Lecture Notes on Stochastic Processes (EE565)

Cheng-Shang Chang
Dept. of Electrical Engineering
National Tsing Hua University
Hsinchu 30043
Taiwan, R.O.C.

Dec. 1994

Brownian motions and stochastic calculus

Contents

1 Brownian motions

Definition 1 *A stochastic process* $\{B(t), t \geq 0\}$ *is called a* Brownian motion *process if*

(i) $B(0) = 0$;

(ii) $\{B(t), t \geq 0\}$ *has stationary and independent increments;*

(iii) *for every* $t > 0$, $B(t)$ *is normally distributed with mean 0 and variance* $\sigma^2 t$.

When $\sigma = 1$, the Brown motion is called standard Brownian motion.

- The first explanation was given by Einstein (1905)
- The above definition was given by Wiener (1918)

Theorem 2 *(Donsker's Theorem (1951)) Let* $\{\alpha_i, i \geq 1\}$ *be a sequence of i.i.d. random variables with mean 0 and variance* σ^2. *Consider the partial sum*

$$S_k = \sum_{i=1}^{k} \alpha_i, \quad S_0 = 0. \tag{1}$$

(Note that $S_k - S_l = \sum_{i=l+1}^{k} \alpha_i$*) Let* Y_t *be the linear interpolation of the partial sum* S_k*, i.e.,*

$$Y_t = S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)\alpha_{\lfloor t \rfloor + 1}, t \geq 0, t \in I\!R. \tag{2}$$

Consider the following sequence of stochastic processes:

$$X_t^{(n)} = \frac{Y_{nt}}{\sigma\sqrt{n}}. \tag{3}$$

Then as $n \to \infty$, $X_t^{(n)}$ *converges in distribution to the standard Brownian motion.*

(Sketch of the proof)

- Independent and stationary increments: Consider two disjoint intervals $(t_1, t_2]$ and $(t_3, t_4]$. Since $\{\alpha_i, i \geq 1\}$ are i.i.d., $X^{(n)}(t_2) - X^{(n)}(t_1)$ and $X^{(n)}(t_4) - X^{(n)}(t_3)$ are (almost) independent. Note that

$$X^{(n)}(t_2) - X^{(n)}(t_1) = \frac{Y(nt_2) - Y(nt_1)}{\sigma\sqrt{n}} \approx \frac{\sum_{i=\lfloor nt_1 \rfloor + 1}^{\lfloor nt_2 \rfloor} \alpha_i}{\sigma\sqrt{n}}. \tag{4}$$

  The argument for stationary increments is similar

- Normal distribution: Since

$$X^{(n)}(t) = \frac{Y(nt)}{\sigma\sqrt{n}} \approx \sqrt{t}\frac{\sum_{l=1}^{\lfloor nt \rfloor} \alpha_i}{\sigma\sqrt{nt}}. \tag{5}$$

  From the central limit theorem, the RHS of (5) converges to the normal distribution with mean 0 and variance $\sqrt{t}$.

- From these two observations, it follows that all the finite joint distributions of $X^n(t)$ converge to those of the standard Brownian motion. (Technical point: need tightness for the weak convergence)

Proposition 3 *(Markov property) A Brownian motion process is a* Markov *process.*

(Sketch of the proof) Let $B(t)$ be a Brownian motion. Then

$$\begin{aligned} &Pr(B(t+s) \leq a | B(s) = x, B(u), 0 \leq u \leq s) \\ &= Pr(B(t+s) - B(s) \leq a - x | B(s) = x, B(u), 0 \leq u \leq s). \end{aligned} \tag{6}$$

From the independent increment property,

$$\begin{aligned} &Pr(B(t+s) \leq a | B(s) = x, B(u), 0 \leq u \leq s) \\ &= Pr(B(t+s) - B(s) \leq a - x) = Pr(B(t+s) \leq a | B(s) = x). \end{aligned} \tag{7}$$

Distribution

- Marginal distribution:

$$Pr(B(t) \in (x, x+dx)) = f_t(x)dx = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} dx \tag{8}$$

- Joint distribution of $B(t_1), B(t_2), \ldots, B(t_n)$:

$$f(x_1, x_2, \ldots, x_n) = f_{t_1}(x_1) f_{t_2 - t_1}(x_2 - x_1) \cdots f_{t_n - t_{n-1}}(x_n - x_{n-1}). \tag{9}$$

- Conditional distribution: for $s < t$,

$$Pr(B(s) \in (x, x+dx) | B(t) = b) \tag{10}$$

$$= \frac{f_s(x) f_{t-s}(b-x)}{f_t(b)} dx \tag{11}$$

$$= c_1 \exp\left\{ \frac{-x^2}{2s} - \frac{(b-x)^2}{2(t-s)} \right\} dx \text{ for some constant } c_1 \tag{12}$$

$$= c_2 \exp\left\{ - \frac{t(x - bs/t)^2}{2s(t-s)} \right\} dx \text{ for some constant } c_2 \tag{13}$$

The conditional distribution of $B(s)$ given $B(t) = b$, $s < t$, is normal with mean bs/t and variance $s(t-s)/t$.

Definition 4 *(Gaussian process) A stochastic process $\{X(t), t \geq 0\}$ is a* Gaussian *process if $X(t_1), \ldots, X(t_n)$ has a multivariate normal distribution for all $t_1, \ldots, t_n$.*

Proposition 5 *The standard Brownian motion, denoted by $B(t)$, is a Gaussian process with $EB(t) = 0$ and $EB(t)B(s) = min(t, s)$. Conversely, if a Gaussian process has zero mean and the covariance function* $\min(s,t)$*, then it is the standard Brownian motion.*

Proof. It is clear that $EB(t) = 0$.

Note from the independent increment property that for $s \leq t$,

$$EB(s)B(t) = E[B(s)(B(t) - B(s) + B(s))] \tag{14}$$

$$= E[B(s)^2] + E[B(s)(B(t) - B(s))] \tag{15}$$

$$= E[B(s)^2] + E[B(s)]E[B(t) - B(s)] \tag{16}$$

$$= E[B(s)^2] = s. \tag{17}$$

To show the converse, note that the joint distribution of a multinormal distribution is determined by the marginal mean and the covariance values. □

Proposition 6 *(Equivalence transformation) If $B(t)$ is a standard Brownian motion, then the following processes are also the standard Brownian motion.*

(i) *Scaling: $Z(t) = B(nt)/\sqrt{n}$ for $n > 0$.*

(ii) *Time-inversion: $Z(t) = tB(1/t)$ for $t > 0$ and $Z(0) = 0$.*

(iii) *Time-reversal: $Z(t) = B(T) - B(T-t)$ for $T > 0$.*

(iv) *Symmetry: $Z(t) = -B(t)$.*

Proof. The proofs are similar. We only show (ii). Note that $Z(t)$ in (ii) is still a zero mean Gaussian process. We only need to check

$$EZ(t)Z(s) = st \min[\frac{1}{s}, \frac{1}{t}] = \min[s,t]. \tag{18}$$

□

Definition 7 *(Brownian bridge). Let $B(t)$ be the standard Brownian motion and $\{B_b(t) = (B(t)|B(1) = 0), 0 \le t \le 1\}$ is called the* Brownian bridge,

Proposition 8 *The Brownian bridge $B_b(t)$ is a Gaussian process with $EB_b(t) = 0$ and $EB_b(s)B_b(t) = s(1-t)$, $s < t < 1$.*

Proof. Using the same argument for (10), one can show that the Brownian bridge is a Gaussian process. Also, it follows (from the conditional mean) that $E(B_b(t)) = E(B(t)|B(1) = 0) = 0 \cdot t = 0$. Now for $s < t < 1$ (cf. the conditional variance),

$$E[B_b(s)B_b(t)] = E[B(s)B(t)|B(1) = 0] \tag{19}$$
$$= E[E[B(s)B(t)|B(t), B(1) = 0]|B(1) = 0] \tag{20}$$
$$= E[E[B(s)B(t)|B(t)]|B(1) = 0] \tag{21}$$
$$= E[B(t)E[B(s)|B(t)]|B(1) = 0] \tag{22}$$
$$= E[B(t)\frac{s}{t}B(t)|B(1) = 0] \tag{23}$$
$$= \frac{s}{t}E[B(t)^2|B(1) = 0] = \frac{s}{t}t(1-t) \tag{24}$$
$$= s(1-t). \tag{25}$$

□

Proposition 9 *Let $B(t)$ be the standard Brownian motion, then $\{Z(t) = B(t) - tB(1), 0 \le t \le 1\}$ is the Brownian bridge.*

Proof. It is clear that $Z(t)$ is Gaussian (why? conditioning on $B(1)$ first). Also $EZ(t) = EB(t) - tEB(1) = 0 - 0 = 0$. Thus, one only needs to show $EZ(s)Z(t) = s(1-t)$ for $s < t < 1$. Note that

$$EZ(s)Z(t) = E(B(s) - sB(1))(B(t) - tB(1)) \tag{26}$$
$$= EB(s)B(t) - sEB(1)B(t) - tEB(s)B(1) + stEB(1)^2 \tag{27}$$
$$= s - st - ts + st = s(1-t) \tag{28}$$

□

Definition 10 *(Reflected Brownian motion) Let $B(t)$ be the standard Brownian motion and $B_r(t) = |B(t)|$. Then $B_r(t)$ is called the reflected Brownian motion (at the origin).*

Note that

$$Pr(B_r(t) \le y) = Pr(-y \le B(t) \le y) \tag{29}$$
$$= Pr(B(t) \le y) - Pr(B(t) \le -y) \tag{30}$$
$$= Pr(B(t) \le y) - 1 + Pr(B(t) \ge -y) \tag{31}$$
$$= 2P(B(t) \le y) - 1 \tag{32}$$
$$= \frac{2}{\sqrt{2\pi t}} \int_{-\infty}^{y} e^{-x^2/2t} dx - 1 \tag{33}$$

One can easily compute (homework)

$$EB_r(t) = \sqrt{2t/\pi} \tag{34}$$
$$Var(B_r(t)) = (1 - \frac{2}{\pi})t \tag{35}$$

Definition 11 *(Geometric Brownian motion) Let $B(t)$ be the standard Brownian motion and $B_g(t) = e^{B(t)}$. Then $B_g(t)$ is called the geometric Brownian motion.*

Note that $E[B_g(t)] = Ee^{B(t)} = e^{t/2}$ (using the moment generating function of a normal distribution). Also,

$$Var(B_g(t)) = E[e^{2B(t)}] - (Ee^{B(t)})^2 = e^{2t} - e^t. \tag{36}$$

Definition 12 *(Integrated Brownian motion) Let $B(t)$ be the standard Brownian motion and $B_i(t) = \int_0^t B(s)ds$. Then $B_i(t)$ is called the integrated Brownian motion.*

Proposition 13 *The integrated Brownian motion $B_i(t)$ is a Gaussian process with $EB_i(t) = 0$ and $EB_i(s)B_i(t) = s^2(\frac{t}{2} - \frac{s}{6})$, $s \le t$.*

Proof. Note that if U_j, $j = 1, \ldots, m$, are independent normal random variables, then $W_i = \sum_{j=1}^m a_{i,j}U_j$, $i = 1, \ldots, m$, are jointly normal. Approximating the integrations by sums, one can show that $B_i(t_1), \ldots, B_i(t_n)$ are jointly normal for all $t_1, \ldots, t_m$. Thus, $B_i(t)$ is Gaussian.

Now compute

$$EB_i(t) = E\int_0^t B(s)ds = \int_0^t EB(s)ds = 0. \tag{37}$$

For $s \leq t$,

$$EB_i(s)B_i(t) = E\Big[\int_0^s B(y)dy \int_0^t B(u)du\Big] \tag{38}$$

$$= E\Big[\int_0^s \int_0^t B(y)B(u)dydu\Big] \tag{39}$$

$$= \int_0^s \int_0^t EB(y)B(u)dydu \tag{40}$$

$$= \int_0^s \int_0^t min(y,u)dydu \tag{41}$$

$$= \int_0^s \Big(\int_0^u ydy + \int_u^t udy\Big)du \tag{42}$$

$$= s^2(\frac{t}{2} - \frac{s}{6}). \tag{43}$$

□

Definition 14 *(Brownian motion with drift) Let $B(t)$ be the standard Brownian motion and $B_d(t) = \mu t + B(t)$. Then $B_d(t)$ is called the Brownian motion with drift coefficient μ.*

Facts: If $B_d(t)$ is the Brownian motion with drift coefficient μ, then

(i) $B_d(0) = 0$;

(ii) $\{B_d(t), t \geq 0\}$ has stationary and independent increments;

(iii) for every $t > 0$, $B_d(t)$ is normally distributed with mean μt and variance t.

Hitting time and the reflection principle (Désiré André)

Let $B(t)$ be the standard Brownian motion and

$$T_b = \inf\{t \geq 0 : B(t) = b\} \tag{44}$$

be the *first time* that $B(t)$ hits the level b. We will argue from "the reflection principle" that

$$Pr(T_b < t) = \sqrt{\frac{2}{\pi}} \int_{b/\sqrt{t}}^{\infty} e^{-x^2/2}dx. \tag{45}$$

Note that

$$Pr(T_b < t) = Pr(T_b < t, B(t) > b) + Pr(T_b < t, B(t) < b) + Pr(T_b < t, B(t) = b). \tag{46}$$

The last term in (46) is 0. Also,

$$Pr(T_b < t, B(t) > b) = Pr(B(t) > b) \tag{47}$$

since the Brownian motion must have crossed the level b some time before t given $B(t) > b$.

By symmetry of Brownian motion (cf. Proposition 6(iv)), for every sample path in $\{T_b > t, B(t) < b\}$, there is a corresponding sample path (with the same probability) in $\{T_b > t, B(t) > b\}$ (see Fig. xxx). Thus,

$$Pr(T_b < t, B(t) < b) = Pr(T_b < t, B(t) > b) = P(B(t) > b). \tag{48}$$

In conjunction with (46), one has

$$Pr(T_b < t) = 2Pr(B(t) > b) = \sqrt{\frac{2}{\pi}} \int_{b/\sqrt{t}}^{\infty} e^{-x^2/2} dx. \tag{49}$$

For further reading on hitting times, see e.g., Karatzas and Shreve [2], Chapter 2.6, and Ross [7], Chapter 6.

2 Sample path integrals and stochastic integrals

Consider the integral

$$\int_0^t \phi(s, \omega) dX(s, \omega). \tag{50}$$

for the stochastic processes $\{\phi(t, \omega), t \geq 0\}$ and $\{X(t, \omega), t \geq 0\}$. For every ω, $\{\phi(t, \omega), t \geq 0\}$ and $\{X(t, \omega), t \geq 0\}$ are *deterministic* functions, i.e., the sample paths of these two stochastic processes. Thus, the integral, conditioning on each ω, is a *Lebesgue-Stieltjes integral*.

