Stochastic Calculus

João Guerra

2016

Contents

Chapter 1

Introduction

The main objective of these notes is to introduce the main concepts of stochastic calculus to the master, post-graduation, or doctorate students in Mathematical Finance. In particular, this is the principal text resource to the *Stochastic Calculus* course of the Mathematical Finance Master degree at Lisbon School of Business and Economics, for the year of 2015/2016.

It is intended to explore the main techniques and methods of stochastic calculus and stochastic differential equations, to relate stochastic differential equations with partial differential equations, as well as applying methods and techniques studied to problems of mathematical finance, such as valuation and hedging of the risk of financial derivatives, such as futures and options.

In order to make the most of these notes, the reader should already have knowledge (up to an undergraduate level) of integral and differential calculus, probability (and some measure) theory and stochastic processes theory. Previous knowledge of partial and ordinary differential equations is useful, although it is not necessary.

1.1 What is Stochastic Calculus?

In a summary way, stochastic calculus is a type of integral and differential calculus that involves continuous-time stochastic processes, like the Brownian motion. It allows us to define integrals of stochastic processes, where the "integrating function" is also a stochastic process. We may also define and solve stochastic differential equations, that are basically ordinary differential equations with an extra random term.

The most important stochastic process when treating with financial applications, paradigmatic in the development of stochastic calculus, is the Brownian motion. Because of this, we'll focus the study of integration theory with respect to the Brownian motion. The fundamental topics presented in these notes are the construction of the stochastic integral, Itô's formula, stochastic differential equations, the Girsanov Theorem, the relation between stochastic and partial differential equations and applications to the Black-Scholes derivatives pricing model. Before we proceed with these topics, we'll cover some fundamental results on stochastic processes, conditional expectation, martingales and Brownian motion.

There are other important financial applications of stochastic calculus besides the valuation and hedging the risk of financial derivatives, as modelling interest rate term-structures and credit risk.

The theory on Brownian motion and stochastic calculus was developed by some of the most important mathematicians and physicists of the 20th century. With strong contributions to the area, we may highlight Louis Bachelier, Albert Einstein, Norbert Wiener, Andrey Kolmogorov, Vincent Doeblin, Kiyosi Itô, Joseph Doob and Paul-André Meyer. A brief history of stochastic calculus and it's financial applications is presented by Jarrow and Protter in their highly recommended article [4].

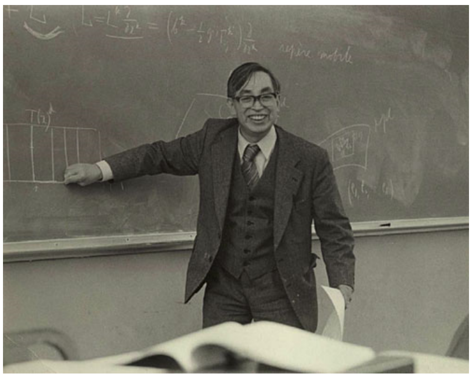

Figure 1.1: Kiyosi Itô

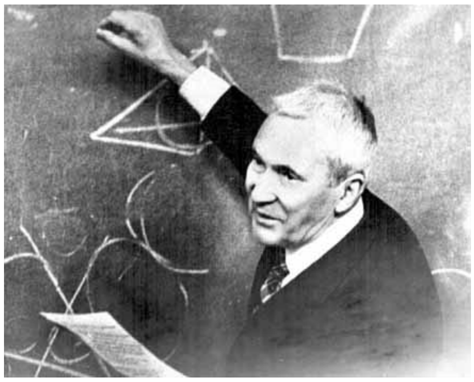

Figure 1.2: Andrey Kolmogorov

Chapter 2

Probability Theory and Stochastic Processes

2.1 Stochastic Processes

We start by the classic definition of a stochastic process.

Definition 2.1 *A stochastic process is a family of random variables* $\{X_t, t \in T\}$ *defined in a probability state* $(\Omega, \mathcal{F}, P)$*, where* T *is the set on which the parameter* t *is defined. The process is said to be discrete-time if* $T = \mathbb{N}$*, and continuous-time if* $T = [a, b] \subset \mathbb{R}$ *or* $T = \mathbb{R}$*.*

A stochastic process may be considered as a map of two variables: $t \in T$ and $\omega \in \Omega$:

$$\{X_t, t \in T\} = \{X_t(\omega), \omega \in \Omega, t \in T\},$$

where X_t represents the state or position of the process at time t. The state space (space where the random variables take values)is usually $\mathbb{R}$ (process with a continuous state space) or $\mathbb{N}$ (discrete state space).

For each fixed $\omega \in \Omega$, the map $t \to X_t(\omega)$ ou $X_\cdot(\omega)$ is called a trajectory of the process. Some examples of trajectories are presented next (one-dimension and two-dimension Brownian motion).

Example 2.2 *Consider a sequence of independent random variables* $\{Z_t, t \in \mathbb{N}\}$*. Then*

$$X_t = Z_1 + Z_2 + \cdots Z_t = X_{t-1} + Z_t$$

is a discrete-time stochastic process. This process is known as random walk.

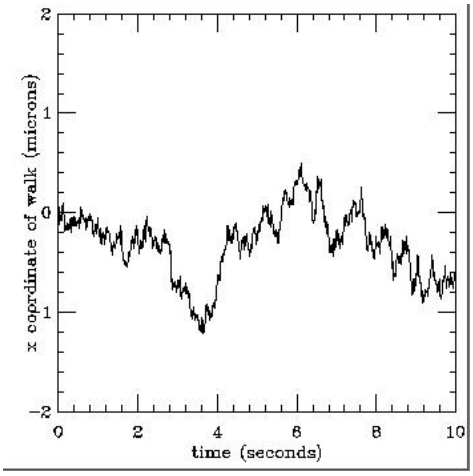

Figure 2.1: Trajectory of a one-dimension Brownian motion

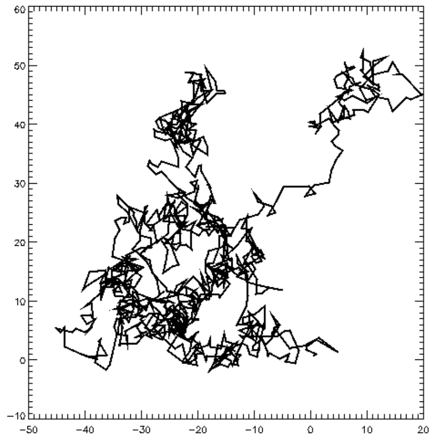

Figure 2.2: Trajectory of a two-dimension Brownian motion

A fundamental concept in the theory of stochastic processes is the concept of Markov process - where, "given the present, the future is independent from the past". In this sense, a Markov process is a process where the probability of obtaining a state in a future time t depends only on the last observed state t_k, that is, if $t_1 < t_2 < \cdots < t_k < t$, then

$$P\left[a < X_t < b | X_{t_1} = x_1, X_{t_2} = x_2, \ldots, X_{t_k} = x_k\right] = P\left[a < X_t < b | X_{t_k} = x_k\right].$$

A Markov process with a discrete state space is called a Markov chain. If the process is a continuous-time process with a continuous state space, it is called a diffusion.

In order to characterize probabilistically a process X, the concept of finite dimension distribution is used.

Definition 2.3 *Let $\{X_t, t \in T\}$ be a stochastic process. The finite dimension distributions (fdd) of X are all the distributions of the vectors*

$$\left(X_{t_1}, X_{t_2}, \ldots, X_{t_n}\right),$$

where $n = 1, 2, 3, \ldots; t_1, t_2, \ldots, t_n \in T$.

The law of probability or distribution of a stochastic process is identified with the family of the finite dimension distributions of that process.

Definition 2.4 *(Gaussian process) A process is called Gaussian when all the fdd are Gaussian.*

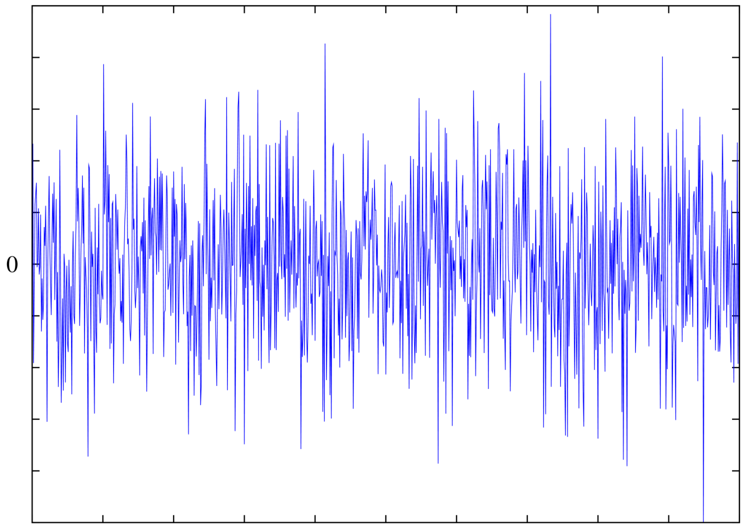

Figure 2.3: A trajectory of the white noise process

Knowing the parameters μ (expected value) and Σ (covariance) is enough to characterize a Gaussian distribution. Hence, to characterize a Gaussian process, one just needs to know μ and Σ for all vectors of type $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$.

