined :

$$\int_0^t w(s)dX(s) = \alpha\int_0^t w(s)X(s)ds + \int_0^t w(s)dw(s)\,,$$

where the integral on the left side has never appeared before. It can also be shown (Problem 7.4) that, for the O-U process (2.63), $E(X(t)) = e^{\alpha t}E(X(0))$ and

$$\text{(2.64)} \qquad \text{Cov}(X(s), X(t)) = e^{\alpha(s+t)}\left[V(X(0)) + \frac{1 - e^{-2\alpha \min(s,t)}}{2\alpha}\right].$$

It turns out that, if α is positive, $V(X(t))$ is greater than $V(w(t))$ and increasing with t. On the other hand, if $\alpha < 0$ and $X(0) \sim \text{N}\left(0, -\frac{1}{2\alpha}\right)$, then $\{X(t)\}$ is a Gaussian, stationary process on [0,1] with $E(X(t)) = 0$ and $\text{Cov}(X(s), X(t)) = -e^{\alpha|s-t|}/(2\alpha)$ so that $V(X(t)) = -1/(2\alpha)$.

An example of the O-U process can be constructed, by substituting (2.4) into (2.63), as

$$\text{(2.65)} \qquad \begin{aligned} X(t) &= e^{\alpha t}X(0) \\ &+\sqrt{2}\sum_{n=1}^{\infty} \frac{\alpha e^{\alpha t} - \alpha\cos\left(n-\frac{1}{2}\right)\pi t + \left(n-\frac{1}{2}\right)\pi\sin\left(n-\frac{1}{2}\right)\pi t}{\left(n-\frac{1}{2}\right)^2\pi^2 + \alpha^2} Z_n. \end{aligned}$$

Figure 2.5 shows a sample path of the nonstationary O-U process simulated from (2.65) with $X(0) = 0$ and $\alpha = 1$, while Figure 2.6 a sample path of the stationary O-U process with $X(0) \sim \text{N}(0, 1/(-2\alpha))$ and $\alpha = -1$. The O-U process covers a wider class of stochastic processes than the Brownian motion.

Figure 2.5 Figure 2.6

The O-U process $\{X(t)\}$ in (2.63) will play an important role in subsequent chapters. It will be shown in Chapter 3 that $\{X(t)\}$ naturally appears in

$$\mathcal{L}\left(\frac{1}{\sqrt{T}}\, y_{[tT]}\right) \longrightarrow \mathcal{L}\left(X(t)\right),$$

where $[x]$ denotes the greatest integer not exceeding x, and

$$\text{(2.66)} \qquad y_t = \left(1 + \frac{\alpha}{T}\right) y_{t-1} + \varepsilon_t, \quad (t = 1, \cdots, T),$$

with $\{\varepsilon_t\} \sim \text{i.i.d.}\,(0,1)$, while the initial value y_0 is of the form $y_0 = \sqrt{T}X(0)$. The discrete-time process $\{y_t\}$ in (2.66) may be called the *near random walk* or the *near integrated process*. We shall also have

$$\text{(2.67)} \qquad \mathcal{L}\left(\frac{1}{T^2}\sum_{t=1}^{T} y_t^2\right) \longrightarrow \mathcal{L}\left(\int_0^1 X^2(t)dt\right).$$

It can be shown (Problem 7.5) that the m.s. Riemann integral in (2.67) can be expressed by using the m.s. Riemann-Stieltjes integral as

$$
\begin{aligned}
(2.68) \qquad \int_0^1 X^2(t)dt &= \frac{e^{2\alpha}-1}{2\alpha}X^2(0) + X(0)\int_0^1 \frac{e^{\alpha(2-t)}-e^{\alpha t}}{\alpha}\,dw(t) \\
&\quad + \int_0^1\int_0^1 \frac{e^{\alpha(2-s-t)}-e^{\alpha|s-t|}}{2\alpha}dw(s)dw(t)\,.
\end{aligned}
$$

Two approaches to deriving the c.f. of (2.68) will be presented in Chapters 4 and 5.

The discussions so far have been restricted to scalar cases, but can be extended to vector cases. We briefly mention an extended version of Ito's theorem (Jazwinski (1970, p.112)).

An Extended Version of Ito's Theorem. *Let the p-dimensional stochastic process $\{Y(t)\}$ be the unique solution of the Ito SDE :*

$$(2.69) \qquad dY(t) = \mu(Y(t),t)dt + G(Y(t),t)dw(t)\,,$$

where μ is $p\times 1$ and G is $p\times q$, while $\{w(t)\}$ is the q-dimensional standard Brownian motion. Let $g(y,t)$ be a real-valued function of $y : p\times 1$ and $t\in[0,1]$ with continuous partial derivatives $g_y = \partial g/\partial y$, $g_{yy} = \partial^2 g/(\partial y\partial y')$ and $g_t = \partial g/\partial t$. Assume further that $g_t(Y(t),t)$ and $G(Y(t),t)G'(Y(t),t)g_{yy}(Y(t),t)$ are m.s. Riemann integrable. Then $g(Y(t),t)$ has the stochastic differential

$$(2.70) \qquad dg = g_y'dY(t) + \left(g_t + \frac{1}{2}\mathrm{tr}\left(g_{yy}GG'\right)\right)dt\,.$$

As an application of this theorem let us consider the differential of $Y_1(t)Y_2(t)$, where $Y(t) = (Y_1(t), Y_2(t))'$ satisfies (2.69). Since $g(y,t) = y_1y_2$, we have

$$g_y = \begin{pmatrix} y_2 \\ y_1 \end{pmatrix}, \quad g_{yy} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad g_t = 0$$

so that (2.70) yields

$$(2.71) \qquad d(Y_1(t)Y_2(t)) = Y_1(t)dY_2(t) + Y_2(t)dY_1(t) + G_1'(Y(t),t)G_2(Y(t),t)dt\,,$$

where $G(Y(t),t) = (G_1(Y(t),t), G_2(Y(t),t))'$.

Similarly, we can show (Problem 7.6) that

$$(2.72)\qquad d\left(e^{\alpha t}w(t)\int_0^t e^{-\alpha s}dw(s)\right) = \\ w(t)dw(t) + dt + e^{\alpha t}(\alpha w(t)dt + dw(t))\int_0^t e^{-\alpha s}dw(s)\,,$$

$$(2.73)\qquad d\left(\exp\left(w(t)\int_0^t e^{-\alpha s}dw(s)\right)\right) = \left[\left(e^{-\alpha t}w(t) + \int_0^t e^{-\alpha s}dw(s)\right)dw(t) \right. \\ \left. + e^{-\alpha t}dt + \frac{1}{2}\left(e^{-\alpha t}w(t) + \int_0^t e^{-\alpha s}dw(s)\right)^2 dt\right]\exp\left(w(t)\int_0^t e^{-\alpha s}dw(s)\right)\,,$$

where $\{w(t)\}$ is the one-dimensional standard Brownian motion.

Problems

7.1 Obtain stochastic differentials in (2.55), (2.56) and (2.57).

7.2 Show that the SDE in (2.60) has a solution $X(t) = X(0)\exp\left\{w(t) - \frac{t}{2}\right\}$.

7.3 Show that the solution to the SDE in (2.61) is given by (2.62).

7.4 Derive the covariance in (2.64).

7.5 Prove that the relation in (2.68) holds.

7.6 Derive the relations in (2.72) and (2.73), where $\{w(t)\}$ is the one-dimensional standard Brownian motion.