2.1 Review of the Lebesgue-Stieltjes integral

Consider a random variable X with the distribution function $F(x)$. Note that a distribution function has the following properties:

- $F(-\infty) = 0$, $F(\infty) = 1$

- $F(x)$ is increasing (or nondecreasing)

$$F(a) \leq F(b), \text{ if } a < b.$$

- $F(x)$ is right continuous.

$$\lim_{x \downarrow a} F(x) = F(a).$$

For a (real valued) function $\phi : I\!R \mapsto I\!R$, one can write $E\phi(X)$ as follows:

$$E\phi(X) = \int_{(-\infty,\infty)} \phi(x) dF(x), \tag{51}$$

where the integral on the right-hand side is the Lebesgue integral. If both $\phi(x)$ and $F(x)$ are continuous, then the Lebesgue integral in (51) is equal to the Riemann integral, i.e.,

$$\int_{(-\infty,\infty)} \phi(x) dF(x) = \int_{-\infty}^{\infty} \phi(x) dF(x). \tag{52}$$

With this mind, one can extend the integral to the case when F is a *bounded* increasing function on $I\!R^+$ which is right continuous. To see this, let $F_1(x) = F(x)/F(\infty)$. Then $F_1(x)$ is a distribution function.

Definition 15 *(Bounded variation) Given a real function G, the variation of G over a finite interval $[a, b]$ is*

$$T_a^b = \sup_{n,(t_0,\ldots,t_n)} \sum_{k=0}^{n-1} |G(t_{k+1}) - G(t_k)|, \tag{53}$$

where $a = t_0 < t_1 \ldots < t_n = b$ is a partition of the interval $[a, b]$. The variation over $I\!R^+$ is equal to

$$T_0^\infty = \lim_{b \to \infty} T_0^b. \tag{54}$$

If $T_a^b < \infty$ (resp. $T_0^\infty < \infty$), then we say G has a bounded *variation over $[a, b]$ (resp. $I\!R^+$).*

Proposition 16 *If G has a bounded variation over $[a, b]$, then there exist two bounded increasing functions G_1 and G_2 such that $G(x) = G(a) + G_1(x) - G_2(x)$ for $a \leq x \leq b$.*

For the proof of Proposition 16, see e.g. Royden [8].

Example 17 *Consider the function* $G(x) = \frac{1}{4} - (x - \frac{1}{2})^2$ *over* $[0, 1]$*. Then one has (see Fig. XXX)*

$$G_1(x) = \begin{cases} \frac{1}{4} - (x - \frac{1}{2})^2 & \text{if } 0 \le x \le 1/2 \\ \frac{1}{4} & \text{if } 1/2 \le x \le 1 \end{cases} \tag{55}$$

and

$$G_2(x) = \begin{cases} 0 & \text{if } 0 \le x \le 1/2 \\ (x - \frac{1}{2})^2 & \text{if } 1/2 \le x \le 1 \end{cases} \tag{56}$$

Now suppose that G is also *right continuous*. Then the Lebesgue-Stieltjes integral for a function with bounded variation over $[a, b]$ can be defined as

$$\int_{[a,b]} \phi(s)dG(s) = \int_{[a,b]} \phi(s)dG_1(s) - \int_{[a,b]} \phi(s)dG_2(s) \tag{57}$$

provided that both integrals on the right-hand side are *finite*.

In particular, if a function G that has a continuous first derivative G' over $[a, b]$, then G has a bounded variation over $[a, b]$ since

$$\sum_{k=0}^{n-1} |G(t_{k+1}) - G(t_k)| = \sum_{k=0}^{n-1} | \int_{t_k}^{t_{k+1}} G'(s)ds| \tag{58}$$

$$\le \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} |G'(s)|ds \tag{59}$$

$$= \int_a^b |G'(s)|ds < \infty \tag{60}$$

In this case,

$$\int_a^b \phi(s)dG(s) = \int_a^b \phi(s)G'(s)ds. \tag{61}$$

2.2 Sample path properties of Brownian motions

In this section, let $B(t)$ be the standard Brownian motion.

Proposition 18 *The sample path of the standard Brownian motion is continuous in probability, i.e., for any* $\epsilon > 0$

$$\lim_{h \to 0} Pr(|B(t + h) - B(t)| > \epsilon) = 0. \tag{62}$$

Proof. Note that $B(t) - B(s)$ is a normal r.v. with mean 0 and variance $t - s$, $s < t$. Thus,

$$E(B(t) - B(s))^2 = t - s. \tag{63}$$

Using Chebychev's inequality,

$$Pr(|B(t+h) - B(t)| > \epsilon) = Pr((B(t+h) - B(t))^2 > \epsilon^2) \leq \frac{h}{\epsilon^2}. \tag{64}$$

Letting $h \to 0$ completes the argument. □

In fact, a stronger continuity result is available (the proof will not be discussed here). With probability one, the sample path of the standard Brownian motion is *uniformly continuous* on every finite interval, i.e.,

$$Pr(\cup_{t\in[a,b]}\{\omega : \lim_{h\to 0} |B(t+h) - B(t)| \neq 0\}) = 0. \tag{65}$$

Proposition 19 *The sample path of the standard Brownian motion is almost surely not differentiable. In fact,*

$$Pr(\limsup_{h\to 0} \Big|\frac{B(t+h) - B(t)}{h}\Big| = \infty) = 1. \tag{66}$$

Proof. From the stationary increment property, equation (66) is equivalent to

$$Pr(\limsup_{h\to 0} \Big|\frac{B(h)}{h}\Big| > d) = 1, \tag{67}$$

for any $d > 0$. Define the events

$$E_h = \{\omega : \sup_{0\leq s\leq h} \Big|\frac{B(s)}{s}\Big| > d\}. \tag{68}$$

For any sequence $h_n \downarrow 0$, we have $E_{h_n} \supset E_{h_{n+1}}$. Thus,

$$Pr(\lim_{n\to\infty} E_{h_n}) = \lim_{n\to\infty} Pr(E_{h_n}). \tag{69}$$

(Note that $\lim_{n\to\infty} E_{h_n} = \cap_{n=1}^{\infty} E_{h_n}$.)

Using the scaling property in Proposition 6(i), one has

$$Pr(E_{h_n}) \geq Pr(|B(h_n)/h_n| > d) = Pr(|B(1)| > \sqrt{h_n} d), \tag{70}$$

which tends to one as $n \to \infty$ ($h_n \to 0$). □

Proposition 20 *(Quadratic variation) Consider the closed interval* $[a,b]$*. Let* $\pi_n = [a = t_0^{(n)} < t_1^{(n)} < \ldots < t_{m_n}^{(n)} = b]$*,* $n = 1, 2, \ldots,$ *be a sequence of partitions of* $[a,b]$ *such that*

$$\Delta_n = \max_{1 \le m \le m_n} (t_{m+1}^{(n)} - t_m^{(n)}) \to 0, \tag{71}$$

as $n \to \infty$*. Then*

$$\lim_{n\to\infty} E\Big(\sum_{m=1}^{m_n} (B(t_{m+1}^{(n)}) - B(t_m^{(n)}))^2 - (b-a)\Big)^2 = 0. \tag{72}$$

If $\sum_{n=1}^{\infty} \Delta_n < \infty$*, then with probability one*

$$\lim_{n\to\infty} \sum_{m=1}^{m_n} (B(t_{m+1}^{(n)}) - B(t_m^{(n)}))^2 = b - a. \tag{73}$$

Proof. To show (72), we write

$$S_n = \sum_{m=1}^{m_n} (B(t_{m+1}^{(n)}) - B(t_m^{(n)}))^2 - (b-a) \tag{74}$$

$$= \sum_{m=1}^{m_n} [(B(t_{m+1}^{(n)}) - B(t_m^{(n)}))^2 - (t_{m+1}^{(n)} - t_m^{(n)})]. \tag{75}$$

From the independent increment property and (63), it follows that S_n is a sum of independent random variables with zero mean. Thus, $ES_n = 0$ and

$$ES_n^2 = \sum_{m=1}^{m_n} E[(B(t_{m+1}^{(n)}) - B(t_m^{(n)}))^2 - (t_{m+1}^{(n)} - t_m^{(n)})]^2, \tag{76}$$

since the cross terms are 0. Note that $B(t_{m+1}^{(n)}) - B(t_m^{(n)})$ is a normal random variable with zero mean and variance $t_{m+1}^{(n)} - t_m^{(n)}$. Since $EX^4 = 3$ for a standard normal random variable X,

$$ES_n^2 = 2\sum_{m=1}^{m_n} (t_{m+1}^{(n)} - t_m^{(n)})^2$$

$$\le 2\Delta_n \sum_{m=1}^{m_n} (t_{m+1}^{(n)} - t_m^{(n)}) = 2\Delta_n (b-a). \tag{77}$$

Since $\Delta_n \to 0$,

$$\lim_{n\to\infty} ES_n^2 = 0. \tag{78}$$

For (73), we have from Chebychev's inequality that

$$Pr(|S_n| > \epsilon) \leq \frac{ES_n^2}{\epsilon^2} \leq 2(b-a)\frac{\Delta_n}{\epsilon^2}. \tag{79}$$

Since we assume that $\sum_{n=1}^{\infty} \Delta_n < \infty$, we have for every $\epsilon > 0$ that

$$\sum_{n=1}^{\infty} Pr(|S_n| > \epsilon) < \infty. \tag{80}$$

As a direct result of the Borel-Cantelli lemma,

$$Pr(|S_n| > \epsilon, i.o.) = 0. \tag{81}$$

Thus,

$$Pr(\lim_{n\to\infty} S_n = 0) = 1. \tag{82}$$

□

To meet the condition that $\sum_{n=1}^{\infty} \Delta_n < \infty$, one can choose $t_m^{(n)} = a + (m(b-a)/2^n)$. In this case, $\Delta_n = 1/2^n$.

Corollary 21 *With probability one, the standard Brownian motion has* un-bounded *variation over every bounded interval.*

Proof. This can be argued by contradiction. Suppose that the sample path of the standard Brownian motion has a bounded variation over $[a, b]$, denoted by T_a^b (almost surely). Since

$$\sum_{m=1}^{m_n} (B(t_{m+1}^{(n)}) - B(t_m^{(n)}))^2 \leq \max_{1\leq m\leq m_n} |B(t_{m+1}^{(n)}) - B(t_m^{(n)})| \sum_{m=1}^{m_n} |B(t_{m+1}^{(n)}) - B(t_m^{(n)})|$$
$$\leq \max_{1\leq m\leq m_n} |B(t_{m+1}^{(n)}) - B(t_m^{(n)})| T_a^b, \tag{83}$$

letting $n \to \infty$ yields ($\max_{1\leq m\leq m_n} |B(t_{m+1}^{(n)}) - B(t_m^{(n)})| \to 0$ from the continuity of the sample path of the Brownian motion)

$$\sum_{m=1}^{m_n} (B(t_{m+1}^{(n)}) - B(t_m^{(n)}))^2 = 0 \;\; a.s. \tag{84}$$

which contradicts Proposition 20 (when $\sum_{n=1}^{\infty} \Delta_n < \infty$). □

2.3 Stochastic integrals

Now we know the standard Brownian motion does not have bounded variations for its sample paths. How do we define the sample path integral $\int_a^b \phi(s)dB(s)$ for the standard Brownian motion $B(t)$?

In view of the integral, we need to clarify the relation between $\{\phi(t), t \geq 0\}$ and $\{B(t), t \geq 0\}$. To do so, we strengthen the definition of the standard Brownian motion as follows:

Definition 22 *A stochastic process $\{B(t), t \geq 0\}$ is called a* Brownian motion *process if*

(i) *Initial condition: $B(0) = 0$;*

(ii) *Independent increments: $B(t)-B(s)$ is independent of $\{B(\tau), \phi(\tau), 0 \leq \tau \leq s\}$;*

(iii) *Normal distribution: $B(t) - B(s)$ is normally distributed with mean 0 and variance $t - s$.*

Note that the above definition is the same Definition 1 except we have an additional assumption on independence between the increment of the standard Brownian motion and $\{\phi(t)\}$ before the increment. In general, one can define a "history" $\{\mathcal{F}_t, t \geq 0\}$ generated by the σ-algebra of $\{B(\tau), \phi(\tau), 0 \leq \tau \leq t\}$ (and other stochastic processes). The independent increment property for the standard Brownian motion then holds with respect to the history $\{\mathcal{F}_t, t \geq 0\}$.

Definition 23 *(Simple processes) A process $\{\phi(t), t \geq 0\}$ is called* simple *if there exists a strictly increasing sequence of real numbers $\{t_n\}_{n=0}^{\infty}$ with $t_0 = 0$ such that $\lim_{n\to\infty} t_n = \infty$, as well as a sequence of* bounded *random variables $\{\eta_n\}_{n=0}^{\infty}$, i.e., $\sup_{n\geq 0} |\eta_n| \leq C < \infty$, such that*

$$\phi(t) = \eta_0 \mathbf{1}_{\{0\}}(t) + \sum_{n=0}^{\infty} \eta_n \mathbf{1}_{\{(t_n, t_{n+1}]\}}(t); \quad 0 \leq t < \infty, \tag{85}$$

where $\mathbf{1}_{(a,b]}(t) = 1$ if $t \in (a, b]$ and 0 otherwise.

In words,

$$\text{if } \ t_n < t \leq t_{n+1}, \ \text{ then } \ \phi(t) = \eta_n. \tag{86}$$

Also, it follows from the independent increment property of the standard Brownian motion that $B(t + s) - B(t)$, $s > 0$, is independent of $\eta_m, m = 1, \ldots, n$. This observation is crucial in the following development.

In view of (86),

$$\int_0^t \phi(s)ds = \sum_{m=0}^{n-1} \eta_m(t_{m+1} - t_m) + \eta_n(t - t_n). \tag{87}$$

Moreover, for any continuous function f, the stochastic process $\{f(\phi(t)), t \geq 0\}$ is also a simple process with the sequence of random variables $\{f(\eta_n)\}_{n=0}^{\infty}$.

Definition 24 *(Stochastic integrals for simple processes) If $\{\phi(t), t \geq 0\}$ is a simple process in Definition 23 and $\{B(t), t \geq 0\}$ is the standard Brownian motion, then we define the stochastic integral over the interval $[0, t]$ as follows:*

$$\int_0^t \phi(s)dB(s) = \sum_{m=0}^{n-1} \eta_m[B(t_{m+1}) - B(t_m)] + \eta_n[B(t) - B(t_n)] \tag{88}$$

where $t_n \leq t < t_{n+1}$.

To simplify the notations, let

$$I_t(\phi) = \int_0^t \phi(s)dB(s). \tag{89}$$

(Note that $I_t(\phi)$ is a random variable.)

Proposition 25 *If $\{\phi(t), t \geq 0\}$ are $\{\psi(t), t \geq 0\}$ are simple processes in Definition 23, then*

(i) $I_0(\phi) = 0$.

(ii) $EI_t(\phi) = 0$.

(iii) $E(I_t(\phi))^2 = E\int_0^t \phi^2(s)ds$.

(iv) $I_t(c_1\phi + c_2\psi) = c_1 I_t(\phi) + c_2 I_t(\psi)$ *for all $c_1, c_2 \in I\!R$ (assuming the independent increment property holds with respect to the history generated by ϕ, ψ and the standard Brownian motion itself).*

Proof. Properties (i) and (iv) are obvious (since the definition in (88) is linear).