Example 2.5 *(white noise) Let $\{X_t, t \geq 0\}$ be a stochastic process where $X_t \sim N(0, \sigma^2)$ and suppose all the random variables of the process are independent. Then the process is Gaussian and its fdd may be described by the distribution fuctions*

$$\begin{aligned} F(x_1, x_2, \ldots, x_n) &= P\left(X_{t_1} \leq x_1, X_{t_2} \leq x_2, \ldots, X_{t_n} \leq x_n\right) \\ &= P(X_{t_1} \leq x_1) P\left(X_{t_2} \leq x_2\right) \ldots P\left(X_{t_n} \leq x_n\right) \\ &= \Phi(x_1) \Phi\left(x_2\right) \ldots \Phi\left(x_n\right). \end{aligned}$$

The expected value and covariance functions of X are:

$$\begin{aligned} \mu_X\left(t\right) &= E\left[X_t\right] = 0, \\ c_X\left(s,t\right) &= \begin{cases} \sigma^2 & se\ s = t \\ 0 & se\ s \neq t \end{cases}. \end{aligned}$$

In general, given a process X, we may define the expected value and covariance functions as

$$\begin{aligned} \mu_X\left(t\right) &= E\left[X_t\right], \\ c_X\left(s,t\right) &= cov(X_t, X_s) = E\left[\left(X_t - \mu_X\left(t\right)\right)\left(X_s - \mu_X\left(s\right)\right)\right]. \end{aligned}$$

An important concept is the stationarity or invariance of a distribution. We'll now define what is meant by stationary process and process of stationary increments.

Definition 2.6 *A stochastic process X is said to be strictly (or strongly) stationary if*

$$\left(X_{t_1}, X_{t_2}, \ldots, X_{t_n}\right) \stackrel{d}{=} \left(X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_n+h}\right),$$

for all possible choices of n; $t_1, t_2, \ldots, t_n \in T$ and h.

Definition 2.7 *A stochastic process X is said to have stationary increments if*

$$X_t - X_s \stackrel{d}{=} X_{t+h} - X_{s+h},$$

for all possible values of s, t and h.

Exercise 2.8 *Show that if a process X is Gaussian and strongly stationary, then $\mu_X(t) = \mu_X(0)$, $\forall t \in T$ and $c_X(s,t) = f(|s-t|)$ depends only on the distance $|s-t|$.*

The independence of increments of a process is a fundamental property and will be vastly used when we discuss the stochastic integral. We now present the definition of this concept.

Definition 2.9 *A stochastic process is said to have independent increments if the random variables*

$$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \ldots, X_{t_n} - X_{t_{n-1}}$$

are independent when $t_1 < t_2 < \cdots < t_n$, $n = 1, 2, \ldots$

All the processes that have independent increments are Markov processes. Let's see an important example of one of these processes: the Poisson process.

Example 2.10 *(Poisson process) A stochastic process $\{X_t, t \geq 0\}$ is called a Poisson process with intensity λ se*

1. $X_0 = 0$,
2. *X has independent and stationary increments,*
3. $X_t \sim Poi(\lambda t)$.

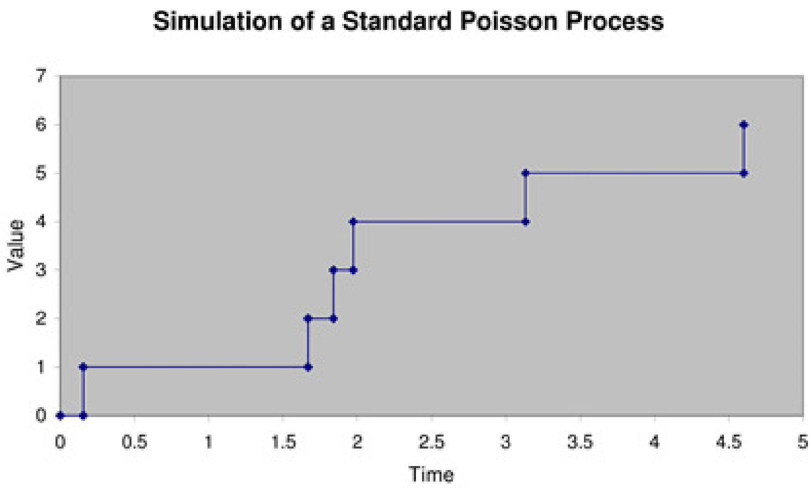

Figure 2.4: A trajectory of a Poisson process

A random variable Y has the (discrete) Poisson distribution of parameter λ, or $Poi(\lambda)$, if

$$P(Y = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Exercise 2.11 *Show that if X is a Poisson process, then $X_t - X_s \sim Poi(\lambda(t-s))$ if $t > s$.*

How may we define an equivalence relationship between processes? The next definition offers an answer to this question.

Definition 2.12 *A stochastic process $\{X_t, t \in T\}$ is said to be equivalent to another stochastic process $\{Y_t, t \in T\}$ if, for each $t \in T$ we have*

$$P\{X_t = Y_t\} = 1.$$

In this case we say that a process is a version of the other.

Note that two equivalent processes may have much different trajectories, as illustrated in the next example.

Example 2.13 *Let φ be a non-negative random variable with continuous distribution, and consider the stochastic processes*

$$X_t = 0,$$
$$Y_t = \begin{cases} 0 & se\ \varphi \neq t \\ 1 & se\ \varphi = t \end{cases}.$$

These processes are equivalent but their trajectories are different. The trajectories of Y always have a discontinuity point.

Definition 2.14 *Two stochastic processes $\{X_t, t \in T\}$ and $\{Y_t, t \in T\}$ are said to be indistinguishable if*

$$X_\cdot(\omega) = Y_\cdot(\omega) \quad \forall \omega \in \Omega \backslash N,$$

where N has null probability ($P(N) = 0$).

Two stochastic processes with continuous from the right (left, or simply continuous) trajectories that are equivalent are also indistinguishable. We may define other concepts of probability for stochastic processes, other than trajectory continuity.

Definition 2.15 *A continuous-time stochastic process $\{X_t; t \in T\}$ that takes values in $\mathbb{R}$ is said to be continuous in probability if, for any $\varepsilon > 0$ and for any $t \in T$, we have*

$$\lim_{s \to t} P\left[|X_s - X_t| > \varepsilon\right] = 0.$$

Definition 2.16 *Let $p \geq 1$. A continuous-time stochastic process $\{X_t; t \in T\}$ that takes values in $\mathbb{R}$, and such that $E\left[|X_t|^p\right] < \infty$, is said to be continuous in mean of order p, if for any $t \in T$, we have*

$$\lim_{s \to t} E\left[|X_s - X_t|^p\right] = 0.$$

Continuity in mean of order p implies continuity in probability. However, the reverse implication does not hold. Also, none of these continuity definitions imply the continuity of trajectories.

Example 2.17 *A Poisson process $N = \{N_t, t \geq 0\}$ with intensity λ is a process with discontinuous trajectories. However, it is continuous in mean of order 2 (mean square) (recall that $N_t - N_s \sim Poi(\lambda(t-s))$), because*

$$\lim_{s \to t} E\left[|N_t - N_s|^2\right] = \lim_{s \to t}\left[\lambda(t-s) + (\lambda(t-s))^2\right] = 0.$$

The Continuity Criterion of Kolmogorov is a useful theoretical tool that allows us to prove a given stochastic process has continuous trajectories. This result is presented below.

Theorem 2.18 *(Continuity Criterion of Kolmogorov): Let* $X = \{X_t; t \in T\}$ *be a stochastic process, where* T *is a bounded interval of* $\mathbb{R}$*, and suppose there exists* $p > 0$*,* $\alpha > 0$ *and* $C > 0$ *such that*

$$E\left[|X_t - X_s|^p\right] \leq C\,|t-s|^{1+\alpha}. \tag{2.1}$$

Then there exists a version of X *with continuous trajectories.*

More precisely, equation (2.1) implies that for each $\varepsilon > 0$ there exists a r.v. G_ε such that

$$|X_t(\omega) - X_s(\omega)| \leq G_\varepsilon(\omega)\,|t-s|^{\frac{1+\alpha}{p}-\varepsilon} \qquad a.s. \tag{2.2}$$

and $E\left[G_\varepsilon^p\right] < \infty$. That is, X has Hölder continuous trajectories of order β, for all $\beta < \frac{1+\alpha}{p}$.

A proof of this theorem may be found in [5], pgs. 53-54.

2.2 Conditional Expectation

Consider a probability space $(\Omega, \mathcal{F}, P)$ and let A and B be two events where A,$B \in \mathcal{F}$ and $P(B) > 0$. The conditional expectation of A given B may be defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2.3}$$

The map $A \to P(A|B)$ defines a probability measure in the σ-algebra $\mathcal{F}$. The expected value or conditional expectation of the integrable r.v. X given B may be computed using the formula

$$E(X|B) = \frac{E\left[X\mathbf{1}_B\right]}{P(B)}. \tag{2.4}$$

Example 2.19 *Let* X *be a uniform r.v. taking values in* $(0,1]$*. Let* $A = \left(0, \frac{1}{4}\right]$*. Let us compute* $E[X]$ *and* $E[X|A]$*.*

$$E[X] = \int_0^1 x f(x)\,dx = \int_0^1 x dx = \frac{1}{2}.$$

$$E[X|A] = \frac{E(X\mathbf{1}_A)}{P(A)} = \frac{\int_0^{1/4} x dx}{1/4} = \frac{1}{8}.$$

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{B} \subset \mathcal{F}$ a σ-algebra.