Since $B(t_{j+1}) - B(t_j)$ is independent of $\{\eta_m, m = 1, \ldots, j\}$ and $\{B(\tau), 0 \leq \tau \leq t_j\}$, and $E[B(t_{j+1} - B(t_j)] = 0$, it then follows from (88) that

$$EI_t(\phi) = \sum_{m=0}^{n-1} E\eta_m E[B(t_{m+1}) - B(t_m)] + E\eta_n E[B(t) - B(t_n)] = 0. \tag{90}$$

To show (iii), note from (88) that

$$E(I_t(\phi))^2 = E\Big(\sum_{m=0}^{n-1} \eta_m[B(t_{m+1}) - B(t_m)] + \eta_n[B(t) - B(t_n)]\Big)^2 \quad (91)$$

$$= E(\sum_{m=0}^{n-1} \eta_m^2[B(t_{m+1}) - B(t_m)]^2 + E\eta_n^2[B(t) - B(t_n)]^2 \quad (92)$$

since all the cross terms are 0. To see this, note for $i < j$

$$E\Big(\eta_i[B(t_{i+1} - B(t_i)]\eta_j[B(t_{j+1} - B(t_j)]\Big)$$
$$= E(\eta_i[B(t_{i+1} - B(t_i)]\eta_j)E[B(t_{j+1} - B(t_j)] = 0, \quad (93)$$

where once again we use the facts that $B(t_{j+1}) - B(t_j)$ is independent of $\{\eta_m, m = 1, \ldots, j\}$ and $\{B(\tau), 0 \leq \tau \leq t_j\}$, and that $E[B(t_{j+1} - B(t_j)] = 0$. Using the independent increment property and $E[B(t_{m+1}) - B(t_m)]^2 = t_{m+1} - t_m$ in (91), one has

$$E(I_t(\phi))^2 = \sum_{m=0}^{n-1} E\eta_m^2(t_{m+1} - t_m) + E\eta_n^2(t - t_n) \quad (94)$$

$$= E\Big(\sum_{m=0}^{n-1} \eta_m^2(t_{m+1} - t_m) + \eta_n^2(t - t_n)\Big) \quad (95)$$

$$= E\int_0^t (\phi(s))^2 ds. \quad (96)$$

□

Now we define the stochastic integrals for processes that satisfy a L^2 condition.

Definition 26 *(Stochastic integrals for L^2 processes) Consider a stochastic process $\{\phi(t), t \geq 0\}$. If*

$$E\int_0^t (\phi(s))^2 ds < \infty, \quad (97)$$

then $\int_0^t \phi(s)dB(s)$ is defined as the limit of the stochastic integrals of the sequence of simple processes $\{\phi^{(n)}\}_{n=1}^{\infty}$ that converges to ϕ, i.e.,

$$I_t(\phi) = \lim_{n\to\infty} I_t(\phi^{(n)}). \quad (98)$$

The convergence of the above limit is in L^2, i.e.,

$$\lim_{n\to\infty} E(I_t(\phi) - I_t(\phi^{(n)}))^2 = 0. \tag{99}$$

We sketch the proof for the L^2 convergence (for the case that ϕ is uniformly continuous) as follows: suppose that ϕ is uniformly continuous. Consider the sequence of simple processes

$$\phi^{(n)}(s) = \phi(0)\mathbf{1}_{\{0\}}(s) + \sum_{k=0}^{2^n-1} \phi(\frac{kt}{2^n})\mathbf{1}_{\{(kt/2^n,(k+1)t/2^n]\}}(s) \quad 0 \le s < t, n \ge 1. \tag{100}$$

Since the sample path is uniformly continuous, one can show the Cauchy convergence criterion(by the bounded convergence theorem) that

$$\lim_{n\to\infty} \sup_{m>n} E\int_0^t |\phi^{(n)}(s) - \phi^{(m)}(s)|^2 ds = 0 \tag{101}$$

Using Proposition 25(iii) and (iv), one has

$$\lim_{n\to\infty} \sup_{m>n} E|I_t(\phi^{(n)}) - I_t(\phi^{(m)})|^2 \tag{102}$$

$$= \lim_{n\to\infty} \sup_{m>n} E|I_t(\phi^{(n)} - \phi^{(m)})|^2 \tag{103}$$

$$= \lim_{n\to\infty} \sup_{m>n} E\int_0^t |\phi^{(n)}(s) - \phi^{(m)}(s)|^2 ds = 0. \tag{104}$$

Proposition 27 *Proposition 25 holds for the stochastic integrals in Definition 26.*

Proof. Properties (i)(ii) and (iv) are obvious. To show (iii), note from Proposition 25(iii) and the L^2 convergence that

$$E(I_t(\phi))^2 = E\lim_{n\to\infty}(I_t(\phi^{(n)}))^2 = \lim_{n\to\infty} E(I_t(\phi^{(n)}))^2 \tag{105}$$

$$= \lim_{n\to\infty} E\int_0^t (\phi^{(n)}(s))^2 ds = E\int_0^t (\phi(s))^2 ds. \tag{106}$$

□

Corollary 28

$$E[I_t(\phi)I_t(\psi)] = E\int_0^t \phi(s)\psi(s) ds \tag{107}$$

(assuming the independent increment property holds with respect to the history generated by ϕ, ψ and the standard Brownian motion itself).

Proof. Note from Proposition 25(iv) that

$$E[I_t(\phi+\psi)]^2 = E[I_t(\phi)+I_t(\psi)]^2 \tag{108}$$

$$= E[I_t(\phi)]^2 + 2E[I_t(\phi)I_t(\psi)] + E[I_t(\psi)]^2. \tag{109}$$

Applying Proposition 25(iii) yields

$$E[I_t(\phi+\psi)]^2 = E\int_0^t (\phi(s)+\psi(s))^2 ds, \tag{110}$$

$$E[I_t(\phi)]^2 = E\int_0^t (\phi(s))^2 ds, \tag{111}$$

$$E[I_t(\psi)]^2 = E\int_0^t (\psi(s))^2 ds. \tag{112}$$

In conjunction with (107), one has

$$E[I_t(\phi)I_t(\psi)] = E\int_0^t \phi(s)\psi(s)ds. \tag{113}$$

□

Proposition 29 *Consider a family of deterministic functions* $\{h(t,s), 0 \le s,t\}$ *and the stochastic process generated by the following stochastic integrals* $\{X(t) = \int_0^t h(t,s)dB(s), t \ge 0\}$. *Then* $\{X(t), t \ge 0\}$ *is a* Gaussian *process with* $EX(t) = 0$ *and*

$$EX(s)X(t) = \int_0^{\min[s,t]} h(t,u)h(s,u)du. \tag{114}$$

Intuitively, one can view $X(t)$ as the output process from a linear time varying filter subject to the standard Brownian motion input. The impulse response of the filter is $h(t,s)$.

Proof.

In view of (88) and the argument used in Proposition 13 for the integrated Brownian motion, $\{X(t), t \ge 0\}$ is a *Gaussian* process ($X(t)$ can be approximated by a sum of independent normal random variables cf. (88)). That $EX(t) = 0$, $t \ge 0$, follows from Proposition 25(ii). To compute the covariance values, let $\phi(u) = h(t,u)$ and $\psi(u) = h(s,u)$. Since $\{h(t,s), t,s \ge 0\}$ are a family of deterministic functions, both ϕ and ψ are deterministic functions. It then follows from the independent increment property of the standard Brownian motion that $I_s(\phi)$ is independent of $I_t(\psi) - I_s(\psi)$ (cf. (88) and note that $I_s(\phi)$ is determined by the standard Brownian motion from 0 to s and that $I_t(\psi) - I_s(\psi)$ is determined by the standard Brownian motion from s to t). Using this "independent increment property" and $EI_s(\phi) = 0$, one has for $0 \le s \le t$,

$$EX(s)X(t) = EI_s(\phi)I_t(\psi) = E\Big(I_s(\phi)[I_s(\psi) + I_t(\psi) - I_s(\psi)]\Big) \tag{115}$$
$$= E[I_s(\phi)I_s(\psi)] + EI_s(\phi)E(I_t(\psi) - I_s(\psi)) \tag{116}$$
$$= E[I_s(\phi)I_s(\psi)]. \tag{117}$$

It then follows from Corollary 28 that

$$E[I_s(\phi)I_s(\psi)] = \int_0^s \phi(u)\psi(u)du \tag{118}$$
$$= \int_0^s h(t,u)h(s,u)du. \tag{119}$$

The case for $0 \le t \le s$ is similar. □

3 Itô's calculus

Let $f : I\!R \mapsto I\!R$ be a function that f' is continuous. For a process ψ with a bounded variation, one has (from the usual calculus)

$$df(\psi(t)) = f'(\psi(t))d\psi(t), \tag{120}$$

or equivalently,

$$f(\psi(t)) = f(\psi(0)) + \int_0^t f'(\psi(s))d\psi(s). \tag{121}$$

One might wonder if the calculus can be extended to Brownian motions. For instance, do we still have the following identity

$$f(B(t)) = f(B(0)) + \int_0^t f'(B(s))dB(s)? \tag{122}$$

The answer is no, and we need an extra correction term. The calculus was carried out by Itô (1944).

Theorem 30 *(Itô's calculus, cf. Wong and Hajek [9], Proposition 4.3.2 and Karatzas and Shreve [2], Theorem 3.3.3) Let $f : I\!R \mapsto I\!R$ be a function that f'' is continuous. Suppose that*

$$\phi(t) = \phi(0) + \psi(t) + B(t), \tag{123}$$

where $\psi(t)$ is a continuous *process with a bounded variation and $B(t)$ is the standard Brown motion. Then*

$$f(\phi(t)) = f(\phi(0)) + \int_0^t f'(\phi(s))d\psi(s) + \int_0^t f'(\phi(s))dB(s)$$
$$+\frac{1}{2}\int_0^t f''(\phi(s))ds, 0 \le t < \infty, \ a.s. \tag{124}$$

Remark 31 Equation (124) can also be written in differential notation:

$$df(\phi(t)) = f'(\phi(t))d\psi(t) + f'(\phi(t))dB(t) + \frac{1}{2}f''(\phi(t))dt, \tag{125}$$
$$= f'(\phi(t))d\phi(t) + \frac{1}{2}f''(\phi(t))dt, 0 \le t < \infty, \ a.s. \tag{126}$$

This is the "chain-rule" for stochastic calculus.

Though the rule stated in Theorem 30 is for $\phi(t) = \phi(0) + \psi(t) + B(t)$, it can be extended to the case that includes stochastic integrals as follows:

$$\phi(t) = \phi(0) + \int_0^t \psi_1(s)ds + \int_0^t \psi_2(s)dB(s), \tag{127}$$

where the increment of the standard Brownian motion is independent of the history of ψ_1, ψ_2, and the standard Brownian motion itself. In this case, $\int_0^t \psi_1(s)ds$ is the term with a bounded variation and $\int_0^t \psi_2(s)dB(s)$ is the generalization from Brownian motions to stochastic integrals. Then for the function f considered in Theorem 30, one has the following chain-rule:

$$f(\phi(t)) = f(\phi(0)) + \int_0^t f'(\phi(s))\psi_1(s)ds + \int_0^t f'(\phi(s))\psi_2(s)dB(s)$$
$$+\frac{1}{2}\int_0^t f''(\phi(s))(\psi_2(s))^2ds, 0 \le t < \infty, \ a.s. \tag{128}$$

In differential notation, one has

$$df(\phi(t)) = f'(\phi(t))\psi_1(t)dt + f'(\phi(t))\psi_2(t)dB(t) + \frac{1}{2}f''(\phi(t))^2(\psi_2(t))^2dt$$
$$= f'(\phi(t))d\phi(t) + \frac{1}{2}f''(\phi(t))^2(\psi_2(t))^2dt, 0 \le t < \infty, \ a.s. \tag{129}$$

We can further allow the dependence through time. Let $f(x_1, x_2)$ be a real valued function that has continuous second partial with respect to x_1 and continuous first partial with respect to x_2. Then

$$f(\phi(t),t) = f(\phi(0),0) + \int_0^t \frac{\partial f(\phi(s),s)}{\partial x_2} ds + \int_0^t \frac{\partial f(\phi(s),s)}{\partial x_1} \psi_1(s) ds$$
$$+ \int_0^t \frac{\partial f(\phi(s),s)}{\partial x_1} \psi_2(s) dB(s) + \frac{1}{2} \int_0^t \frac{\partial^2 f(\phi(s),s)}{\partial {x_1}^2} (\psi_2(s))^2 ds,$$
$$0 \le t < \infty, \ a.s. \tag{130}$$

In differential notation, one has

$$df(\phi(t),t) = \frac{\partial f(\phi(t),t)}{\partial x_2} dt + \frac{\partial f(\phi(t),t)}{\partial x_1} \psi_1(t) dt + \frac{\partial f(\phi(t),t)}{\partial x_1} \psi_2(t) dB(t)$$
$$+ \frac{1}{2} \frac{\partial^2 f(\phi(t),t)}{\partial x_1^2} (\psi_2(t))^2 dt, 0 \le t < \infty, \ a.s. \tag{131}$$

Proof. (Sketch of the proof for Theorem 30) Fix $t > 0$ and consider a partition $\pi = [0 = t_0 < t_1 < \ldots < t_m = t]$. Applying Taylor's expansion yields

$$f(\phi(t)) - f(\phi(0)) = \sum_{k=1}^{m} \{f(\phi(t_k)) - f(\phi(t_{k-1}))\}$$
$$= \sum_{k=1}^{m} f'(\phi(t_{k-1}))[\phi(t_k) - \phi(t_{k-1})] + \frac{1}{2} \sum_{k=1}^{m} f''(\xi_k)[\phi(t_k) - \phi(t_{k-1})]^2, \tag{132}$$

where $\xi_k = \phi(t_{k-1}) + \theta_k(\phi(t_k) - \phi(t_{k-1}))$ for some $0 \le \theta_k \le 1$ (note that both x_k and θ_k are random variables). Thus, we have

$$f(\phi(t)) - f(\phi(0)) = J_1(\pi) + J_2(\pi) + \frac{1}{2} J_3(\pi), \tag{133}$$

where

$$J_1(\pi) = \sum_{k=1}^{m} f'(\phi(t_{k-1}))[\psi(t_k) - \psi(t_{k-1})] \tag{134}$$
$$J_2(\pi) = \sum_{k=1}^{m} f'(\phi(t_{k-1}))[B(t_k) - B(t_{k-1})] \tag{135}$$
$$J_3(\pi) = \sum_{k=1}^{m} f''(\xi_k)[\phi(t_k) - \phi(t_{k-1})]^2. \tag{136}$$

It is easy to see that $J_1(\pi)$ converges to the Lebesgue-Stieltjes integral $\int_0^t f'(\phi(s))d\psi(s)$ as $\max_{1\le k\le m}|t_k - t_{k-1}| \to 0$. Similarly, using the approximation of simple processes, $J_2(\pi)$ converges to the stochastic integral $\int_0^t f'(\phi(s))dB(s)$. Now write $J_3(\pi)$ as

$$J_3(\pi) = J_4(\pi) + J_5(\pi) + J_6(\pi), \tag{137}$$

where

$$J_4(\pi) = \sum_{k=1}^{m} f''(\xi_k)[\psi(t_k) - \psi(t_{k-1})]^2 \tag{138}$$

$$J_5(\pi) = 2\sum_{k=1}^{m} f''(\xi_k)[\psi(t_k) - \psi(t_{k-1})][B(t_k) - B(t_{k-1})] \tag{139}$$

$$J_6(\pi) = \sum_{k=1}^{m} f''(\xi_k)[B(t_k) - B(t_{k-1})]^2 \tag{140}$$

Since we assume that f'' is continuous, $\sup_{0\le s\le t} f''(s) \le c_1 < \infty$ for some constant c_1 (a continuous function in a bounded interval is bounded). Also, from the assumption that ψ has a bounded variation, we have

$$\sum_{k=1}^{m} |\psi(t_k) - \psi(t_{k-1})| \le c_2 < \infty \tag{141}$$

for some constant c_2. Thus,

$$\begin{aligned} &|J_4(\pi)| + |J_5(\pi)| \\ &\le 2c_1c_2(\max_{1\le k\le m}|\psi(t_k) - \psi(t_{k-1})| + \max_{1\le k\le m}|B(t_k) - B(t_{k-1})|). \end{aligned} \tag{142}$$

From the continuity of the processes ψ and B, the last term in (142) converges to 0 as $\max_{1\le k\le m}|t_k - t_{k-1}| \to 0$. Write

$$J_6(\pi) = J_6^*(\pi) - J_7(\pi) + J_6(\pi) - J_6^*(\pi) + J_7(\pi), \tag{143}$$

where

$$J_6^*(\pi) = \sum_{k=1}^{m} f''(\phi(t_{k-1}))[B(t_k) - B(t_{k-1})]^2 \tag{144}$$

$$J_7(\pi) = \sum_{k=1}^{m} f''(\phi(t_{k-1}))[t_k - t_{k-1}] \tag{145}$$

Observe that

$$|J_6^*(\pi) - J_6(\pi)| \le \max_{1\le k\le m} |f''(\xi_k) - f''(\phi(t_k))| \sum_{k=1}^{m} [B(t_k) - B(t_{k-1})]^2. \quad (146)$$

In view of (73) and the continuity of f'', $|J_6^*(\pi) - J_6(\pi)|$ converges to 0. Following the same argument for (73), one can show that $|J_6^*(\pi) - J_7(\pi)|$ converges to 0. Thus, we are left with the extra term $J_7(\pi)$. □

Example 32 *Let $f(x) = x^2$ and $\phi(t) = B(t)$. Then $f'(x) = 2x$ and $f''(x) = 2$. Thus,*

$$dB^2(t) = 2B(t)dB(t) + dt, \quad (147)$$

or

$$B^2(t) = 2\int_0^t B(s)dB(s) + t. \quad (148)$$

Example 33 *Let*

$$\phi(t) = \int_0^t \psi(s)dB(s) - \frac{1}{2}\int_0^t (\psi(s))^2 ds \quad (149)$$

and $f(x) = e^x$. Then

$$e^{\phi(t)} = 1 + \int_0^t e^{\phi(s)}\psi(s)dB(s) - \int_0^t e^{\phi(s)}\frac{1}{2}(\psi(s))^2 ds + \frac{1}{2}\int_0^t e^{\phi(s)}(\psi(s))^2 ds$$
$$= 1 + \int_0^t e^{\phi(s)}\psi(s)dB(s) \quad (150)$$

In particular, for $\psi(t) = 1$, one has the stochastic integral representation for the geometric Brownian motion

$$e^{B(t)} = e^{t/2} + e^{t/2}\int_0^t e^{B(s)}e^{-\frac{1}{2}s}dB(s). \quad (151)$$

	dt	dB(t)
dt	0	0
$dB(t)$	0	dt

Table 1: The rule of thumb

Remark 34 (Rule of thumb) It is usually more convenient to perform computations using differential notation. For instance, in Example 33, one has

$$d\phi(t) = \psi(t)dB(t) - \frac{1}{2}(\psi(t))^2 dt. \tag{152}$$

Now use the Taylor's expansion

$$df(\phi(t)) \approx f'(\phi(t))d\phi(t) + \frac{1}{2}f''(\phi(t))(d\phi(t))^2, \tag{153}$$

and the "multiplication table" It then follows from (153) that

$$d(\phi(t))^2 = (\psi(t)dB(t) - \frac{1}{2}(\psi(t))^2 dt)^2 \tag{154}$$

$$= (\psi(t))^2(dB(t))^2 - (\psi(t))^3 dB(t)dt + \frac{1}{4}(\phi(t))^4(dt)^2 \tag{155}$$

$$= (\psi(t))^2 dt, \tag{156}$$

and that

$$de^{\phi(t)} = e^{\phi(t)}d\phi(t) + \frac{1}{2}e^{\phi(t)}(d\phi(t))^2 \tag{157}$$

$$= e^{\phi(t)}\psi(t)dB(t). \tag{158}$$

Example 35 (Integration by parts (special cases)) *Let* $f(x_1, x_2) = x_1 g(x_2)$ *and*

$$\phi(t) = \phi(0) + \int_0^t \psi_1(s)ds + \int_0^t \psi_2(s)dB(s). \tag{159}$$

where $g : I\!R \mapsto I\!R$ *has continuous first derivative and the increment of the standard Brownian motion is independent of the history of* ψ_1*,* ψ_2*, and the* *standard Brownian motion itself. In this case,* $\frac{\partial f(x_1,x_2)}{\partial x_1} = g(x_2)$, $\frac{\partial f(x_1,x_2)}{\partial x_2} = x_1 g'(x_2)$ *and* $\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2} = 0$. *Then it follows from (130) that*

$$\phi(t)g(t) = \phi(0)g(0) + \int_0^t \phi(s)g'(s)ds + \int_0^t g(s)\psi_1(s)ds$$
$$+ \int_0^t g(s)\psi_2(s)dB(s), \quad 0 \le t < \infty, \ a.s. \tag{160}$$

In differential notation, one has

$$d\phi(t)g(t) = \phi(t)g'(t)dt + g(t)d\phi(t). \tag{161}$$

4 Stochastic differential equations and diffusion equations

4.1 Stochastic differential equations

Let $B(t)$ be the standard Brownian motion as in the previous section. Consider the following equation:

$$d\phi(t) = \mu(\phi(t), t)dt + \sigma(\phi(t), t)dB(t), \tag{162}$$

for some functions $\mu : I\!R^2 \to I\!R$ and $\sigma : I\!R^2 \to I\!R$. The above equation is called a *stochastic differential equation*. As in calculus, one might wonder if there is a solution for the stochastic differential equation and if the solution in unique. For instance, consider the differential equation:

$$d\phi(t) = |\phi(s)|^\alpha ds, \quad \phi(0) = 0. \tag{163}$$

It can be shown that for $0 < \alpha < 1$, all functions of the form

$$\phi(t) = \begin{cases} 0 & \text{if } 0 \le t \le \tau \\ (\frac{t-\tau}{\beta})^\beta & \tau \le t < \infty \end{cases} \tag{164}$$

with $\beta = 1/(1-\alpha)$ and arbitrary $0 \le \tau < \infty$ solve (163). Thus, it is reasonable to develop a theory by assuming Lipschitz-type conditions.

Theorem 36 *(cf. Wong and Hajek [9], Proposition 4.7.1 and Karatzas and Shreve [2], Theorem 5.2.9 and Proposition 5.2.13) Suppose that the* *coefficients $\mu(x,t)$ and $\sigma(x,t)$ satisfy the global Lipschitz and linear growth conditions*

$$|\mu(x,t)-\mu(y,t)|+|\sigma(x,t)-\sigma(y,t)| \le K|x-y|, \tag{165}$$

$$|\mu(x,t)|^2+|\sigma(x,t)|^2 \le K^2(1+x^2), \tag{166}$$

for every $0 \le t < \infty$ and $x, y \in I\!R$, where K is a positive constant. Then there is a unique solution for (162).

The idea of the proof for Theorem 36 is to mimic the deterministic case that construct recursively a sequence of successive approximations by setting $\phi^{(0)}(t) = \eta$ and

$$\phi^{(n+1)}(t) = \eta + \int_0^t \mu(\phi^{(n)}(s), s)ds + \int_0^t \sigma(\phi^{(n)}(s), s)dB(s). \tag{167}$$

The one shows the sequence converge to a solution of (162).

Example 37 (Brownian motion with drift) *Consider the stochastic differential equation*

$$d\phi(t) = -\mu dt + dB(t), \;\; \phi(0) = 0. \tag{168}$$

Then the obvious solution is

$$\phi(t) = \mu t + B(t), \tag{169}$$

which is the Brownian motion with drift coefficient μ.

Example 38 (The Ornstein-Uhlenbeck process) *Consider the stochastic differential equation*

$$d\phi(t) = -\alpha\phi(t)dt + \sigma dB(t), \tag{170}$$

for some $\alpha > 0$ and $\sigma > 0$. Then the solution of this equation is

$$\phi(t) = \phi(0)e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)}dB(s); \;\; 0 \le t < \infty. \tag{171}$$

(Verify this via Itô's calculus using Example 35.) Intuitively, one can view $\phi(t)$ as the output process from the linear time invariant filter with the impulse response $h(t) = e^{-\alpha t}$ (when $\phi(0) = 0$). If $\phi(0)$ is normally distributed,

then it follows from Proposition 29 that $\phi(t)$ is a Gaussian process with $E\phi(t) = e^{-\alpha t}E\phi(0)$ and

$$\begin{aligned}
&E\phi(s)\phi(t) \\
&= E\Big(\phi(0)e^{-\alpha s} + \sigma\int_0^s e^{-\alpha(s-u)}dB(u)\Big)\Big(\phi(0)e^{-\alpha t} + \sigma\int_0^t e^{-\alpha(t-u)}dB(u)\Big) \\
&= E(\phi(0))^2 e^{-\alpha(s+t)} + \sigma^2 E\int_0^s e^{-\alpha(s-u)}dB(u)\int_0^t e^{-\alpha(t-u)}dB(u) \qquad (172) \\
&= [E(\phi(0))^2 + \frac{\sigma^2}{2\alpha}(e^{2\alpha\min[t,s]} - 1)]e^{-\alpha(t+s)}. \qquad (173)
\end{aligned}$$

Example 39 (Brownian bridge) *Consider the following stochastic differential equation*

$$d\phi(t) = -\frac{\phi(t)}{1-t}dt + dB(t), \;\; \phi(0) = 0. \qquad (174)$$

Then the solution of this equation is

$$\phi(t) = \int_0^t \frac{1-t}{1-s}dB(s); \;\; 0 \le t < 1. \qquad (175)$$

(Verify this via Itô's calculus using Example 35.) From Proposition 29, ϕ is a Gaussian process with zero mean and

$$E\phi(t)\phi(s) = \min[s,t] - st. \qquad (176)$$

Thus, ϕ is indeed the Brownian bridge in Definition 7.

Example 40 (Linear equations (see Karatzas and Shreve [2], Section 5.5.6)) *Consider the stochastic differential equation*

$$\begin{aligned}
&d\phi(t) = [A(t)\phi(t) + U(t)]dt + \sigma(t)dB(t), \;\; 0 \le t < \infty, \qquad (177) \\
&\phi(0) = \eta, \qquad (178)
\end{aligned}$$

where $A(t), U(t)$ and $\sigma(t)$ are deterministic functions. Note that the corresponding linear system equation (without noise) is

$$\begin{aligned}
&\eta'(t) = A(t)\eta(t) + U(t), \;\; 0 \le t < \infty, \qquad (179) \\
&\eta(0) = \eta, \qquad (180)
\end{aligned}$$

The solution of (179) is known to be

$$\eta(t) = \Phi(t)[\eta(0) + \int_0^t \frac{U(s)}{\Phi(s)} ds], \tag{181}$$

where $\Phi(t)$ *solves*

$$\Phi'(t) = A(t)\Phi(t), \quad \Phi(0) = 1. \tag{182}$$

Using the Itô rule, it can be shown that

$$\phi(t) = \Phi(t)[\phi(0) + \int_0^t \frac{U(s)}{\Phi(s)} ds + \int_0^t \frac{\sigma(s)}{\Phi(s)} dB(s)]. \tag{183}$$

Again, it follows from Proposition 29 that ϕ *is a Gaussian process with*

$$E\phi(t) = \Phi(t)[E\phi(0) + \int_0^t \frac{U(s)}{\Phi(s)} ds] \tag{184}$$

$$E\Big([\phi(t) - E\phi(t)][\phi(s) - E\phi(s)]\Big) = \Phi(s)[V(0) + \int_0^{\min[s,t]} (\frac{\sigma(u)}{\Phi(u)})^2 du]\Phi(t), \tag{185}$$

where $V(0) = E(\phi(0) - E\phi(0))^2$.

4.2 Diffusion equations

In this section, we derive intuitively the diffusion equations (forward and backward equations) associated with the stochastic differential equation

$$d\phi(t) = \mu(\phi(t), t)dt + \sigma(\phi(t), t)dB(t), \tag{186}$$

where $\mu : I\!R^2 \to I\!R$ and $\sigma : I\!R^2 \to I\!R$ and $B(t)$ is the standard Brownian motion.

In view of (186), the solution (if it exists) is a Markov process (since the small change in the near future, i.e., $d\phi(t)$, only depends on its current "state", $\phi(t)$). Moreover, one can derive intuitively from Proposition 27(ii) and (iii) that

$$E[\phi(t+h) - \phi(t) | \phi(t) = y] = \mu(y, t)h + o(h), \tag{187}$$

and

$$E[(\phi(t+h) - \phi(t))^2 | \phi(t) = y] = \sigma^2(y, t)h + o(h). \tag{188}$$

To see these, use the approximation

$$\phi(t+h) - \phi(t) = \int_t^{t+h} \mu(\phi(s), s)ds + \int_t^{t+h} \sigma(\phi(s), s)dB(s) \quad (189)$$

$$\approx \mu(\phi(t), t)h + \sigma(\phi(t), t)(B(t+h) - B(t)). \quad (190)$$

To simplify the derivation, we further assume that $\mu(x_1, x_2) = \mu(x_1)$ and that $\sigma(x_1, x_2) = \sigma(x_1)$. Under these assumptions, the solution of

$$d\phi(t) = \mu(\phi(t))dt + \sigma(\phi(t))dB(t) \quad (191)$$

is a *time homogeneous* Markov process.