Definition 2.20 *The conditional expectation of the integrable random variable X given $\mathcal{B}$ (or $E(X|\mathcal{B})$) is an integrable random variable Z sucha that*

1. *Z is $\mathcal{B}$-measurable.*
2. *for all $A \in \mathcal{B}$ we have*

$$E\left(Z\mathbf{1}_A\right) = E\left(X\mathbf{1}_A\right). \tag{2.5}$$

If X is integrable then $Z = E(X|\mathcal{B})$ exists and it is unique (almost surely).

Definition 2.21 *(generated σ-algebra): Let $\mathcal{C}$ be a family of subsets of Ω. Then, the smaller σ-algebra that contains $\mathcal{C}$ is represented by $\sigma\left(\mathcal{C}\right))$ and it is called the σ-algebra generated by $\mathcal{C}$.*

Definition 2.22 *(σ-algebra generated by X): Let X be a random variable. The σ-algebra generated by X is defined as $\{X^{-1}(B) : B \in \mathcal{B}_{\mathbb{R}}\}$.*

We will now see some essential properties of the conditional expectation.

Proposition 2.23 *Let X, Y and Z be integrable random variables, $\mathcal{B}$ a σ-algebra and $a, b \in \mathbb{R}$. Then,*

1. $$E(aX + bY|\mathcal{B}) = aE(X|\mathcal{B}) + bE(Y|\mathcal{B}). \tag{2.6}$$
2. $$E\left(E(X|\mathcal{B})\right) = E\left(X\right). \tag{2.7}$$
3. *If X and the σ-algebra $\mathcal{B}$ are independent, then:*
$$E(X|\mathcal{B}) = E\left(X\right) \tag{2.8}$$
4. *If X is $\mathcal{B}$-measurable (or if $\sigma\left(X\right) \subset \mathcal{B}$) then:*
$$E(X|\mathcal{B}) = X. \tag{2.9}$$
5. *If Y is $\mathcal{B}$-measurable (or if $\sigma\left(X\right) \subset \mathcal{B}$) then:*
$$E(YX|\mathcal{B}) = YE(X|\mathcal{B}) \tag{2.10}$$
6. *Given two σ-algebras $\mathcal{C} \subset \mathcal{B}$, then:*
$$E(E\left(X|\mathcal{B}\right)|\mathcal{C}) = E(E\left(X|\mathcal{C}\right)|\mathcal{B}) = E(X|\mathcal{C}) \tag{2.11}$$

7. *Consider two r.v.'s X and Z such that Z is $\mathcal{B}$-measurable and X is independent from $\mathcal{B}$. Let $h(x, z)$ be a measurable function such that $h(X, Z)$ is an integrable random variable. Then,*

$$E\left(h\left(X, Z\right)|\mathcal{B}\right) = E\left(h\left(X, z\right)\right)|_{z=Z}. \tag{2.12}$$

   Remark: at first one computes $E\left(h\left(X, z\right)\right)$ for an arbitrarily fixed value z, and then z is substituted by the r.v. Z.

Proposition 2.24 *(Jensen's inequality): Let X be an integrable random variable and $\mathcal{B}$ a σ-algebra. If φ is a convex function such that $E\left[|\varphi\left(X\right)|\right] < \infty$, then*

$$\varphi\left(E(X|\mathcal{B})\right) \leq E(\varphi\left(X\right)|\mathcal{B}). \tag{2.13}$$

A particular case of the Jensen's inequality is obtained by considering $\varphi\left(x\right) = |x|^p$. If $E(|X|^p) < \infty$, $p \geq 1$, then

$$|E(X|\mathcal{B})|^p \leq E(|X|^p|\mathcal{B}).$$

Hence, if $p \geq 1$,

$$E\left[|E(X|\mathcal{B})|^p\right] \leq E(|X|^p). \tag{2.14}$$

The set of all random variables that are square integrable - $L^2\left(\Omega, \mathcal{F}, P\right)$ - is an Hilbert space with the inner product

$$\langle X, Y\rangle = E\left[XY\right].$$

The space $L^2\left(\Omega, \mathcal{B}, P\right)$ is a subspace of $L^2\left(\Omega, \mathcal{F}, P\right)$. Given a random variable $X \in L^2\left(\Omega, \mathcal{F}, P\right)$, $E(X|\mathcal{B})$ is the orthogonal projection of X in the subspace $L^2\left(\Omega, \mathcal{B}, P\right)$ and minimizes the mean-squared distance of X to $L^2\left(\Omega, \mathcal{B}, P\right)$ in the sense that

$$E\left[\left(X - E(X|\mathcal{B})\right)^2\right] = \min_{Y \in L^2(\Omega, \mathcal{B}, P)} E\left[\left(X - Y\right)^2\right] \tag{2.15}$$

Exercise 2.25 *Show that if X and the σ-algebra $\mathcal{B}$ are independent, then $E(X|\mathcal{B}) = E\left(X\right)$*

Solution 2.26 *If X and $\mathbf{1}_A$ are independent then if $A \in \mathcal{B}$,*

$$E\left[X\mathbf{1}_A\right] = E\left[X\right]E\left[\mathbf{1}_A\right] = E\left[E\left[X\right]\mathbf{1}_A\right]$$

and, by the definition of conditional expectation,$E(X|\mathcal{B}) = E\left(X\right)$.

Exercise 2.27 *Show that if Y is a $\mathcal{B}$-measurable random variable, then*

$$E(YX|\mathcal{B}) = YE(X|\mathcal{B}).$$

Solution 2.28 *If $Y = \mathbf{1}_A$ and $A, B \in \mathcal{B}$, then, by the definition of conditional expectation,*

$$\begin{aligned} E\left[\mathbf{1}_A E(X|\mathcal{B})\mathbf{1}_B\right] &= E\left[\mathbf{1}_{A\cap B} E(X|\mathcal{B})\right] \\ &= E\left[X\mathbf{1}_{A\cap B}\right] = E\left[\mathbf{1}_B \mathbf{1}_A X\right]. \end{aligned}$$

Therefore, $\mathbf{1}_A E(X|\mathcal{B}) = E\left[\mathbf{1}_A X|\mathcal{B}\right]$. In the same way we'll obtain the result if $Y = \sum_{j=1}^m a_j \mathbf{1}_{A_j}$ (i.e., Y is a $\mathcal{B}$-measurable stair function). The general result is proven by aproximating Y by a sequence of $\mathcal{B}$-measurable stair functions.

Example 2.29 *Given the random variable $X \in L^2(\Omega, \mathcal{F}, P)$, we shall show that $E(X|\mathcal{B})$ is the orthogonal projection of X in the subspace $L^2(\Omega, \mathcal{B}, P)$, and*

$$E\left[(X - E(X|\mathcal{B}))^2\right] = \min_{Y\in L^2(\Omega,\mathcal{B},P)} E\left[(X-Y)^2\right]$$

(1) $E(X|\mathcal{B}) \in L^2(\Omega, \mathcal{B}, P)$, because X is $\mathcal{B}$-measurable. By (2.14),

$$E\left[|E(X|\mathcal{B})|^2\right] \le E(|X|^2) < \infty.$$

(2) If $Z \in L^2(\Omega, \mathcal{B}, P)$ then, by the properties 2 and 5 of the conditional expectation,

$$\begin{aligned} E\left[(X - E(X|\mathcal{B}))\, Z\right] &= E\left[XZ\right] - E\left[E(X|\mathcal{B})Z\right] \\ &= E\left[XZ\right] - E\left[E(XZ|\mathcal{B})\right] \\ &= 0 \end{aligned}$$

hence $(X - E(X|\mathcal{B}))$ is orthogonal to $L^2(\Omega, \mathcal{B}, P)$.

(3) As

$$E\left[(X-Y)^2\right] = E\left[(X - E(X|\mathcal{B}))^2\right] + E\left[(E(X|\mathcal{B}) - Y)^2\right]$$

we have that $E\left[(X-Y)^2\right] \ge E\left[(X - E(X|\mathcal{B}))^2\right]$ and therefore

$$E\left[(X - E(X|\mathcal{B}))^2\right] = \min_{Y\in L^2(\Omega,\mathcal{B},P)} E\left[(X-Y)^2\right].$$

Exercise 2.30 *Prove properties 1,2,4 and 6 of the conditional expectation (Proposition 2.23)*

2.3 Discrete-time Martingales

The concept of martingale is one of the most fruitful in stochastic analysis. To define a martingale, one should previously define the concept of filtration. Consider the probability space $(\Omega, \mathcal{F}, P)$.

Definition 2.31 *A sequence of σ-algebras $\{\mathcal{F}_n, n \geq 0\}$ such that*

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n \subset \cdots \subset \mathcal{F}$$

is called a filtration.

A filtration may be interpreted as representing the flow of information generated by a random experiment or by a stochastic process.

Definition 2.32 *A discrete-time stochastic process $M = \{M_n; n \geq 0\}$ is said to be a martingale with respect to the filtration $\{\mathcal{F}_n, n \geq 0\}$ if:*

1. *For each n, M_n is a $\mathcal{F}_n$-measurable r.v. (i.e., M is a stochastic process adapted to the filtration $\{\mathcal{F}_n, n \geq 0\}$).*
2. *For all n, $E[|M_n|] < \infty$.*
3. *For all n, we have*
$$E[M_{n+1}|\mathcal{F}_n] = M_n. \tag{2.16}$$

The stochastic process $M = \{M_n; n \geq 0\}$ is called a supermartingale (resp. submartingale) if it verifies conditions 1 and 2 of the previous definition and if condition 3 is substituted by (3') $E[M_{n+1}|\mathcal{F}_n] \leq M_n$ (resp. (3") $E[M_{n+1}|\mathcal{F}_n] \geq M_n$).

From condition (3) (or eq. (2.16)) it is easy to show that

$$E[M_n] = E[M_0]$$

for all $n \geq 1$. That is, the expected value of a martingale is constant with time.

Condition (3) - eq. (2.16) - is equivalent to

$$E[\Delta M_n|\mathcal{F}_{n-1}] = 0.$$

for all $n \geq 1$, where $\Delta M_n := M_n - M_{n-1}$.

The martingale condition - eq. (2.16) - may be interpreted as: given the information $\mathcal{F}_n$, M_n is the best estimation for M_{n+1}.

Exercise 2.33 *Show that, if* $M = \{M_n; n \geq 0\}$ *is a martingale, then*

$$E[M_n] = E[M_0], \quad \forall n \geq 1.$$

Example 2.34 *(random walk): Let* $\{Z_n; n \geq 0\}$ *be a sequence of integrable and independent random variables with 0 expected value. Let* $M = \{M_n; n \geq 0\}$ *be defined as*

$$M_n = Z_0 + Z_1 + \cdots + Z_n.$$

The process M *is a random walk. Consider the natural filtration generated by* $\{Z_n; n \geq 0\}$*, i.e.,*

$$\mathcal{F}_n := \sigma\{Z_0, Z_1, \ldots, Z_n\}.$$

As $M_0, M_1, \ldots, M_n$ *and* $Z_0, Z_1, \ldots, Z_n$ *contain the same information, then both should generate the same* σ*-algebra* $\mathcal{F}_n$*. Let us show that* M *is a martingale.*

1. M *is adapted to the filtration* $\{\mathcal{F}_n, n \geq 0\}$ *, because* M_n *is* $\mathcal{F}_n$*-measurable, as* $\mathcal{F}_n$ *is also generated by* M_n*.*

2. $E[|M_n|] < \infty$*, because all r.v.* Z_n *are integrable (i.e.* $E[|Z_n|] < \infty$ *for all* n*).*

Example 2.35 *Pelas propriedades básicas da esperança condicionada:*

$$\begin{aligned} E[M_{n+1}|\mathcal{F}_n] &= E[M_n + Z_{n+1}|\mathcal{F}_n] \\ &= M_n + E[Z_{n+1}|\mathcal{F}_n] \\ &= M_n + E[Z_{n+1}] \\ &= M_n. \end{aligned}$$

Note that a σ-algebra $\sigma(X_1, X_2, \ldots, X_n)$ generated by the r.v. $(X_1, X_2, \ldots, X_n)$ contains all the essensial information on the structure of the random vector $(X_1, X_2, \ldots, X_n)$ (as a function of $\omega \in \Omega$). Any martingale with respect to a filtration $\mathcal{G}$ is also a martingale with respect to the filtration generated by the process itself (smaller filtration).

Lema 2.36 *Let* $M = \{M_n; n \geq 0\}$ *be a martingale with respect to* $\{\mathcal{G}_n, n \geq 0\}$ *and* $\mathcal{F}_n = \sigma\{M_0, M_1, \ldots, M_n\} \subset \mathcal{G}_n$ *the filtration generated by the process* M*. Then* M *is a martingale with respect to* $\{\mathcal{F}_n, n \geq 0\}$*.*

Proof. By property 6 of the conditional expectation and by the martingale property,

$$\begin{aligned} E\left[M_{n+1}|\mathcal{F}_n\right] &= E\left[E\left[M_{n+1}|\mathcal{G}_n\right]|\mathcal{F}_n\right] \\ &= E\left[M_n|\mathcal{F}_n\right] \\ &= M_n. \end{aligned}$$

■

Proposition 2.37 *1. Seja $M = \{M_n; n \geq 0\}$ uma $\{\mathcal{F}_n\}$-martingala. Então, para $m \geq n$, temos*

$$E\left[M_m|\mathcal{F}_n\right] = M_n.$$

2. $\{M_n; n \geq 0\}$ é submartingala se e só se $\{-M_n; n \geq 0\}$ é supermartingala.

3. Se $\{M_n; n \geq 0\}$ é martingala e φ é função convexa tal que $E\left[|\varphi(M_n)|\right] < \infty \quad \forall n \geq 0$, então $\{\varphi(M_n), n \geq 0\}$ é uma submartingala.

Property 3 is a consequence of the Jensen inequality and has as corollary: if $\{M_n; n \geq 0\}$ and $E\left[|M_n|^p\right] < \infty \quad \forall n \geq 0$ and some $p \geq 1$, then $\{|M_n|^p, n \geq 0\}$ is a submartingale.

Exercise 2.38 *Let $M = \{M_n; n \geq 0\}$ be a $\{\mathcal{F}_n\}$-martingale. Show that if $m \geq n$ then $E\left[M_m|\mathcal{F}_n\right] = M_n$.*

2.4 The Martingale Transform

Let $\{\mathcal{F}_n, n \geq 0\}$ be a filtration on the probability space $(\Omega, \mathcal{F}, P)$.

Definition 2.39 *A stochastic process $\{H_n, n \geq 1\}$ is said to be predictable if H_n is $\mathcal{F}_{n-1}$-measurable (i.e., if H_n is "known" at time $n-1$).*

Definition 2.40 *Given a $\{\mathcal{F}_n\}$-martingale $M = \{M_n; n \geq 0\}$ and a predictable process $\{H_n, n \geq 1\}$, the process $\{(H \cdot M)_n, n \geq 1\}$ defined as*

$$(H \cdot M)_n = M_0 + \sum_{j=1}^{n} H_j \Delta M_j$$

is called the martingale transform of M by $\{H_n, n \geq 1\}$.

The martingale transform of a predictable sequence is the discrete version of the stochastic integral, that is:

$$(H \cdot M)_n - M_0 = \sum_{j=1}^{n} H_j \Delta M_j \approx \int_0^n H_s dM_s.$$

Proposition 2.41 *If $M = \{M_n; n \geq 0\}$ is a martingale and $\{H_n, n \geq 0\}$ is a predictable process with bounded random variables, then the martingale transform $\{(H \cdot M)_n, n \geq 1\}$ is a martingale.*

Proof. 1. $(H \cdot M)_n$ is $\{\mathcal{F}_n\}$-measurable because $\sum_{j=1}^{n} H_j \Delta M_j$ is $\mathcal{F}_n$- measurable.

2. $(H \cdot M)_n$ is integrable, as the r.v. M_n are integrable and the r.v. H_n are bounded.

3. By the properties of conditional expectation,

$$\begin{aligned} E\left[(H \cdot M)_{n+1} - (H \cdot M)_n \,|\mathcal{F}_n\right] &= E\left[H_{n+1}\left(M_{n+1} - M_n\right)|\mathcal{F}_n\right] \\ &= H_{n+1} E\left[M_{n+1} - M_n|\mathcal{F}_n\right] \\ &= 0. \end{aligned}$$

■

Consider now the following gambling system H_n:

- The amount bet by player at move n is H_n;
- $\Delta M_n = M_n - M_{n-1}$ represents the gains at move n;
- M_n represents the accumulated fortune at instant n;
- $(H \cdot M)_n$ represents the accumulated fortune of the player if he uses the betting system $\{H_n, n \geq 1\}$.

If $\{M_n; n \geq 0\}$ is a martingale the game is said to be fair, and then $(H \cdot M)_n$ is also a martingale, that is, the game remains fair no matter the betting system used, as long as $\{H_n, n \geq 0\}$ satisfies the conditions of proposition 2.41.

Example 2.42 *(double-betting): Suppose that*

$$M_n = M_0 + Z_1 + \cdots + Z_n,$$

where $\{Z_n; n \geq 1\}$ are independent r.v. that represent the heads $(+1)$ or tails (-1) of a coin. Then, $P(Z_i = 1) = P(Z_i = -1) = \frac{1}{2}$. Suppose a player starts *by betting a single Euro and doubles its bet every time he loses (if the coin comes up tails (-1)) ending the game otherwise (if the coin comes up heads $(+1)$). So the betting system is given by*

$$\begin{aligned} &H_1 = 1, \\ &H_n = 2H_{n-1} \quad \textit{if } Z_{n-1} = -1, \\ &H_n = 0 \quad \textit{if } Z_{n-1} = +1. \end{aligned}$$

If the player loses k moves and wins at move $k+1$, he gets

$$(H \cdot M)_k = -1 - 2 - 4 - \cdots - 2^{k-1} + 2^k = 1.$$

This may seem like an always-win strategy, but pay attention to the fact that, for this to be true (with probability 1), unlimited funds and time are required (unbounded betting strategy). In fact, in this case proposition 2.41 does not apply because the variables H_n (of the gambling system) are not bounded.