Define the transition (density) function

$$p(x, t, y)dx = Pr(x \le \phi(t) \le x + dx | \phi(0) = y). \quad (192)$$

It then follows from the Chapman-Kolmogorov equation that for all $0 \le s \le t$

$$p(x, t, y) = \int_{-\infty}^{\infty} p(z, s, y)p(x, t - s, z)dz. \quad (193)$$

Note that the random variable $\phi(s)$ (when conditioning on $\phi(0) = y$) has the density $p(z, s, y)$. One can write (193) as

$$p(x, t, y) = E[p(x, t - s, \phi(s)]. \quad (194)$$

To derive the backward equation, we use Taylor's expansion on $p(x, t - s, \phi(s))$, i.e.,

$$p(x, t - s, \phi(s)) = p(x, t, y) - s\frac{\partial}{\partial t}p(x, t, y) \quad (195)$$

$$+(\phi(s) - y)\frac{\partial}{\partial y}p(x, t, y) + \frac{(\phi(s) - y)^2}{2}\frac{\partial^2}{\partial y^2}p(x, t, y) + ... \quad (196)$$

Assuming the expectation and the expansion can be exchanged (in general, this is not true), one has from (187) and (188) that

$$p(x, t, y) = p(x, t, y) - s\frac{\partial}{\partial t}p(x, t, y) \quad (197)$$

$$+\mu(y)s\frac{\partial}{\partial y}p(x, t, y) + \frac{\sigma^2(y)}{2}s\frac{\partial^2}{\partial y^2}p(x, t, y) + o(s). \quad (198)$$

Dividing by s and letting $s \to 0$ yields the *backward diffusion equation*

$$\frac{\partial}{\partial t}p(x, t, y) = \mu(y)\frac{\partial}{\partial y}p(x, t, y) + \frac{\sigma^2(y)}{2}\frac{\partial^2}{\partial y^2}p(x, t, y). \quad (199)$$

To derive the forward equation, we rewrite the Chapman-Kolmogorov equation as follows:

$$p(x, t+s, y) = \int_{-\infty}^{\infty} p(z, t, y) p(x, s, z) dz. \tag{200}$$

Note that

$$\frac{\partial p(x, t+s, y)}{\partial t} = \frac{\partial p(x, t+s, y)}{\partial s}. \tag{201}$$

Thus,

$$\frac{\partial p(x, t+s, y)}{\partial t} = \int_{-\infty}^{\infty} p(z, t, y) \frac{\partial p(x, s, z)}{\partial s} dz. \tag{202}$$

Now use the backward equation in (202) and one has

$$\frac{\partial p(x, t+s, y)}{\partial t} = \int_{-\infty}^{\infty} p(z, t, y) \Big[\mu(z) \frac{\partial}{\partial z} p(x, s, z) + \frac{\sigma^2(z)}{2} \frac{\partial^2}{\partial z^2} p(x, s, z)\Big] dz. \tag{203}$$

Apply integration by parts (once for the first term and twice for the second term) and assume all the boundary terms vanishes. We have

$$\frac{\partial p(x, t+s, y)}{\partial t} = \int_{-\infty}^{\infty} \Big[- \frac{\partial}{\partial z}(p(z, t, y)\mu(z)) + \frac{\partial^2}{\partial z^2}(p(z, t, y)\frac{\sigma^2(z)}{2})\Big] p(x, s, z) dz. \tag{204}$$

Note $p(x, s, z)$ becomes the delta function $\delta(x - z)$ as $s \to 0$. Thus, letting $s \to 0$ yields the *forward diffusion equation*

$$\frac{\partial p(x, t, y)}{\partial t} = -\frac{\partial}{\partial x}(p(x, t, y)\mu(x)) + \frac{\partial^2}{\partial x^2}(p(x, t, y)\frac{\sigma^2(x)}{2}). \tag{205}$$

Analogous to the fundamental limit theorem for ergodic Markov chains, one should also expect that

$$\lim_{t \to \infty} p(x, t, y) = \pi(x), \tag{206}$$

where $\pi(x)$ is the *stationary* distribution of the Markov process associated with the stochastic differential equation in (191). Note that the stationary distribution should satisfy

$$\pi(x) = \int_{-\infty}^{\infty} \pi(y) p(x, t, y) dy. \tag{207}$$

Mimicking the derivation for the Markov chains, we can let

$$\lim_{t \to \infty} \frac{\partial p(x, t, y)}{\partial t} = 0. \tag{208}$$

In conjunction with (206) and the forward diffusion equation in (205), one has (when $t \to \infty$)

$$0 = -\frac{\partial}{\partial x}(\pi(x)\mu(x)) + \frac{\partial^2}{\partial x^2}(\pi(x)\frac{\sigma^2(x)}{2}). \tag{209}$$

Integrating (209) yields

$$\frac{c_1}{2} = -\pi(x)\mu(x) + \frac{d}{dx}(\pi(x)\frac{\sigma^2(x)}{2}), \tag{210}$$

where c_1 is a constant. Multiplying by the integrating factor

$$s(x) = \exp\Big(- \int^x \frac{2\mu(z)}{\sigma^2(z)} dz\Big), \tag{211}$$

we can write (210) in

$$\frac{d(s(x)\sigma^2(x)\pi(x))}{dx} = c_1 s(x). \tag{212}$$

Another integration yields

$$\pi(x) = c_1 \frac{S(x)}{s(x)\sigma^2(x)} + c_2 \frac{1}{s(x)\sigma^2(x)}, \tag{213}$$

where

$$S(x) = \int^x s(z) dz. \tag{214}$$

The (normalization) constants c_1 and c_2 are determined to guarantee the constraints $\pi(x) \geq 0$ and $\int_{-\infty}^{\infty} \pi(x) = 1$.

Example 41 (The Ornstein-Uhlenbeck process) *Consider the O-U process in Example 38. In this example, $\mu(x) = -\alpha x$ and $\sigma(x) = \sigma$. Thus, $s(x) = exp(\gamma x^2)$ with $\gamma = \alpha/\sigma^2$ and*

$$\pi(x) = c_1 (\int^x e^{\gamma z^2} dz) e^{-\gamma x^2} + c_2 e^{-\gamma x^2}. \tag{215}$$

To insure that $\int_{-\infty}^{\infty} \pi(x) = 1$, one must have $\lim_{x \to \infty} \pi(x) = 0$. Thus, $c_1 = 0$ and the stationary distribution of the O-U process is the normal distribution, $\pi(x) = ce^{-\gamma x^2}$. This agrees with the direct computation from (172) (letting $s = t \to \infty$).

Example 42 *Consider the following stochastic differential equation*

$$d\phi(t) = -\phi(t)dt + \sqrt{2(1+\phi^2(t))}dB(t). \tag{216}$$

Then $\mu(x) = -x$ *and* $\sigma(x) = \sqrt{2(1+x^2)}$. *Thus,* $s(x) = \sqrt{1+x^2}$. *To ensure that* $\int_{-\infty}^{\infty} \pi(x) = 1$, $c_1 = 0$ *in (213) (Why? If* $c_1 \neq 0$, *then* $\pi(x) \approx 1/x$ *for large* x*). Thus,*

$$\pi(x) = c_2 \frac{1}{(1+x^2)^{3/2}}. \tag{217}$$

From $\int_{-\infty}^{\infty} \pi(x) = 1$, *one has* $c_2 = 1/2$

For further reading on diffusion, consult Wong and Hajek [9], Chapter 5, Karatzas and Shreve [2] Chapter 5, and Karlin and Taylor [3], Chapter 15.

5 Filtering and Control

5.1 Kalman-Bucy filter

Consider the linear system

$$\eta'(t) = -\alpha\eta(t), \quad 0 \leq t < \infty, \tag{218}$$

$$\eta(0) = \eta, \tag{219}$$

where η is normally distributed with mean μ_0 and variance σ_0^2. Suppose that we are not able to observe the system directly and that our observation is perturbed by a noise (modelled by the standard Brownian motion) as follows:

$$\phi(t) = \int_0^t \eta(s)ds + B(t), \tag{220}$$

where $\{\eta(s), 0 \leq s \leq t\}$ and $\{B(s), 0 \leq s \leq t\}$ are independent. The question is how we obtain a good (or optimal) estimate for $\eta(t)$ based on the observation process $\{\phi(s), 0 \leq s \leq t\}$.

To formalize the problem, we consider the family of linear filters in Proposition 29. An estimate of $\eta(t)$ can be obtained by passing the observation process through a linear time varying filter, i.e.,

$$\hat{\eta}(t) = \int_0^t h(t,s)d\phi(s). \tag{221}$$

The mathematical question is then to find the best filter $h(t,s)$ that minimizes the mean square error (MSE), i.e.,

$$\min E(\eta(t) - \hat{\eta}(t))^2. \tag{222}$$

Proposition 43 *(Wiener-Hopf equation) A necessary and sufficient condition for the optimal filter is that*

$$h(t,s) = E[\eta(t)\eta(s)] - \int_0^t h(t,u)E[\eta(s)\eta(u)]du. \tag{223}$$

Proof. Consider the variation problem.

$$J(\epsilon) = E[\eta(t) - \int_0^t (h(t,u) + \epsilon\delta(t,u))d\phi(u)]^2 \tag{224}$$

$$= E\Big[\eta(t)^2 - 2\eta(t)\int_0^t (h(t,u) + \epsilon\delta(t,u))d\phi(u) \tag{225}$$

$$+ \int_0^t \int_0^t (h(t,u_1) + \epsilon\delta(t,u_1))(h(t,u_2) + \epsilon\delta(t,u_2))d\phi(u_1)d\phi(u_2)\Big] \tag{226}$$

$$= A\epsilon^2 + B\epsilon + C. \tag{227}$$

A necessary condition for the optimal solution is

$$\frac{dJ(\epsilon)}{\epsilon}|_{\epsilon=0} = B = 0. \tag{228}$$

Thus,

$$B = -2E[\eta(t)\int_0^t \delta(t,u)d\phi(u)] + 2E[\int_0^t \int_0^t h(t,u_1)\delta(t,u_2)d\phi(u_1)d\phi(u_2)] = 0. \tag{229}$$

Note from (220) that $d\phi(t) = \eta(t)dt + dB(t)$. Since $\{\eta(s), 0 \le s \le t\}$ and $\{B(s), 0 \le s \le t\}$ are independent,

$$E[\eta(t)\int_0^t \delta(t,u)d\phi(u)] \tag{230}$$

$$= E[\eta(t)\int_0^t \delta(t,u)\eta(u)d(u)] + E[\eta(t)\int_0^t \delta(t,u)dB(u)] \tag{231}$$

$$= \int_0^t \delta(t,u)E[\eta(t)\eta(u)]d(u) + E[\eta(t)]E[\int_0^t \delta(t,u)dB(u)] \tag{232}$$

$$= \int_0^t \delta(t,u)E[\eta(t)\eta(u)]d(u)], \tag{233}$$

where $E[\int_0^t \delta(t,u)dB(u)] = 0$ follows from Proposition 27(ii). Similarly,

$$E[\int_0^t \int_0^t h(t,u_1)\delta(t,u_2)d\phi(u_1)d\phi(u_2)] \tag{234}$$

$$= E[\int_0^t \int_0^t h(t,u_1)\delta(t,u_2)\eta(u_1)du_1\eta(u_2)du_2] + E[\int_0^t \int_0^t h(t,u_1)\delta(t,u_2)\eta(u_1)du_1 dB(u_2)]$$

$$+E[\int_0^t \int_0^t h(t,u_1)\delta(t,u_2)dB(u_1)\eta(u_2)du_2] + E[\int_0^t \int_0^t h(t,u_1)\delta(t,u_2)dB(u_1)dB(u_2)] \tag{235}$$

$$= \int_0^t \int_0^t h(t,u_1)\delta(t,u_2)E[\eta(u_1)\eta(u_2)]du_1du_2] + \int_0^t h(t,u)\delta(t,u)du, \tag{236}$$

where we apply Corollary 28 in the last identity. In conjunction with (229), one has

$$\int_0^t \delta(t,s)\Big[E[\eta(t)\eta(s)] - \int_0^t h(t,u)E[\eta(s)\eta(u)]du - h(t,s)\Big]ds = 0. \tag{237}$$

Since (237) holds for arbitrary $\delta(t,s)$, it must

$$E[\eta(t)\eta(s)] - \int_0^t h(t,u)E[\eta(s)\eta(u)]du = h(t,s). \tag{238}$$

To prove the sufficient part, one needs to verify that

$$E[\eta(t) - \int_0^t h_1(t,s)d\phi(s)]^2 \geq E[\eta(t) - \int_0^t h(t,s)d\phi(s)]^2, \tag{239}$$

when $h(t,s)$ satisfies the Wiener-Hopf equation. Write

$$E[\eta(t) - \int_0^t h_1(t,s)d\phi(s)]^2 = E[\eta(t) - \int_0^t h(t,s)d\phi(s) + \int_0^t h(t,s)d\phi(s) - \int_0^t h_1(t,s)d\phi(s)]^2$$

$$= E[\eta(t) - \int_0^t h(t,s)d\phi(s)]^2 + E[\int_0^t h(t,s)d\phi(s) - \int_0^t h_1(t,s)d\phi(s)]^2$$

$$+2E[\eta(t) - \int_0^t h(t,s)d\phi(s)][\int_0^t h(t,s)d\phi(s) - \int_0^t h_1(t,s)d\phi(s)]. \tag{240}$$

It is obvious that

$$E[\int_0^t h(t,s)d\phi(s) - \int_0^t h_1(t,s)d\phi(s)]^2 \tag{241}$$

$$= E[\int_0^t (h(t,s) - h_1(t,s))d\phi(s)]^2 \geq 0. \tag{242}$$

Also, using the Wiener-Hopf equation, one can show the last term in (240) is 0.

□

Proposition 44 *If $h(t,s)$ is the optimal filter, then*

$$h(t,t) = E(e(t))^2, \tag{243}$$

where $e(t) = \eta(t) - \hat{\eta}(t)$ is the error of the estimate $\hat{\eta}(t) = \int_0^t h(t,s)d\phi(s)$.

Proof. Note that

$$E(e(t))^2 = E(\eta(t))^2 - 2E[\eta(t)\int_0^t h(t,u)d\phi(u)] \tag{244}$$

$$+ \int_0^t \int_0^t h(t,u)h(t,s)d\phi(u)d\phi(s). \tag{245}$$

Recall that $d\phi(t) = \eta(t)dt + dB(t)$. Analogous to the computations in the proof of Proposition 43, one has

$$E[\eta(t)\int_0^t h(t,u)d\phi(u)] = \int_0^t h(t,u)E[\eta(t)\eta(u)]du, \tag{246}$$

and

$$E\int_0^t \int_0^t h(t,u)h(t,s)d\phi(u)d\phi(s)$$
$$= \int_0^t \int_0^t h(t,u)h(t,s)E[\eta(u)\eta(s)]duds + \int_0^t (h(t,u))^2 du. \tag{247}$$

Thus, applying the Wiener-Hopf equation in Proposition 43 yields

$$E(e(t))^2 = E(\eta(t))^2 - \int_0^t h(t,u)E[\eta(t)\eta(u)]du$$
$$- \int_0^t h(t,u)\Big(E[\eta(t)\eta(u)] - \int_0^t h(t,s)E[\eta(u)\eta(s)]ds - h(t,u)\Big)du \tag{248}$$
$$= h(t,t). \tag{249}$$

□

Note that both Proposition 43 and Proposition 44 are general results for the observation process in (220). We have not used the knowledge of the linear system in (218). With the knowledge of the linear system, the Wiener-Hopf equation can be further simplified as in the following proposition.

Proposition 45 *A sufficient condition for the Wiener-Hopf equation in Proposition 43 is*

$$\frac{\partial h(t,s)}{\partial t} = (-\alpha - h(t,t))h(t,s). \tag{250}$$

Proof. Taking the partial derivatives on both sides of the Wiener-Hopf equation yields

$$\frac{\partial h(t,s)}{\partial t} = \frac{\partial E[\eta(t)\eta(s)]}{\partial t} - \frac{\partial}{\partial t}(\int_0^t h(t,u)E[\eta(s)\eta(u)]du). \tag{251}$$

Recall that $\eta'(t) = -\alpha\eta(t)$. Thus,

$$\frac{\partial E[\eta(t)\eta(s)]}{\partial t} = E[\eta'(t)\eta(s)] = -\alpha E[\eta(t)\eta(s)], \tag{252}$$

where the interchange of the derivative and the expectation can be justified by the Lipschitz continuous condition (in general, if the derivative exists, then it is Lipschitz continuous). By the chain rule

$$\frac{\partial}{\partial t}(\int_0^t h(t,u)E[\eta(s)\eta(u)]du) \tag{253}$$

$$= h(t,t)E[\eta(s)\eta(t)] + \int_0^t \frac{\partial h(t,u)}{\partial t}E[\eta(s)\eta(u)]du. \tag{254}$$

$$\tag{255}$$

In conjunction with (251) and (252), one has

$$\frac{\partial h(t,s)}{\partial t} = (-\alpha - h(t,t))E[\eta(s)\eta(t)] - \int_0^t \frac{\partial h(t,u)}{\partial t}E[\eta(s)\eta(u)]du. \tag{256}$$

Applying the Wiener-Hopf equation in (256) yields

$$0 = -\frac{\partial h(t,s)}{\partial t} + (-\alpha - h(t,t))(h(t,s) + \int_0^t h(t,u)E[\eta(s)\eta(u)]du) - \int_0^t \frac{\partial h(t,u)}{\partial t}E[\eta(s)\eta(u)]du \tag{257}$$

It is clear that (250) is a sufficient condition of (257). □

Proposition 46 *If $\hat{\eta}(t)$ is the optimal estimate, then it satisfies the recursive equation*

$$d\hat{\eta}(t) = -\alpha\hat{\eta}(t)dt + h(t,t)d\nu(t), \tag{258}$$

where

$$d\nu(t) = d\phi(t) - \hat{\eta}(t)dt = (\eta(t) - \hat{\eta}(t))dt + dB(t). \tag{259}$$

Proof. Recall that

$$\hat{\eta}(t) = \int_0^t h(t,s)d\phi(s). \tag{260}$$

By the chain rule,

$$d\hat{\eta}(t) = h(t,t)d\phi(t) + \int_0^t \frac{\partial h(t,u)}{\partial t} dt d\phi(u). \tag{261}$$

It then follows from (250) that

$$d\hat{\eta}(t) = h(t,t)d\phi(t) + \int_0^t (-\alpha - h(t,t))h(t,u)d\phi(u)dt \tag{262}$$

$$= h(t,t)d\phi(t) + (-\alpha - h(t,t)) \int_0^t h(t,u)d\phi(u)dt \tag{263}$$

$$= h(t,t)d\phi(t) + (-\alpha - h(t,t))\hat{\eta}(t)dt \tag{264}$$

$$= -\alpha\hat{\eta}(t)dt + h(t,t)d\nu(t). \tag{265}$$

□

Now the only unknown is $h(t,t)$.

Proposition 47 *Let $k(t) = h(t,t)$. Then it satisfies the Riccati equation*

$$\frac{dk(t)}{dt} = -2\alpha k(t) - k^2(t). \tag{266}$$

Note that the equation can be solved offline .

Proof. Note from (258) and (259) that

$$de(t) = d\eta(t) - d\hat{\eta}(t) \tag{267}$$

$$= -\alpha\eta(t)dt + \alpha\hat{\eta}(t)dt - k(t)((\eta(t) - \hat{\eta}(t))dt + dB(t)) \tag{268}$$

$$= (-\alpha - k(t))e(t)dt - k(t)dB(t). \tag{269}$$

By Itô's rule (see Remark 31) and (129),

$$d(e(t))^2 = 2e(t)de(t) + \frac{1}{2}2(k(t))^2 dt \tag{270}$$

$$= 2e(t)[(-\alpha - k(t))e(t)dt - k(t)dB(t)] + k^2(t)dt. \tag{271}$$

Thus,

$$(e(t))^2 = 2\int_0^t (e(s))^2(-\alpha - k(s))ds - 2\int_0^t e(s)k(s)dB(s) + \int_0^t k^2(s)ds. \tag{272}$$

Taking expectation on both sides yields

$$E(e(t))^2 = 2\int_0^t E(e(s))^2(-\alpha - k(s))ds + \int_0^t k^2(s)ds, \tag{273}$$

where the stochastic integral term vanishes. It then follows from Proposition 44 that

$$k(t) = 2\int_0^t k(s)(-\alpha - k(s))ds + \int_0^t k^2(s)ds \tag{274}$$

$$= -2\alpha\int_0^t k(s)ds - \int_0^t k^2(s)ds \tag{275}$$

□

Theorem 48 *(Kalman-Bucy filter) The optimal estimate that minimizes the MSE can be generated by the recursive equation*

$$d\hat{\eta}(t) = -\alpha\hat{\eta}(t)dt + k(t)(d\phi(t) - \hat{\eta}(t)dt), \tag{276}$$

where $k(t)$ is obtained by the Riccati equation

$$\frac{dk(t)}{dt} = -2\alpha k(t) - k^2(t). \tag{277}$$

Though we only show that the Kalman-Bucy filter is optimal among all the time varying linear filters. This result can be extended to all nonlinear filters. This is due to the fact that the optimal filter for a *Gaussian* process is linear.

For further reading on filtering, we refer to Wong and Hajek [9], Chapter 7.

5.2 Stochastic control

In this section, we discuss Linear Quadratic Gaussian (LQG) regulators and the associated control problems.

Consider the linear system in (218). Suppose that we add control $\{u(t), t \geq 0\}$ to the system with the constant gain b, i.e.,

$$\eta'(t) = -\alpha\eta(t) + bu(t), \quad 0 \leq t < \infty, \tag{278}$$
$$\eta(0) = \eta_0, \tag{279}$$

The objective of adding control is to minimize the following cost function

$$\min \int_0^t (g\eta^2(s) + hu^2(s))ds \tag{280}$$

for some constants g and h. In view of the quadratic form of the objective function, the control problem is called a quadratic regulator. To solve the control problem, we use an approach in dynamic programming. Let $V(s, x)$ be the optimal value of the cost function for the problem that starts from time s and $\eta(s) = x$. Thus,

$$V(0, \eta_0) = \min \int_0^t (g\eta^2(s) + hu^2(s))ds. \tag{281}$$

Also, one has the boundary condition

$$V(t, x) = 0, \tag{282}$$

for all states x.

Analogous to the derivation of the backward equation, one has

$$V(s - ds, x) = \min_{\{u(\tau), s-ds \leq \tau \leq s\}} \Big[\int_{s-ds}^{s} (g\eta^2(\tau) + hu^2(\tau))d\tau + V(s, \eta(s)) \Big]. \tag{283}$$

Applying Taylor's expansion yields

$$V(s, \eta(s)) = V(s, x) + \frac{\partial V(s, x)}{\partial x} d\eta(s). \tag{284}$$

In conjunction with (283), one has

$$0 = \min_{\{u(\tau), s-ds \leq \tau \leq s\}} \Big[\frac{1}{ds} \int_{s-ds}^{s} (g\eta^2(\tau) + hu^2(\tau))d\tau + \frac{V(s, x) - V(s - ds, x)}{ds} + \frac{\partial V(s, x)}{\partial x} \frac{d\eta(s)}{ds} \Big]. \tag{285}$$

Letting $ds \to 0$,

$$\frac{d\eta(s)}{ds} \longrightarrow -\alpha x + bu(s) \tag{286}$$
$$\frac{1}{ds} \int_{s-ds}^{s} g\eta^2(\tau) + hu^2(\tau)d\tau \longrightarrow gx^2 + hu^2(s) \tag{287}$$
$$\frac{V(s, x) - V(s - ds, x)}{ds} \longrightarrow \frac{\partial V(s, x)}{\partial s}. \tag{288}$$

Thus,

$$0 = \min_u [gx^2 + hu^2 + \frac{\partial V(s,x)}{\partial s} + \frac{\partial V(s,x)}{\partial x}(-\alpha x + bu)]. \tag{289}$$

The above equation is known as the Hamilton-Jacobi-Bellman (HJB) equation.

Proposition 49 *(Verification Theorem) If there exists $V^*(t,x)$ that satisfies the HJB equation in (289), then*

$$V(s,x) \geq V^*(s,x), 0 \leq s \leq t. \tag{290}$$

Proof. Consider the problem that starts from time s and $\eta(s) = x$. For a particular control $\{u(\tau), s \leq \tau \leq t\}$, there is an associate path $\{\eta(\tau), s \leq \tau \leq t\}$. Integrating along the path yields

$$0 = V^*(t,\eta(t)) = V^*(s,\eta(s)) + \int_s^t \frac{\partial V^*(\tau,\eta(\tau))}{\partial \tau} d\tau + \int_s^t \frac{\partial V^*(\tau,\eta(\tau))}{\partial x} d\eta(\tau). \tag{291}$$

Using (278), one has

$$0 = V^*(s,\eta(s)) + \int_s^t \Big[\frac{\partial V^*(\tau,\eta(\tau))}{\partial \tau} + \frac{\partial V^*(\tau,\eta(\tau))}{\partial x}(-\alpha\eta(\tau) + bu(\tau))\Big] d\tau. \tag{292}$$

Since we assume V^* satisfies the HJB equation, for any control $u(\cdot)$

$$0 \leq gx^2 + hu^2 + \frac{\partial V^*(s,x)}{\partial s} + \frac{\partial V^*(s,x)}{\partial x}(-\alpha x + bu). \tag{293}$$

In conjunction with (292), one has

$$\int_s^t (g\eta^2(\tau) + hu^2(\tau)) d\tau \geq V^*(s,x). \tag{294}$$

□

In view of Proposition 49, one could solve the quadratic regulator problem by guessing the right solution. Guess

$$V(s,x) = k(s)x^2 \tag{295}$$

for some function $k(s)$. Note that $\frac{\partial V(s,x)}{\partial s} = k'(s)x^2$ and $\frac{\partial V(s,x)}{\partial x} = 2k(s)x$. Thus, the HJB equation in (289) is

$$0 = \min_u [gx^2 + hu^2 + k'(s)x^2 + 2xk(s)(-\alpha x + bu)]. \tag{296}$$

The optimization problem can then be solved by

$$u = -\frac{bk(s)x}{h}. \tag{297}$$

Replacing (297) in (296) yields

$$k'(s) = -g + 2\alpha k(s) + \frac{b^2}{h}k^2(s), \tag{298}$$

with the boundary condition $k(t) = 0$ (since $V(t,x) = 0$).

Now we extend the control problem to the stochastic setting. Assume there is perturbation in the linear system in (278) and the system is governed by the following equation

$$d\eta(t) = (-\alpha\eta(t) + bu(t))dt + cdB(t), \quad 0 \leq t < \infty, \tag{299}$$
$$\eta(0) = \eta_0, \tag{300}$$

where $B(t)$ is the standard Brownian motion. The objective of this control problem is to minimize the *expected* cost as follows:

$$\min E \int_0^t (g\eta^2(s) + hu^2(s))ds \tag{301}$$

for some constants g and h.

Analogous to the deterministic setting, let $V(s,x)$ be the optimal value of the cost function for the problem that starts from time s and $\eta(s) = x$. Similarly, one has the boundary condition

$$V(t,x) = 0, \tag{302}$$

for all states x and the "backward" equation

$$V(s-ds,x) = \min_{\{u(\tau), s-ds \leq \tau \leq s\}} E\Big[\int_{s-ds}^{s} (g\eta^2(\tau) + hu^2(\tau))d\tau + V(s,\eta(s))\Big]. \tag{303}$$

Applying Ito's rule (Taylor's expansion to the second order term) yields

$$V(s,\eta(s)) = V(s,x) + \frac{\partial V(s,x)}{\partial x}d\eta(s) + \frac{1}{2}\frac{\partial^2 V(s,x)}{\partial x^2}(d\eta(s))^2. \tag{304}$$

When ds is small,

$$d\eta(s) \longrightarrow (-\alpha x + bu(s))ds + cdB(s) \tag{305}$$
$$(d\eta(s))^2 \longrightarrow c^2 ds \tag{306}$$
$$\int_{s-ds}^{s} g\eta^2(\tau) + hu^2(\tau)d\tau \longrightarrow (gx^2 + hu^2(s))ds \tag{307}$$
$$V(s,x) - V(s-ds,x) \longrightarrow \frac{\partial V(s,x)}{\partial s}ds. \tag{308}$$

Note that $E[cdB(s)] = 0$ and one has the following HJB equation

$$0 = \min_u [gx^2 + hu^2 + \frac{\partial V(s,x)}{\partial s} + \frac{\partial V(s,x)}{\partial x}(-\alpha x + bu) + \frac{\partial^2 V(s,x)}{\partial x^2}\frac{c^2}{2}]. \tag{309}$$

Proposition 50 *(Verification Theorem) If there exists $V^*(t,x)$ that satisfies the HJB equation in (309), then*

$$V(s,x) \geq V^*(s,x), 0 \leq s \leq t. \tag{310}$$

Guess the solution

$$V(s,x) = k_0(s) + k_2(s)x^2 \tag{311}$$

for some function $k_0(s)$ and $k_2(s)$. One can show that

$$u = -\frac{bk_2(s)x}{h} \tag{312}$$
$$k_2'(s) = -g + 2\alpha k_2(s) + \frac{b^2}{h}k_2^2(s) \tag{313}$$
$$k_0'(s) = -c^2 k_2(s) \tag{314}$$

with the boundary conditions $k_2(t) = k_0(t) = 0$.

The above problem is with complete observation. Let us now consider the problem with incomplete observation as in (218) and (220) (the observation is perturbed by a noise). Assume the system is governed by the following equations:

$$d\eta(t) = (-\alpha\eta(t) + bu(t))dt, \quad 0 \leq t < \infty, \tag{315}$$
$$\eta(0) = \eta_0, \tag{316}$$
$$\phi(t) = \int_0^t \eta(s)ds + B(t), \tag{317}$$

where η_0 is normally distributed and $B(t)$ is the standard Brownian motion. In this system, ϕ is the process we can observe. The objective of this control problem is once again to minimize the *expected* cost as follows:

$$\min E \int_0^t (g\eta^2(s) + hu^2(s))ds \tag{318}$$

for some constant g and h.

The solution of this problem is based on the *separation principle*: the estimation part and the control part can be optimized separately. In view of the Kalman-Bucy filter in Theorem 48, the estimation part can be carried out by the recursive equation (though we have the extra control term)

$$d\hat{\eta}(t) = (-\alpha\hat{\eta}(t) + bu(t))dt + k(t)d\nu(t), \tag{319}$$

where $d\nu(t) = d\phi(t) - \hat{\eta}(t)dt$ is the innovation process and $k(t)$ is the Kalman gain obtained by the Riccati equation

$$\frac{dk(t)}{dt} = -2\alpha k(t) - k^2(t). \tag{320}$$

It can be shown that $\nu(t)$ is in fact the standard Brownian motion. In view of (309), one has the corresponding HJB equation for (319)

$$0 = \min_u [gx^2 + hu^2 + \frac{\partial V(s,x)}{\partial s} + \frac{\partial V(s,x)}{\partial x}(-\alpha x + bu) + \frac{\partial^2 V(s,x)}{\partial x^2}\frac{k^2(s)}{2}], \tag{321}$$

where $V(s,x)$ is the best performance starting at $\hat{\eta}(s) = x$. The problem can then be solved by guessing the solution

$$V(t,x) = v_0(t) + v_2(t)x^2, \tag{322}$$

where

$$v_2'(s) = -g + 2\alpha v_2(s) + \frac{b^2}{h}v_2^2(s) \tag{323}$$

$$v_0'(s) = -k^2(s)v_2(s) \tag{324}$$

with the boundary conditions $v_2(t) = v_0(t) = 0$.

For further reading on stochastic control, see e.g. M.H.A. Davis and R.B. Vinter [1] and V. Krishnan [4].