Example 2.43 *(financial application) Let $S_n := \{S_n^0, S_n^1, n \geq 1\}$ be adapted processes that represent the prices of two financial assets. Let*

$$S_n^0 = (1+r)^n$$

be the price of a risk-free asset (bond), where r is the interest rate (the process S_n^0 é is deterministic). A portfolio is $\phi_n := \{\phi_n^0, \phi_n^1, n \geq 1\}$, that represents the number of units of each asset. The value of the portfolio at period n is given by

$$V_n = \phi_n^0 S_n^0 + \phi_n^1 S_n^1 = \phi_n \cdot S_n$$

The porfolio is said to be self-financed if, for any n,

$$V_n = V_0 + \sum_{j=1}^{n} \phi_j \Delta S_j.$$

This condition is equivalent to have, for any n,

$$\phi_n \cdot S_n = \phi_{n+1} \cdot S_n$$

Define the discounted prices as

$$\widetilde{S}_n = (1+r)^{-n} S_n = \left(1, (1+r)^{-n} S_n^1\right).$$

Thus,

$$\begin{aligned} \widetilde{V}_n &= (1+r)^{-n} V_n = \phi_n \cdot \widetilde{S}_n, \\ \phi_n \cdot \widetilde{S}_n &= \phi_{n+1} \cdot \widetilde{S}_n, \\ \widetilde{V}_n &= V_0 + \sum_{j=1}^{n} \phi_j \Delta \widetilde{S}_j \end{aligned}$$

The process $\widetilde{V}_n = \left(\phi_n^1 \cdot \widetilde{S}^1\right)_n$ *is the martingale transform of* $\left\{\widetilde{S}_n^1\right\}$ *by the predictable process* $\{\phi_n^1\}$*. If* $\left\{\widetilde{S}_n^1\right\}$ *is a martingale and if* $\{\phi_n^1\}$ *is a bounded sequence, then, by Proposition 2.41, the process* $\left\{\widetilde{V}_n\right\}$ *is also a martingale.*

Example 2.44 *(binomial model) A probability measure* Q *equivalent to* P *is called a risk-neutral probability measure if, in the probability space* $(\Omega, \mathcal{F}, Q)$*, the process* $\left\{\widetilde{S}_n^1\right\}$ *is a* $\{\mathcal{F}_n\}$*-martingale. On that case, if* $\{\phi_n^1\}$ *is bounded,* $\left\{\widetilde{V}_n\right\}$ *will also be a martingale, as seen in the previous example.*

In the binomial model, assume the r.v.

$$T_n = \frac{S_n}{S_{n-1}}$$

are independent and take the values $1+a$ *and* $1+b$ *with probabilities* p *and* $1-p$*, respectively, with* $a < r < b$*. Let us determine* p *(that is, the probability measure* Q*) in such way that* $\left\{\widetilde{S}_n^1\right\}$ *is a martingale.*

$$\begin{aligned} E\left[\widetilde{S}_{n+1}|\mathcal{F}_n\right] &= (1+r)^{-n-1} E\left[S_n T_{n+1}|\mathcal{F}_n\right] \\ &= \widetilde{S}_n (1+r)^{-1} E\left[T_{n+1}|\mathcal{F}_n\right] \\ &= \widetilde{S}_n (1+r)^{-1} E\left[T_{n+1}\right] \end{aligned}$$

Hence, $\left\{\widetilde{S}_n^1\right\}$ *is a martingale if* $E\left[T_{n+1}\right] = (1+r)$*, that is, if*

$$E\left[T_{n+1}\right] = p(1+a) + (1-p)(1+b) = 1+r$$

thus

$$p = \frac{b-r}{b-a}.$$

Consider now a $\{\mathcal{F}_N\}$*-measurable r.v.* H *that represents the payoff of a derivative on asset 1 that matures at time* N*. As an example, an European call option with exercise price* K *has a payoff* $H = (S_N - K)^+$*. The derivative is said to be replicable if there exists a self-financing portfolio such that*

$$V_N = H.$$

The price of the derivative will be the value of this portfolio. As $\left\{\widetilde{V}_n\right\}$ *is a martingale on the probability space,*

$$\begin{aligned} V_n = (1+r)^n \widetilde{V}_n &= (1+r)^n E_Q\left[\widetilde{V}_N|\mathcal{F}_n\right] \\ &= (1+r)^{-(N-n)} E_Q\left[H|\mathcal{F}_n\right] \end{aligned}$$

If $n = 0$*, then* $\mathcal{F}_0 = \{\Omega, \varnothing\}$ *and*

$$V_0 = (1+r)^{-N} E_Q[H].$$

2.5 Continuous-time Martingales

Continuous-time martingales are defined in a similar way as we defined discrete-time martingales, and the majority of properties still hold in this case.

Definition 2.45 *Consider the probability space* $(\Omega, \mathcal{F}, P)$*. A family of* σ*-algebras* $\{\mathcal{F}_t, t \geq 0\}$ *such that*

$$\mathcal{F}_s \subset \mathcal{F}_t, \quad 0 \leq s \leq t.$$

is called a filtration.

Let $\mathcal{F}_t^X$ be the σ-algebra generated by the process X on the interval $[0,t]$, i.e. $\mathcal{F}_t^X = \sigma(X_s, 0 \leq s \leq t)$. Then $\mathcal{F}_t^X$ may be interpreted as the information generated by the process X on the interval $[0,t]$. Claiming that $A \in \mathcal{F}_t^X$ means that it is possible to decide wether the event A has happened or not, by observing the trajectories of X in $[0,t]$.

Example 2.46 *If* $A = \{\omega : X(5) > 1\}$ *then* $A \in \mathcal{F}_5^X$ *but* $A \notin \mathcal{F}_4^X$*.*

Definition 2.47 *A stochastic process* $M = \{M_t; t \geq 0\}$ *is said to be a martingale with respect to the filtration* $\{\mathcal{F}_t, t \geq 0\}$ *if:*

1. *For all* $t \geq 0$*,* M_t *is a* $\mathcal{F}_t$*-measurable r.v. (i.e.,* M *is a stochastic process adapted to the filtration* $\{\mathcal{F}_t, t \geq 0\}$*).*
2. *For all* $t \geq 0$*,* $E[|M_t|] < \infty$*.*
3. *For all* $s \leq t$*, we have*
$$E[M_t|\mathcal{F}_s] = M_s.$$

Condition (3) is equivalent to $E[M_t - M_s|\mathcal{F}_s] = 0$. If $t \in [0,T]$ then, by the martingale property, $M_t = E[M_T|\mathcal{F}_t]$. Like in the discrete-time case, condition (3) implies $E[M_t] = E[M_0]$ for all t.

The definitions of supermartingale and submartingale are analogous to their respective discrete-time definitions.

We also have the following generalization of the Chebyshev inequality (analogous to the discrete-time version).

Theorem 2.48 *(Doob's maximal (or martingale) inequality): If* $M = \{M_t; t \geq 0\}$ *is a martingale with continuous trajectories, then, for all* $p \geq 1$, $T \geq 0$ *and* $\lambda > 0$,

$$P\left[\sup_{0 \leq t \leq T} |M_t| \geq \lambda\right] \leq \frac{1}{\lambda^p}\left[E\left|M_T\right|^p\right]$$

To a proof of the discrete version of this theorem (based on the optional stopping theorem), see [9] . To a more detailed analysis of martingales and its properties, it is recommended the reading of [2] and [5].

Chapter 3

Brownian motion

3.1 Definition

The name “Brownian motion” is given in honor of the botanist Robert Brown who was, in 1827, the first to observe, using a microscope, the erratic physical movement of pollen grains suspended in water droplets. This movement is provoked by molecular shocks and it is a physical example of what was later known as Brownian motion. In 1900, Louis Bachelier, in his thesis “Théorie de la spéculation” used the Brownian motion as a model for the evolution of financial asset prices. Five years later, Albert Einstein used this motion in one of it’s famous articles to confirm the existence and find the size and mass of atoms and molecules. The proof that the Brownian motion, as a stochastic process, exists and is soundly defined was done in 1923 by Norbert Wiener - which is why it is also called a Wiener process.

Definition 3.1 *A s.p.* $B = \{B_t; t \geq 0\}$ *is called a Brownian motion if it satisfies the following conditions:*

1. $B_0 = 0$.
2. *B has independent increments.*
3. *Se $s < t$, $B_t - B_s$ é is a r.v. with distribution $N(0, t-s)$.*
4. *The process B has continuous trajectories.*

The Brownian motion is a Gaussian process. In fact, the finite dimension distributions of B, i.e. the distribution of the vectors $(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$ is

Gaussian. From condition 3 of the definition, we may conclude that $B_t \sim N(0,t)$ and

$$E\left[B_t\right] = 0, \quad \forall t \geq 0, \tag{3.1}$$
$$E\left[B_t^2\right] = t, \quad \forall t \geq 0 \tag{3.2}$$

Proposition 3.2 *Let $B = \{B_t; t \geq 0\}$ be a Brownian motion. Then the covariance function of B is*

$$c_B\left(s,t\right) = E\left[B_s B_t\right] = \min\left(s,t\right). \tag{3.3}$$

Proof. If $s \leq t$, using the independence of increments of the Brownian motion and eq.(3.1),

$$\begin{aligned} E\left[B_s B_t\right] &= E\left[B_s\left(B_t - B_s\right) + B_s^2\right] \\ &= E\left[B_s\left(B_t - B_s\right)\right] + E\left[B_s^2\right] \\ &= E\left[B_s\right]E\left[B_t - B_s\right] + s = s, \end{aligned}$$

■

3.2 Main properties

Proposition 3.3 *A stochastic process that satisfies conditions 1,2 and 3. of the definition 3.1 has a version with continuous trajectories.*

Proof. As $(B_t - B_s) \sim N(0, t-s)$, it is possible to show that

$$E\left[\left(B_t - B_s\right)^{2k}\right] = \frac{(2k)!}{2^k \cdot k!}\left(t-s\right)^k. \tag{3.4}$$

To prove this result one may use integration by parts and the mathematical induction method in k (see [10]). For $k = 2$, we get

$$E\left[\left(B_t - B_s\right)^4\right] = 3\left(t-s\right)^2.$$

Therefore, by the Continuity Criterion of Kolmogorov (Theorem 2.18, there exists a version of B with continuous trajectories. ■

In order to define the Brownian motion we could just demand the first three conditions of Definition 3.1, and by the previous proposition, there would exist a version with continuous trajectories.

It may be shown that there exists a s.p. that satisfies conditions 1,2,3. To do so, one may use the Kolmogorov Existence Theorem (see [10]). Then, by the previous proposition, there exists a stochastic process that satisfies the 4 conditions given in the definition. Hence, the Brownian motion exists as a soundly defined mathematical object.

In the definition of Brownian motion, the probability space is arbitrary. However, it is possible to describe the structure of this space by considering the map:

$$\begin{aligned} \Omega &\to C\left([0,\infty),\mathbb{R}\right) \\ \omega &\to B_{\cdot}\left(\omega\right) \end{aligned}$$

that, for each element ω corresponds a continuous function that takes values in $\mathbb{R}$ (the trajectory). The probability space is the space of continuous functions $C\left([0,\infty),\mathbb{R}\right)$ equipped with the Borel σ-algebra $\mathcal{B}_C$ and the probability measure induced by the above map: $P \circ B^{-1}$. This probability is called the Wiener measure.

A corollary may be obtained from the Kolmogorov Continuity Criterion and the the formula

$$E\left[\left(B_t - B_s\right)^{2k}\right] = \frac{(2k)!}{2^k \cdot k!}\left(t-s\right)^k, \tag{3.5}$$

We get

$$\left|B_t\left(\omega\right) - B_s\left(\omega\right)\right| \leq G_\varepsilon\left(\omega\right)\left|t-s\right|^{\frac{1+\alpha}{p}-\varepsilon} \leq G_\varepsilon\left(\omega\right)\left|t-s\right|^{\frac{1}{2}-\varepsilon},$$

for all $\varepsilon > 0$ where $G_\varepsilon\left(\omega\right)$ is a r.v.. Therefore we conclude that the trajectories of the Brownian motion are Hölder continuous of order $\delta = \frac{1}{2} - \varepsilon$. Informally, for $\Delta t > 0$,

$$\left|B_{t+\Delta t} - B_t\right| \approx \left(\Delta t\right)^{\frac{1}{2}}.$$

On the other hand, we already know that

$$E\left[\left(B_{t+\Delta t} - B_t\right)^2\right] = \Delta t.$$

Considering the interval $[0,t]$ and partitions of it such that $0 = t_0 < t_1 < \cdots < t_n = t$, where $t_j = \frac{tj}{n}$, we may then heuristically deduce that:

- B has infinite total variation: $\sum_{k=1}^{n}\left|\Delta B_k\right| \approx n\left(\frac{t}{n}\right)^{1/2} \to \infty$ as $n \to \infty$.
- B has finite quadratic variation: $\sum_{k=1}^{n}\left|\Delta B_k\right|^2 = n\left(\frac{t}{n}\right) = t$.

We now present an important result on the irregularity of the trajectories of the Brownian motion.

Proposition 3.4 *The trajectories of the Brownian motion are not differentiable in any point (almost surely).*

We only need to show that, fora given point t, the derivative of the trajectory does not exist on that point. We have that

$$\frac{B_{t+\Delta t} - B_t}{\Delta t} \approx \frac{\sqrt{\Delta t} Z}{\Delta t} = \frac{Z}{\sqrt{\Delta t}},$$

where $Z \sim N(0,1)$. Then, this ratio tends to ∞ as $\Delta t \to 0$ in probability, because $P\left(\frac{Z}{\sqrt{\Delta t}} > K\right) \to 1$ for any K, when $\Delta t \to 0$. Hence, the derivative does not exist in point t.

Proposition 3.5 *(self-similarity) If $B = \{B_t; t \geq 0\}$ is a Brownian motion then, for all $a > 0$, the process $\left\{a^{-1/2} B_{at}; t \geq 0\right\}$ is also a Brownian motion.*

Exercise 3.6 *Prove proposition 3.5.*

3.3 Processes related to the Brownian motion

We now show some processes that are defined from the Brownian motion $B = \{B_t; t \geq 0\}$.

- Brownian motion with drift:

  $$Y_t = \mu t + \sigma B_t,$$

  where $\sigma > 0$ and μ are constant. Clearly, this is a Gaussian process with $E[Y_t] = \mu t$ and $cov(s,t) = \sigma^2 \min(s,t)$.

- Geometric Brownian motion (model proposed by Samuelson, and later by Black, Scholes and Merton to describe prices of financial assets):

  $$X_t = e^{\mu t + \sigma B_t},$$

  where $\sigma > 0$ and μ are constant. The distribution of X is lognormal, that is, $\ln(X_t)$ has a normal distribution.

- Brownian bridge:

  $$Z_t = B_t - t B_1, \quad t \in [0,1].$$

  Note that $Z_1 = Z_0 = 0$. This process is Gaussian with $E[Z_t] = 0$ and $cov(s,t) = E[Z_s Z_t] = \min(s,t) - st$.

Define the filtration generated by B

$$\mathcal{F}_t^B = \sigma\left\{B_s, s \leq t\right\}.$$

It is considered that $\mathcal{F}_t^B$ also contains the sets with null probability ($N \in \mathcal{F}_0$ if $P(N) = 0$). Some consequences of the inclusion of null probability sets in the filtration are given below:

1. Any version of an adapted process is still an adapted process.

2. The filtration is continuous from the right, i.e.

$$\bigcap_{s>t} \mathcal{F}_s^B = \mathcal{F}_t^B.$$

Example 3.7 *If B is a Brownian motion then the process $X_t = \sup\limits_{0\leq s\leq t} B_s$ is adapted to the filtration generated by B but the process $Y_t = \sup\limits_{0\leq s\leq t+1} B_s$ is not.*

Proposition 3.8 *If $B = \{B_t; t \geq 0\}$ is a Brownian motion and $\left\{\mathcal{F}_t^B, t \geq 0\right\}$ is the filtration generated by B, then the following processes are $\left\{\mathcal{F}_t^B, t \geq 0\right\}$-martingales:*

1. B_t.

2. $B_t^2 - t$.

3. $\exp\left(aB_t - \frac{a^2 t}{2}\right)$.

Proof. 1. Clearly B_t is $\mathcal{F}_t^B$-measurable and integrable. Also, as $B_t - B_s$ is independent from $\mathcal{F}_s^B$ (by the independence of the increments of B), we get

$$E\left[B_t - B_s|\mathcal{F}_s^B\right] = E\left[B_t - B_s\right] = 0.$$

2. The process $B_t^2 - t$ is $\mathcal{F}_t^B$-measurable and integrable. By the properties of B and the conditional expectation,

$$\begin{aligned} E\left[B_t^2 - t|\mathcal{F}_s^B\right] &= E\left[(B_t - B_s + B_s)^2|\mathcal{F}_s^B\right] - t \\ &= E\left[(B_t - B_s)^2\right] + 2B_s E\left[B_t - B_s|\mathcal{F}_s^B\right] + B_s^2 - t \\ &= t - s + B_s^2 - t = B_s^2 - s. \end{aligned}$$

■

Exercise 3.9 *Let $B = \{B_t; t \geq 0\}$ be a Brownian motion and $\left\{\mathcal{F}_t^B, t \geq 0\right\}$ the filtration generated by B. Show that the process $X_t = \exp\left(aB_t - \frac{a^2 t}{2}\right)$ is a martingale.*

3.4 The quadratic variation of the Brownian Motion

The results on the total and quadratic variation of the Brownian motion were deduced heuristically. We dedicate this section to prove rigorously these results.
Let us fix an interval $[0, t]$ and a partition π of this interval, with

$$0 = t_0 < t_1 < \cdots < t_n = t.$$

The norm of the partition is defined as

$$|\pi| = \max_k \Delta t_k,$$

where $\Delta t_k = t_k - t_{k-1}$ and let $\Delta B_k = B_{t_k} - B_{t_{k-1}}$.

Proposition 3.10 *The Brownian motion B has finite quadratic variation in the interval $[0, t]$ (and equal to t), in the sense that*

$$E\left[\left(\sum_{k=1}^{n} (\Delta B_k)^2 - t\right)^2\right] \longrightarrow 0,$$

when $|\pi| \to 0$.

Proof. Using the independence of the increments, the fact that $E\left[(\Delta B_k)^2\right] = \Delta t_k$ and the formula $E\left[(B_t - B_s)^{2j}\right] = \frac{(2j)!}{2^j \cdot j!}(t-s)^j$, we have

$$E\left[\left(\sum_{k=1}^{n} (\Delta B_k)^2 - t\right)^2\right] = E\left[\sum_{k=1}^{n} \left[(\Delta B_k)^2 - \Delta t_k\right]^2\right]$$

$$\sum_{j=1}^{n} \left[3(\Delta t_k)^2 - 2(\Delta t_k)^2 + (\Delta t_k)^2\right]$$

$$= 2\sum_{j=1}^{n} (\Delta t_k)^2 \leq 2t\,|\pi| \underset{|\pi| \to 0}{\to} 0.$$

■

Proposition 3.11 *The Brownian motion B has infinite total variation in the interval $[0, t]$, in the sense that $V = \sup_{\pi} \sum_{k=1}^{n} |\Delta B_k| = \infty$ with probability 1.*

Proof. Using the continuity of the trajectories of the Brownian motion, we have that

$$\sum_{k=1}^{n} (\Delta B_k)^2 \leq \sup_k |\Delta B_k| \sum_{k=1}^{n} |\Delta B_k| \leq V \sup_k |\Delta B_k| \underset{|\pi| \to 0}{\to} 0,$$

if $V < \infty$. But this contradicts the fact that $\sum_{k=1}^{n} (\Delta B_k)^2$ converges in mean squared to t. Therefore $V = \infty$. ■

For a more detailed analysis on the Brownian motion, it is recommended the reading of [5].