Homework

1. Let $\phi(t) = (b - \frac{1}{2}\sigma^2)t + \sigma B(t)$. Find $e^{\phi(t)}$.

2. Solve the stochastic differential equation $d\psi(t) = b\psi(t)dt + \sigma\psi(t)dB(t)$.

3. (Fractional Brownian motion [5]) Let $X(t) = \int_0^t (t-s)^{H-1/2} dB(s)$, where $0 < H < 1$ and $B(t)$ is the standard Brownian motion. Show that $\frac{X(nt)}{n^H}$ has the same distribution as $X(t)$. Find the covariance process $EX(s)X(t)$. Note that $X(t)$ does not have independent increment. When $t \to \infty$, its increment is stationary and is called the fractional Brownian motion.

4. (Orthogonality principle) Use the Wiener-Hopf equation in (223) to verify the last term in (240) is 0, i.e.,

$$E[\eta(t) - \int_0^t h(t,s)d\phi(s)][\int_0^t h(t,s)d\phi(s) - \int_0^t h_1(t,s)d\phi(s)] = 0 \quad (325)$$

This shows that the error term $e(t)$ and the observation process $\{\phi(s), 0 \leq s \leq t\}$ are uncorrelated. Since the process is Gaussian, they are in fact independent.

5. Show Proposition 50 and derive (312)-(314).

5.3 Applications to economics: optimal consumption and investment

Let us consider a market in which two assets (or securities) are traded continuously. The first asset, called the *bond*, has a price $P_0(t)$ which evolves according to the differential equation

$$dP_0(t) = rP_0(t)dt, \quad P_0(0) = p_0. \quad (326)$$

The constant r is simply the *interest rate*. The second asset, call the *stock*, is risky and has a price $P_1(t)$ that is modelled by the linear stochastic differential equation

$$dP_1(t) = bP_1(t)dt + \sigma P_1(t)dB(t), \quad P_1(0) = p_1. \quad (327)$$

The constants b and σ are called the *mean rate of return* and the *dispersion* coefficient, respectively.

Suppose an investor who starts with some initial endowment $x \geq 0$ and invest it in the two assets. Let $N_i(t)$, $i = 0$ and 1, be the number of shares of asset i owned by the investor at time t, and $X(t)$ be the investor's wealth at time t. Then $X(0) = x = N_0(0)p_0 + N_1(0)p_1$ and

$$X(t) = N_0(t)P_0(t) + N_1(t)P_1(t). \tag{328}$$

If trading of shares happens at discrete time, and there is no infusion or withdrawal of funds, then

$$X(t+h)-X(t) = N_0(t)(P_0(t+h)-P_0(t))+N_1(t)(P_1(t+h)-P_1(t)). \tag{329}$$

If, furthermore, the investor chooses at time $t+h$ to consume an amount $hC(t+h)$, then one should modify (329) as

$$X(t+h)-X(t) = N_0(t)(P_0(t+h)-P_0(t))+N_1(t)(P_1(t+h)-P_1(t))-hC(t+h). \tag{330}$$

For the continuous time model, one has from (330) that

$$dX(t) = N_0(t)dP_0(t) + N_1(t)dP_1(t) - C(t)dt. \tag{331}$$

Let $\pi(t) = N_1(t)P_1(t)$ be the amount invested in the stock at time t. Using (326) and (327), one has

$$dX(t) = (rX(t) - C(t) + (b-r)\pi(t))dt + \sigma\pi(t)dB(t). \tag{332}$$

In view of (327), we have two controls: the *portfolio* process $\pi(\cdot)$ and the *consumption* process $C(\cdot)$. The portfolio process decides how the investor invests his wealth in the stock and the consumption process decides how he consumes his wealth. (The rest is of course in the bond.) The objective of the control problem is to *maximum* the following value function

$$\max E \int_0^t (C(s))^\delta ds, \quad 0 < \delta < 1. \tag{333}$$

Note that for $0 < \delta < 1$, the function $g(c) = c^\delta$ is an increasing concave function. The function $g(c)$, usually called the *utility* function, measures how happy the investor is when he consumes at the rate c. The objective is then equivalent to maximize the expected "happiness" over the period $[0,t]$. The reason that $g(c)$ should be increasing is obvious. The concavity of $g(c)$ implies that the gain of happiness is small when he is already happy.

To solve this control problem, let $V(s,x)$ be the optimal value of the value function when the investor starts at time s with the initial endowment x. Analogous to the derivation of the HJB equation in (309), one has

$$0 = \max_{\pi,c}[c^\delta + \frac{\partial V(s,x)}{\partial s} + \frac{\partial V(s,x)}{\partial x}(rx-c+(b-r)\pi) + \frac{\partial^2 V(s,x)}{\partial x^2}\frac{\sigma^2\pi^2}{2}]. \tag{334}$$

Now guess the solution of the HJB equation in (334) with

$$V(s,x) = (k(s))^{1-\delta} x^{\delta}, \tag{335}$$

for some function $k(s)$. Note that

$$\frac{\partial V(s,x)}{\partial s} = (1-\delta)(k(s))^{-\delta} k'(s) x^{\delta} \tag{336}$$

$$\frac{\partial V(s,x)}{\partial x} = (k(s))^{1-\delta} \delta x^{\delta-1} \tag{337}$$

$$\frac{\partial^2 V(s,x)}{\partial x^2} = (k(s))^{1-\delta} \delta(\delta-1) x^{\delta-2}. \tag{338}$$

Replacing them in (334) yields

$$0 = \max_{\pi,c}[c^{\delta} + (1-\delta)(k(s))^{-\delta} k'(s) x^{\delta}$$

$$+(k(s))^{1-\delta}\delta x^{\delta-1}(rx - c + (b-r)\pi) + (k(s))^{1-\delta}\delta(\delta-1)x^{\delta-2}\frac{\sigma^2\pi^2}{2}]. \tag{339}$$

The optimization problem can be solved by

$$\pi = \frac{(b-r)x}{\sigma^2(1-\delta)} \tag{340}$$

$$c = \frac{x}{k(s)}. \tag{341}$$

Replacing π and c in (339) yields

$$k'(s) = \alpha k(s) - 1, \tag{342}$$

where

$$\alpha = -\frac{1}{1-\delta}(r\delta + \frac{1}{2}\frac{(b-r)^2}{\sigma^2}\frac{\delta}{1-\delta}). \tag{343}$$

Since the boundary condition $V(t,x) = 0$ for all x implies $k(t) = 0$, one has

$$k(s) = \frac{1 - e^{-\alpha(t-s)}}{\alpha}. \tag{344}$$

The results in (340) and (341) show both the optimal portfolio process $\pi(s)$ and the consumption process $C(s)$ are proportional to the wealth process $X(t)$.

Now substituting the optimal controls $\pi(\cdot)$ in (340) and $C(\cdot)$ in (341) into (332) yields

$$dX(s) = X(s)(r - \frac{1}{k(s)} + \frac{(b-r)^2}{\sigma^2(1-\delta)})ds + X(s)\frac{(b-r)}{\sigma(1-\delta)}dB(s). \tag{345}$$

Solve the stochastic differential equation (cf. Homework 2) and one has

$$X(s) = x\exp\Big(\int_0^s r - \frac{1}{k(u)} + \frac{(b-r)^2}{\sigma^2(1-\delta)} - \frac{(b-r)^2}{2\sigma^2(1-\delta)}du + \int_0^s \frac{(b-r)}{\sigma(1-\delta)}dB(u)\Big). \tag{346}$$

Observe that

$$\int_0^t \frac{1}{k(u)}du = \int_0^t \frac{\alpha}{1-e^{-\alpha(t-u)}}du = \int_0^t \frac{\alpha}{1-e^{-\alpha u}}du \tag{347}$$

$$\geq \int_0^t \frac{\alpha}{\alpha u}du = \infty. \tag{348}$$

Thus, we have from (346) that (with probability one) $X(t) = 0$. This result is expected since there is no reason to have any wealth left at the end of the optimization period (the investor should consume all in order to maximize his happiness over $[0, t]$).

For further reading on applications to economics, see e.g. Karatzas and Shreve [2], Section 5.8.

6 Extensions and generalizations

To extend from the one dimensional Brownian motion to the multidimensional Brownian motion, it is better to work on a common "history", $\{\mathcal{F}_t, t \geq 0\}$. The history is generated by the σ-algebra of all stochastic processes of interest.

Definition 51 *A stochastic process* $\{(B_1(t), \ldots, B_d(t)), t \geq 0\}$ *is called a d-dimensional* Brownian motion *process if*

(i) *Initial condition:* $B_i(0) = 0$, $i = 1, \ldots, d$;

(ii) *Independent increments:* $B_i(t) - B_i(s)$, $i = 1, \ldots, d$, *are independent of the history* $\mathcal{F}_s$;

(iii) *Normal distribution:* $B_i(t) - B_i(s)$, $i = 1, \ldots, d$, *are normally distributed with mean 0 and covariance matrix* $(t-s)\mathbf{I}_d$, *where* $\mathbf{I}_d$ *is the* $d \times d$ *identity matrix.*

	dt	$dB_i(t)$	$dB_j(t)$
dt	0	0	0
$dB_i(t)$	0	dt	0
$dB_j(t)$	0	0	dt

Table 2: The rule of thumb

Theorem 52 *(Itô's calculus) Let $f : I\!R^{m+1} \mapsto I\!R$ be a function that has continuous second partial with respect to $x_1, \ldots, x_m$ and continuous first partial with respect to x_{m+1}. Suppose that*

$$\phi_i(t) = \phi_i(0) + \int_0^t \psi_i(s)ds + \sum_{k=1}^{d} \int_0^t \xi_{i,k}(s)dB_k(s), \quad i = 1, \ldots, m, \tag{349}$$

where $\int_0^t \psi_i(s)ds$ is the term with a bounded variation and $\sum_{k=1}^d \int_0^t \xi_{i,k}(s)dB_k(s)$ is the sum of stochastic integrals. Then

$$\begin{aligned} f(\phi_1(t), \ldots, \phi_m(t), t) &= f(\phi_1(0), \ldots, \phi_m(0), 0) + \int_0^t \frac{\partial f(\phi_1(s), \ldots, \phi_m(s), s)}{\partial s} ds \\ &+ \sum_{i=1}^{m} \int_0^t \frac{\partial f(\phi_1(s), \ldots, \phi_m(s), s)}{\partial x_i} \psi_i(s)ds + \sum_{i=1}^{m} \int_0^t \frac{\partial f(\phi_1(s), \ldots, \phi_m(s), s)}{\partial x_i} \sum_{k=1}^{d} \xi_{i,k}(s)dB_k(s) \\ &+ \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \int_0^t \frac{\partial^2 f(\phi_1(s), \ldots, \phi_m(s), s)}{\partial x_i \partial x_j} \sum_{k=1}^{d} \xi_{i,k}(s)\xi_{j,k}(s)ds \end{aligned} \tag{350}$$

This can be easily verified using the multiplication table in Table 2.

Now consider a set of stochastic differential equations:

$$d\phi_i(t) = \mu_i(\phi_1(t), \ldots, \phi_m(t))dt + \sum_{k=1}^{d} \sigma_{i,k}(\phi_1(t), \ldots, \phi_m(t))dB_k(t), \quad i = 1, \ldots, m, \tag{351}$$

where $\mu_i : I\!R^m \mapsto I\!R$ and $\sigma_{i,k} : I\!R^m \mapsto I\!R$. Define the transition (density) function

$$\begin{aligned} & p(x_1, \ldots, x_m, t, y_1, \ldots, y_m)dx_1 \cdots dx_m \\ & = Pr(x_i \le \phi_i(t) \le x_i + dx_i, i = 1, \ldots, m | \phi_i(0) = y_i, i = 1, \ldots, m) \end{aligned} \tag{352}$$

Then the corresponding backward equation is

$$\begin{aligned}&\frac{\partial p(x_1,\ldots,x_m,t,y_1,\ldots,y_m)}{\partial t}\\&=\sum_{i=1}^m \mu_i(y_1,\ldots,y_m)\frac{\partial}{\partial y_i}p(x_1,\ldots,x_m,t,y_1,\ldots,y_m)\\&+\frac{1}{2}\sum_{i=1}^m\sum_{j=1}^m a_{i,j}(y_1,\ldots,y_m)\frac{\partial^2}{\partial y_i\partial y_j}p(x_1,\ldots,x_m,t,y_1,\ldots,y_m),\end{aligned}\tag{353}$$

where

$$a_{i,j}(y_1,\ldots,y_m)=\sum_{k=1}^d \sigma_{i,k}(y_1,\ldots,y_m)\sigma_{j,k}(y_1,\ldots,y_m).\tag{354}$$

The corresponding forward equation is

$$\begin{aligned}&\frac{\partial p(x_1,\ldots,x_m,t,y_1,\ldots,y_m)}{\partial t}\\&=-\sum_{i=1}^m \frac{\partial}{\partial x_i}[\mu_i(x_1,\ldots,x_m)p(x_1,\ldots,x_m,t,y_1,\ldots,y_m)]\\&+\frac{1}{2}\sum_{i=1}^m\sum_{j=1}^m \frac{\partial^2}{\partial x_i\partial x_j}[a_{i,j}(x_1,\ldots,x_m)p(x_1,\ldots,x_m,t,y_1,\ldots,y_m)]\end{aligned}\tag{355}$$

One can also extend the Kalman-Bucy filter and the LQG regulators to the multidimensional case. See e.g. V. Krishnan [4]. The extension is based on the notion of *conditional expectation*. For the filtering problem in Section 5.1, let $\mathcal{F}_t^o$ be the σ-algebra generated by $\{\phi(s), 0\leq s\leq t\}$, i.e., $\mathcal{F}_t^o$ is the observed history. We will argue that the best estimate that minimizes the mean square error is $E(\eta(t)|\mathcal{F}_t^o)$, i.e.,

$$E(\eta(t)-E(\eta(t)|\mathcal{F}_t^o))^2\leq E(\eta(t)-\hat{\eta}(t))^2\tag{356}$$

for all $\hat{\eta}(t)$ that is $\mathcal{F}_t^o$-measurable ($\hat{\eta}(t)$ is determined by the observed history). To see this, note that

$$\begin{aligned}&E(\eta(t)-\hat{\eta}(t))^2\\&=E(\eta(t)-E(\eta(t)|\mathcal{F}_t^o)+E(\eta(t)|\mathcal{F}_t^o)-\hat{\eta}(t))^2\\&=E\Big(\eta(t)-E(\eta(t)|\mathcal{F}_t^o))^2+E(E(\eta(t)|\mathcal{F}_t^o)-\hat{\eta}(t)\Big)^2\\&+2E\Big((\eta(t)-E(\eta(t)|\mathcal{F}_t^o))(E(\eta(t)|\mathcal{F}_t^o)-\hat{\eta}(t))\Big).\end{aligned}\tag{357}$$

Note that $E(E(X|Y)) = E(X)$. Conditioning on $\mathcal{F}_t^o$, the last term in (357) becomes

$$\begin{aligned} & E\Big(E\Big((\eta(t) - E(\eta(t)|\mathcal{F}_t^o))(E(\eta(t)|\mathcal{F}_t^o) - \hat{\eta}(t))\Big|\mathcal{F}_t^o\Big)\Big) \\ &= E\Big((E(\eta(t)|\mathcal{F}_t^o) - \hat{\eta}(t))E\Big((\eta(t) - E(\eta(t)|\mathcal{F}_t^o))\Big|\mathcal{F}_t^o\Big)\Big) \\ &= E\Big((E(\eta(t)|\mathcal{F}_t^o) - \hat{\eta}(t))\Big(E(\eta(t)|\mathcal{F}_t^o) - E(\eta(t)|\mathcal{F}_t^o)\Big)\Big) \\ &= 0. \end{aligned} \tag{358}$$

Since $E(E(\eta(t)|\mathcal{F}_t^o) - \hat{\eta}(t))^2 \geq 0$, one derives (356). Thus, the Kalman-Bucy filter is equivalent to finding the conditional expectation $E(\eta(t)|\mathcal{F}_t^o)$. Let $e(t) = \eta(t) - E(\eta(t)|\mathcal{F}_t^o)$ be the error term. This is also equivalent to finding the error term such that

$$E(e(t)|\mathcal{F}_t^o) = 0. \tag{359}$$

Equation (359) is exactly the notion of *martingale*. Based on the notion of *conditional expectation*, the stochastic integral with respect to the Brownian motion can be extended to the integrals with respect to martingales. The Itô calculus is then enlarged and known as the *martingale calculus*. For further reading on martingale calculus, see e.g., Karatzas and Shreve [2], Wong and Hajek [9].