Chapter 4

Itô integral

4.1 Motivation

Let $B = \{B_t; t \geq 0\}$ be a Brownian motion and consider a "stochastic" differential equation of the type

$$\frac{dX}{dt} = b(t, X_t) + \sigma(t, X_t) \text{``}\frac{dB_t}{dt}\text{''},$$

where "$\frac{dB_t}{dt}$" is stochastic noise. This process does not exist in the classic sense, because B is not differentiable. We may express the equation above in the integral form

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \text{``}\int_0^t \sigma(s, X_s) dB''_s$$

The problem now is how to define $\int_0^t \sigma(s, X_s) dB_s$, or, in a broader way, to define stochastic integrals of the form

$$\int_0^T u_s \mathrm{d}B_s,$$

where B is a Brownian motion and u is an adequate stochastic process.

A strategy that could be followed to define this type of integrals would be to consider them as Riemann-Stieltjes integrals. Let us see how to define an integral of this type. Consider a sequence of partitions of $[0, T]$ and a sequence of interior points in those partitions:

$$\begin{aligned} &\tau_n\colon\ 0 = t_0^n < t_1^n < t_2^n < \cdots < t_{k(n)}^n = T \\ &s_n\colon\ t_i^n \leq s_i^n \leq t_{i+1}^n, \quad i = 0, \ldots, k(n) - 1, \end{aligned}$$

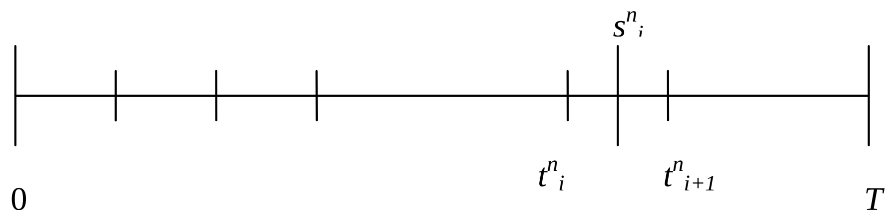

Figure 4.1: Partition of $[0,T]$

such that $\lim_{n\to\infty}\sup_i\left(t^n_{i+1}-t^n_i\right)=0$.

The Riemann-Stieltjes (R-S) integral is defined as the limit of the Riemann sums:

$$\int_0^T f dg := \lim_{n\to\infty}\sum_{i=0}^{n-1} f\left(s^n_i\right)\Delta g_i,$$

where $\Delta g_i := g(t^n_{i+1}) - g(t^n_i)$, if the limit exists and is independent from the sequences τ_n and s_n.

- The Riemann-Stieltjes (R-S) integral $\int_0^T f dg$ exists if f is continuous and g has limited total variation, i.e.

$$\sup_{\tau_n}\sum_i |\Delta g_i| < \infty.$$

- If f is continuous g is of class C^1 then the (R-S) integral $\int_0^T f dg$ exists and

$$\int_0^T f dg := \int_0^T f\left(t\right) g'(t) dt,$$

- In the case of the Brownian motion B, clearly the derivative $B'(t)$ does not exist, so we cannot define the trajectory integral

$$\int_0^T u_t\left(\omega\right) dB_t\left(\omega\right) \overset{\times}{\neq} \int_0^T u_t\left(\omega\right) B'_t\left(\omega\right) dt.$$

In general, we know that the Brownian motion does not have bounded total variation, so we may not define the (R-S) integral $\int_0^T u_t\left(\omega\right) dB_t\left(\omega\right)$. However, if u has C^1 trajectories, using integration by parts, the trajectory (R-S) integral exists and

$$\int_0^T u_t\left(\omega\right) dB_t\left(\omega\right) = u_T\left(\omega\right) B_T\left(\omega\right) - \int_0^T u'_t\left(\omega\right) B_t\left(\omega\right) dt.$$

Still, a problem withstands. For example, the integral $\int_0^T B_t(\omega)\, dB_t(\omega)$ does not exist as an R-S integral. It is useful to consider processes that are more irregular than those with class C^1 trajectories, so how may we define the stochastic integral for these processes? We should abandon the R-S strategy and follow a new reasoning: define the stochastic integral $\int_0^T u_t dB_t$ through a probabilistic approach.

Consider a Brownian motion $B = \{B_t; t \geq 0\}$ and the filtration $\{\mathcal{F}_t, t \geq 0\}$ generated by it.

Definition 4.1 *Consideraremos processos u da classe $L^2_{a,T}$, que se define como a classe de processos estocásticos $u = \{u_t,\ t \in [0,T]\}$, tais que:*

1. *u is measurable and adapted, i.e: u_t is $\mathcal{F}_t$-measurable for all $t \in [0,T]$, and the map $(t,\omega) \to u_t(\omega)$, defined in $[0,T] \times \Omega$ is measurable with respect to the σ-algebra $\mathcal{B}_{[0,T]} \times \mathcal{F}_T$.*

2. $E\left[\int_0^T u_t^2 dt\right] < \infty$.

- Condition 2. allows one to show that u as a function of the variables t and ω belongs to the space $L^2([0,T] \times \Omega)$ and that (using Fubini's theorem)

$$E\left[\int_0^T u_t^2 dt\right] = \int_0^T E\left[u_t^2\right] dt = \int_{[0,T]\times\Omega} u_t^2(\omega)\, dtP(d\omega).$$

4.2 The stochastic integral of simple processes

The strategy to define the stochastic integral consists in defining $\int_0^T u_t dB_t$ for $u \in L^2_{a,T}$ as the mean-squared limit (limit in $L^2(\Omega)$) of integrals of simple process.

Definition 4.2 *A stochastic process u is said to be a simple process if*

$$u_t = \sum_{j=1}^{n} \phi_j \mathbf{1}_{(t_{j-1},t_j]}(t), \tag{4.1}$$

where $0 = t_0 < t_1 < \cdots < t_n = T$, and the r.v. ϕ_j are square-integrable ($E\left[\phi_j^2\right] < \infty$) and $\mathcal{F}_{t_{j-1}}$-measurable.

From now on, we shall denote the class of all simple processes by $\mathcal{S}$.

Definition 4.3 *If u is a simple process of the form(4.1) ($u \in \mathcal{S}$) then we define the stochastic integral of u with respect to the Brownian motion B as*

$$\int_0^T u_t dB_t := \sum_{j=1}^n \phi_j \left(B_{t_j} - B_{t_{j-1}} \right).$$

Example 4.4 *Consider the simple process*

$$u_t = \sum_{j=1}^n B_{t_{j-1}} \mathbf{1}_{(t_{j-1}, t_j]} (t).$$

Then

$$\int_0^T u_t dB_t = \sum_{j=1}^n B_{t_{j-1}} \left(B_{t_j} - B_{t_{j-1}} \right).$$

It is clear that, from the independent increments of B, we have

$$\begin{aligned} E\left[\int_0^T u_t dB_t \right] &= \sum_{j=1}^n E\left[B_{t_{j-1}} \left(B_{t_j} - B_{t_{j-1}} \right) \right] \\ &= \sum_{j=1}^n E\left[B_{t_{j-1}} \right] E\left[B_{t_j} - B_{t_{j-1}} \right] = 0. \end{aligned}$$

Proposition 4.5 *(Proprerty of isometry). Let $u \in \mathcal{S}$ be a simple process. Then the following isometry property holds:*

$$E\left[\left(\int_0^T u_t dB_t \right)^2 \right] = E\left[\int_0^T u_t^2 dt \right]. \tag{4.2}$$

Proof. Taking $\Delta B_j := B_{t_j} - B_{t_{j-1}}$,

$$\begin{aligned} E\left[\left(\int_0^T u_t dB_t \right)^2 \right] &= E\left[\left(\sum_{j=1}^n \phi_j \Delta B_j \right)^2 \right] \\ &= \sum_{j=1}^n E\left[\phi_j^2 \left(\Delta B_j \right)^2 \right] + 2 \sum_{i<j}^n E\left[\phi_i \phi_j \Delta B_i \Delta B_j \right]. \end{aligned}$$

Note that as $\phi_i \phi_j \Delta B_i$ is $\mathcal{F}_{j-1}$-measurable and ΔB_j is independent from $\mathcal{F}_{j-1}$, we have

$$\sum_{i<j}^n E\left[\phi_i \phi_j \Delta B_i \Delta B_j \right] = \sum_{i<j}^n E\left[\phi_i \phi_j \Delta B_i \right] E\left[\Delta B_j \right] = 0.$$

On the other hand, as ϕ_j^2 is $\mathcal{F}_{j-1}$-measurable and ΔB_j is independent from $\mathcal{F}_{j-1}$,

$$\begin{aligned}\sum_{j=1}^{n} E\left[\phi_j^2 (\Delta B_j)^2\right] &= \sum_{j=1}^{n} E\left[\phi_j^2\right] E\left[(\Delta B_j)^2\right] \\ &= \sum_{j=1}^{n} E\left[\phi_j^2\right](t_j - t_{j-1}) = \\ &= E\left[\int_0^T u_t^2 dt\right].\end{aligned}$$

■

Proposition 4.6 *Let $u \in \mathcal{S}$.*

- *1. Linearity: If u,$v \in \mathcal{S}$:*

$$\int_0^T (au_t + bv_t)\, dB_t = a\int_0^T u_t dB_t + b\int_0^T v_t dB_t. \tag{4.3}$$

2. Null expected value:

$$E\left[\int_0^T u_t dB_t\right] = 0. \tag{4.4}$$

Exercise 4.7 *Prove properties 1. and 2. from the previous proposition.*

4.3 Itô integral for adapted processes

The following lemma is fundamental to define stochastic integral of adapted processes.