7 The Karhunen-Loéve expansions and Einstein's construction of the Brownian motion

Consider a complex valued stochastic process $X(t)$. Let $X^*(t)$ be the conjugate of $X(t)$ and

$$R(t,s) = E[X(t)X^*(s)] \tag{360}$$

be its autocorrelation function.

Theorem 53 *(The Karhunen-Loéve expansion) Suppose that* $\{X(t), 0 \leq t \leq T\}$, *is* continuous *in the* L^2 *sense, i.e.,*

$$\lim_{h\to 0} E|X(t+h) - X(t)|^2 = 0. \tag{361}$$

(i) *If* $\{\phi_n(t)\}$ *are* orthonormal *eigenfunctions of*

$$\int_0^T R(t,s)\phi(s)ds = \lambda\phi(t), \tag{362}$$

and $\{\lambda_n\}$ *the eigenvalues, then*

$$E|X(t) - \hat{X}(t)|^2 = 0, \quad 0 \le t \le T, \tag{363}$$

where

$$\hat{X}(t) = \sum_{n=1}^{\infty} c_n \phi_n(t) \tag{364}$$

$$c_n = \int_0^T X(t)\phi_n^*(t)dt \tag{365}$$

and

$$E c_n c_m^* = \lambda_n \delta_{n,m}. \tag{366}$$

(ii) *Conversely, if* $X(t)$ *has an expansion of the form (364) with*

$$\int_0^T \phi_m(t)\phi_n^*(t)dt = \delta_{n,m} = \frac{E c_n c_m^*}{\lambda_n}, \tag{367}$$

then $\{\phi_n(t)\}$ *and* $\{\lambda_n\}$ *must be the eigenfunctions and eigenvalues of (362).*

The proof of Theorem 53 is based on Mercer's theorem on the eigenfunctions and eigenvalues of the integral equation in (362).

1. Any eigenvalue of (362) must be *real* and *positive*.
To see this, note that $R^*(t,s) = R(s,t)$. Thus, taking the conjugate on both sides of (362) yields

$$\int_0^T R(s,t)\phi^*(s)ds = \lambda^*\phi^*(t). \tag{368}$$

This implies

$$\int_0^T \int_0^T \phi(t)R(s,t)\phi^*(s)dsdt = \lambda^* \int_0^T \phi^*(t)\phi(t)dt. \tag{369}$$

Similarly, we have from (362) that

$$\int_0^T \int_0^T \phi^*(t)R(t,s)\phi(s)dsdt = \lambda \int_0^T \phi(t)\phi^*(t)dt. \tag{370}$$

Note that the right hand sides of (369) and (370) are equal. Thus, $\lambda = \lambda^*$ (since $\int_0^T \phi^*(t)\phi(t)dt > 0$).

To see that λ is positive, note that the autocorrelation function is *positive definite*, i.e., for any m, and $a_i, i = 1, \ldots, m$,

$$\sum_{i=1}^{m} \sum_{j=1}^{m} a_i a_j^* R(t_i, t_j) \geq 0. \tag{371}$$

This is due to the identity

$$0 \leq E|\sum_{i=1}^{m} a_i X(t_i)|^2 = \sum_{i=1}^{m} \sum_{j=1}^{m} a_i a_j^* E X(t_i) X^*(t_j). \tag{372}$$

Thus, the right hand side of (370) is nonnegative and this in turn implies λ is positive (the case $\lambda = 0$ is trivial).

2. If $R(t, s)$ is not identical zero, there is at least one eigenvalue for (362). The largest eigenvalue λ_1 is given by

$$\lambda_1 = \max_{||\phi(\cdot)||=1} \int_0^T \int_0^T \phi^*(t) R(t, s) \phi(s) ds dt, \tag{373}$$

where

$$||\phi(\cdot)|| = \Big(\int_0^T \phi^*(t) \phi(t) dt \Big)^{\frac{1}{2}}. \tag{374}$$

The proof is not easy and will not be presented here.

3. Let $\phi_1(t)$ be the normalized eigenfunction corresponding to the eigenvalue λ_1, i.e.,

$$\int_0^T \phi_1(t) \phi_1^*(t) dt = 1. \tag{375}$$

4. Let $R_2(t, s) = R(t, s) - \lambda_1 \phi_1(t) \phi_1^*(s)$. Then $R_2(t, s)$ is the autocorrelation function of the process

$$Y(t) = X(t) - \phi_1(t) \int_0^T X(s) \phi_1^*(s) ds, \tag{376}$$

and thus $R_2(t,s)$ is positive definite. To see this, note that

$$
\begin{aligned}
&EY(t)Y^*(s) = EX(t)X^*(s) - E\Big[X(t)\phi_1^*(s)\int_0^T X^*(u_1)\phi_1(u_1)du_1\Big] \\
&-E\Big[X^*(s)\phi_1(t)\int_0^T X(u_2)\phi_1^*(u_2)du_2\Big] \\
&+E\Big[\phi_1(t)\phi_1^*(s)\int_0^T\int_0^T X(u_2)X^*(u_1)\phi_1(u_1)\phi_1^*(u_2)du_1du_2\Big] \\
&= R(t,s) - \lambda_1\phi_1(t)\phi_1^*(s) \\
&= R_2(t,s)
\end{aligned}
\tag{377}
$$

5. Observe that

$$\int_0^T R_2(t,s)\phi_1(s)ds = 0. \tag{378}$$

Now repeat step 2 and find λ_2 and $\phi_2(\cdot)$ for $R_2(t,s)$. Then

$$\int_0^T R_2(t,s)\phi_2(s)ds = \lambda_2\phi_2(t). \tag{379}$$

It follows that

$$\int_0^T \phi_2(t)\phi_1^*(t)dt = \frac{1}{\lambda_2}\int_0^T \phi_2(s)(\int_0^T R_2(s,t)\phi_1(t)dt)^*ds = 0. \tag{380}$$

Thus, $\phi_2(\cdot)$ is orthogonal to $\phi_1(\cdot)$. Moreover,

$$
\begin{aligned}
&\int_0^T R(t,s)\phi_2(s)ds \\
&= \int_0^T R_2(t,s)\phi_2(s)ds + \lambda_1\phi_1(t)\int_0^T \phi_2(s)\phi_1^*(s)ds \\
&= \lambda_2\phi_2(t).
\end{aligned}
\tag{381}
$$

This implies λ_2 and $\phi_2(\cdot)$ are eigenvalue and eigenfunction of (362).

6. Iterating the procedure yields a decreasing sequence of eigenvalues $\lambda_1, \lambda_2, \ldots, ...$ and a corresponding sequence of eigenfunctions $\phi_1(\cdot), \phi_2(\cdot), \ldots$. Moreover, these eigenfunctions are orthonormal, i.e.,

$$\int_0^T \phi_n(t)\phi_m(t)dt = \delta_{n,m}. \tag{382}$$

7. (Mercer's theorem)

$$R(t,s) = \sum_{n=1}^{\infty} \lambda_n \phi_n(t)\phi_n^*(s), \;\; 0 \le s,t \le T. \tag{383}$$

The number of terms in the sum might be finite. If it is infinite, then $\lambda_n \to 0$ as $n \to \infty$.

Proof. (Sketch of the Proof of Theorem 53) (Below, we assume that all expectations and all infinite sums can be interchanged.) (i) From (365) and the orthonormality of $\phi_n(\cdot)$, it follows that

$$Ec_n X^*(s) = \int_0^T R^*(s,t)\phi_n^*(t)dt = \lambda_n \phi_n^*(s), \tag{384}$$

where we use the fact that λ_n is real. This implies

$$Ec_n c_m^* = Ec_n \int_0^T X^*(t)\phi_m(t)dt = \lambda_m \int_0^T \phi_n^*(t)\phi_m(t)dt = \lambda_n \delta_{n,m}. \tag{385}$$

Hence

$$Ec_n \hat{X}^*(t) = Ec_n \sum_{m=1}^{\infty} c_m^* \phi_m^*(t) = \sum_{m=1}^{\infty} Ec_n c_m^* \phi_m^*(t) = \lambda_n \phi_n^*(t). \tag{386}$$

Thus, it follows from the Mercer theorem in (383) that

$$\begin{aligned} E\hat{X}(t)\hat{X}^*(t) &= E\sum_{n=1}^{\infty} c_n \phi_n(t)\hat{X}^*(t) = \sum_{n=1}^{\infty} \lambda_n \phi_n(t)\phi_n^*(t) \\ &= R(t,t) = E\hat{X}^*(t)X(t) = E|X(t)|^2 = E|\hat{X}(t)|^2. \end{aligned} \tag{387}$$

Thus, we derive (363).

(ii) To see the converse, assume that $X(t)$ has the expansion in (364). From the expansion and (367), it follows that

$$EX(t)c_m^* = \sum_{n=1}^{\infty} Ec_n c_m^* \phi_m(t) = \lambda_m \phi_m(t). \tag{388}$$

Using c_n in (365) yields

$$EX(t)c_m^* = \int_0^T EX(t)X^*(s)\phi_m(s)ds = \int_0^T R(t,s)\phi_m(s)ds. \tag{389}$$

From (388) and (389), we show that λ and $\phi(\cdot)$ are eigenvalue and eigenfunction of (362). $\square$

For the standard Brownian motion, we have shown in Proposition 5 that the autocorrelation function is

$$R(t,s) = \min(t,s). \tag{390}$$

The corresponding integral equation is

$$\int_0^T \min(t,s)\phi(s)ds = \lambda\phi(t), \quad 0 \le t \le T, \tag{391}$$

or

$$\int_0^t s\phi(s)ds + t\int_t^T \phi(s)ds = \lambda\phi(t). \tag{392}$$

Differentiating with respect to t yields

$$\int_t^T \phi(s)ds = \lambda\phi'(t). \tag{393}$$

Differentiating once more, we obtain

$$\lambda\phi''(t) + \phi(t) = 0, \tag{394}$$

with the obvious boundary conditions $\phi(0) = 0$ (see (392)) and $\phi'(T) = 0$ (see (393)). The solution of (394) with $\phi(0) = 0$ is

$$\phi(t) = A\sin\frac{1}{\sqrt{\lambda}}t \tag{395}$$

for some constant A. Using the condition $\phi'(T) = 0$ yields

$$\cos\frac{1}{\sqrt{\lambda}}T = 0. \tag{396}$$

Thus, the eigenvalues are

$$\lambda_n = \frac{T^2}{(n-\frac{1}{2})^2\pi^2}, \ n = 1, 2, \ldots \tag{397}$$

The normalized eigenfunctions are

$$\phi_n(t) = \sqrt{\frac{2}{T}}\sin[(n-\frac{1}{2})\pi(\frac{t}{T})]. \tag{398}$$

From Theorem 53, we have the expansion of the standard Brownian motion

$$\hat{B}(t) = \sum_{n=1}^{\infty} c_n \phi_n(t), \tag{399}$$

where

$$c_n = \int_0^T B(t)\phi_n^*(t)dt = \int_0^t B(t)\phi_n(t)dt, \tag{400}$$

with $\phi_n(t)$ in (398). Moreover, c_n's are *uncorrelated.* In view of the Gaussian nature of the standard Brownian motion, c_n's are (joint) Gaussian random variables and hence they are in fact *independent.* This leads to a way to construct the standard Brownian motion, namely the Einstein construction.

Theorem 54 *(Einstein's construction of the standard Brownian motion) Let c_n, $n = 1, 2, \ldots$, be a sequence of* independent *Gaussian random variables with mean 0 and variance*

$$\lambda_n = \frac{T^2}{(n - \frac{1}{2})^2 \pi^2}. \tag{401}$$

Then

$$B(t) = \sum_{n=1}^{\infty} c_n \phi_n(t), \quad 0 \leq t \leq T \tag{402}$$

is the standard Brownian motion, where

$$\phi_n(t) = \sqrt{\frac{2}{T}} \sin[(n - \frac{1}{2})\pi(\frac{t}{T})]. \tag{403}$$

Proof. (Sketch of the proof of Theorem 54) Since c_n's are independent Gaussian random variables, $B(t)$ in (402) is a Gaussian process. Since $Ec_n = 0$, $EB(t) = 0$. It remains to verify that $EB(t)B(s) = \min(t, s)$. From Mercer's theorem in (383) and the Karhunen-Loéve expansion in (399), it follows that

$$\min(t, s) = \sum_{n=1}^{\infty} \lambda_n \phi_n(t)\phi_n(s), \tag{404}$$

with $\phi_n(\cdot)$ and λ_n specified in (403) and (401) (note that $\phi_n(\cdot)$'s are real-valued functions). Since $Ec_n c_m = \lambda_n \delta_{n,m}$, one has

$$Ec_n B(s) = Ec_n \sum_{m=1}^{\infty} c_m \phi_m(s) = \sum_{m=1}^{\infty} Ec_n c_m \phi_m(s) = \lambda_n \phi_n(s). \tag{405}$$

Thus,

$$EB(t)B(s) = E\sum_{n=1}^{\infty} c_n\phi_n(t)B(s) = \sum_{n=1}^{\infty} \lambda_n\phi_n(s)\phi_n(t). \tag{406}$$

In conjunction with (404), we derive $EB(t)B(s) = \min(t, s)$. □

From the construction, one can "simulate" a sample path of the standard Brownian motion by simulating a large number of Gaussian random variables as in Theorem 54. We note that other L^2-complete orthonormal basis can also be used to construct the standard Brownian motion. For other constructions of the standard Brownian motion, see e.g., Karatzas and Shreve [2]. For further reading on Karhunen-Loéve expansion, see e.g. Wong and Hajek [9], or Papoulis [6].

References

[1] M.H.A. Davis and R.B. Vinter, *Stochastic Modelling and Control*. New York: Chapman and Hall, 1985.

[2] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*. New York: Springer-Verlag, 1988.

[3] S. Karlin and H.M. Taylor. *A second course in stochastic processes*. New York: Academic Press, 1975.

[4] V. Krishnan, *Nonlinear filtering and smoothing: an introduction to martingales, stochastic integrals, and estimation*. New York: Wiley, 1984.

[5] B.B. Mandelbrot and J.V. Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, Vol. 10, No. 4, pp. 422-437, 1968.

[6] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. New York: MaGraw-Hill, 1984.

[7] S.M. Ross, *Stochastic Processes*. New York: J. Wiley & Sons, 1983.

[8] H.L. Royden. *Real Analysis*. New York: The Macmillan Company, 1986.

[9] E. Wong and B. Hajek. *Stochastic processes in engineering systems*. New York: Springer-Verlag, 1984.