Lema 4.8 *If $u \in L^2_{a,T}$ then there exists a sequence of simple processes $\left\{u^{(n)}\right\} \in \mathcal{S}$ such that*

$$\lim_{n\to\infty} E\left[\int_0^T \left|u_t - u_t^{(n)}\right|^2 dt\right] = 0. \tag{4.5}$$

Proof. 1. Suppose u is a mean-squared continuous process, that is:

$$\lim_{s\to t} E\left[|u_t - u_s|^2\right] = 0.$$

Define $t_j^n := \frac{j}{n}T$ and

$$u_t^n = \sum_{j=1}^{n} u_{t_{j-1}^n} \mathbf{1}_{\left(t_{j-1}^n, t_j^n\right]}(t). \tag{4.6}$$

By applying Fubini's theorem, we get

$$\begin{aligned} E\left[\int_0^T \left|u_t - u_t^{(n)}\right|^2 dt\right] &= \left[\int_0^T E\left[\left|u_t - u_t^{(n)}\right|^2\right] dt\right] \\ &= \sum_{j=1}^{n} \int_{t_{j-1}^n}^{t_j^n} E\left[\left|u_{t_{j-1}^n} - u_t\right|^2 dt\right] \\ &\leq T \sup_{|t-s|\leq \frac{T}{n}} E\left[\left|u_s - u_t\right|^2\right] \underset{n\to\infty}{\to} 0. \end{aligned}$$

Step 2. Suppose now that $u \in L_{a,T}^2$ and consider the sequence of processes $\left\{v^{(n)}\right\}$ defined by

$$v_t^n = n \int_{t-\frac{1}{n}}^{t} u_s ds.$$

These processes are mean-squared continuous (they even have continuous trajectories) and belong to the class $L_{a,T}^2$. On the other hand,

$$\lim_{n\to\infty} E\left[\int_0^T \left|u_t - v_t^{(n)}\right|^2 dt\right] = 0,$$

because

$$\lim_{n\to\infty} \int_0^T \left|u_t(\omega) - v_t^{(n)}(\omega)\right|^2 dt = 0.$$

and we may apply the Dominated Convergence Theorem in the space $[0,T] \times \Omega$, since we get, using Cauchy-Schwarz inequality and by changing the integration order,

$$\begin{aligned} E\left[\int_0^T \left|v_t^{(n)}\right|^2 dt\right] &= E\left[n^2 \int_0^T \left|\int_{t-\frac{1}{n}}^{t} u_s ds\right|^2 dt\right] \\ &\leq nE\left[\int_0^T \left(\int_{t-\frac{1}{n}}^{t} u_s^2 ds\right) dt\right] \\ &= nE\left[\int_0^T u_s^2 \left(\int_s^{s+1/n} dt\right) ds\right] \\ &= E\left[\int_0^T u_s^2 ds\right]. \end{aligned}$$

■

Definition 4.9 *The stochastic integral (or Itô integral) of the process* $u \in L^2_{a,T}$ *is defined as the limit (in* $L^2(\Omega)$*)*

$$\int_0^T u_t dB_t = \lim_{n\to\infty} (L^2) \int_0^T u_t^{(n)} dB_t, \tag{4.7}$$

where $\{u^{(n)}\}$ *is a sequence of simple processes that verifies (4.5).*

Note that the limit (4.7) exists since, due to the isometry property for simple processes, the sequence $\left\{\int_0^T u_t^{(n)} dB_t\right\}$ is a Cauchy sequence in $L^2(\Omega)$ and therefore is convergent.

Proof.

$$E\left[\left(\int_0^T u_t^{(n)} dB_t - \int_0^T u_t^{(m)} dB_t\right)^2\right] = E\left[\int_0^T \left(u_t^{(n)} - u_t^{(m)}\right)^2 dt\right]$$
$$\leq 2E\left[\int_0^T \left(u_t^{(n)} - u_t\right)^2 dt\right] + 2E\left[\int_0^T \left(u_t - u_t^{(m)}\right)^2 dt\right] \overset{n,m\to\infty}{\to} 0.$$

■

Proposition 4.10 *Let* $u \in L^2_{a,T}$. *Then, the following properties hold*

- 1. Isometry:

$$E\left[\left(\int_0^T u_t dB_t\right)^2\right] = E\left[\int_0^T u_t^2 dt\right]. \tag{4.8}$$

2. Null expected value:

$$E\left[\int_0^T u_t dB_t\right] = 0 \tag{4.9}$$

3. Linearity:

$$\int_0^T (au_t + bv_t) dB_t = a\int_0^T u_t dB_t + b\int_0^T v_t dB_t. \tag{4.10}$$

Proof. These properties are easily shown for simple processes $u \in \mathcal{S}$. After doing so, the broader case where $u \in L^2_{a,T}$ is proven by considering the process as a limit of a sequence simple processes. ■

Example 4.11 *Let us see that*

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

As the process $u_t = B_t$ is mean-squared continuous, we consider the sequence of approaching simple processes (4.6), i.e.

$$u_t^n = \sum_{j=1}^{n} B_{t_{j-1}^n} \mathbf{1}_{\left(t_{j-1}^n, t_j^n\right]}(t),$$

where $t_j^n := \frac{j}{n}T$.

$$\begin{aligned}
\int_0^T B_t dB_t &= \lim_{n\to\infty} (L^2) \int_0^T u_t^{(n)} dB_t = \\
&= \lim_{n\to\infty} (L^2) \sum_{j=1}^{n} B_{t_{j-1}^n} \left(B_{t_j^n} - B_{t_{j-1}^n}\right) \\
&= \lim_{n\to\infty} (L^2) \frac{1}{2} \sum_{j=1}^{n} \left[\left(B_{t_j^n}^2 - B_{t_{j-1}^n}^2\right) - \left(B_{t_j^n} - B_{t_{j-1}^n}\right)^2\right] \\
&= \frac{1}{2}\left(B_T^2 - T\right),
\end{aligned}$$

we use the fact that $E\left[\left(\sum_{j=1}^n \left(\Delta B_{t_j^n}\right)^2 - T\right)^2\right] \to 0$ and $\frac{1}{2}\sum_{j=1}^n \left(B_{t_j^n}^2 - B_{t_{j-1}^n}^2\right) = \frac{1}{2}B_T^2$.

4.4 Undefined stochastic integrals

Consider a stochastic process $u \in L_{a,T}^2$. Then, for any $t \in [0,T]$, the process $u\mathbf{1}_{[0,t]}$ is also in $L_{a,T}^2$, hence we may define the undefined stochastic integral:

$$\int_0^t u_s \mathrm{d}B_s := \int_0^T u_s \mathbf{1}_{[0,t]}(s)\, \mathrm{d}B_s.$$

The stochastic process $\left\{\int_0^t u_s \mathrm{d}B_s, 0 \le t \le T\right\}$ is the undefined stochastic integral of u with respect to B.

Proposition 4.12 *Main properties of the undefined stochastic integral:*

1. Additivity: For $a \leq b \leq c$, we have:
$$\int_a^b u_s \mathrm{d}B_s + \int_b^c u_s \mathrm{d}B_s = \int_a^c u_s \mathrm{d}B_s.$$

2. Factorization: If $a < b$ and $A \in \mathcal{F}_a$, then:
$$\int_a^b \mathbf{1}_A u_s \mathrm{d}B_s = \mathbf{1}_A \int_a^b u_s \mathrm{d}B_s.$$
This property still holds when $\mathbf{1}_A$ is substituted by any bounded and $\mathcal{F}_a$-measurable random variable.

3. Martingale property: If $u \in L^2_{a,T}$ Then the process $M_t = \int_0^t u_s \mathrm{d}B_s$ is a martingale with respect to the filtration $\mathcal{F}_t$.

4. Continuity: If $u \in L^2_{a,T}$ then the process $M_t = \int_0^t u_s \mathrm{d}B_s$ has a version with continuous trajectories.

5. Maximal inequality for the undefined stochastic integral: If $u \in L^2_{a,T}$ e $M_t = \int_0^t u_s \mathrm{d}B_s$, them, for any $\lambda > 0$,
$$P\left[\sup_{0 \leq t \leq T} |M_t| > \lambda\right] \leq \frac{1}{\lambda^2} E\left[\int_0^T u_t^2 \mathrm{d}t\right].$$

Proof. 3: Let $u^{(n)}$ be a sequence of simple processes such that
$$\lim_{n \to \infty} E\left[\int_0^T \left|u_t - u_t^{(n)}\right|^2 dt\right] = 0.$$
Let $M_n(t) = \int_0^t u_s^{(n)} \mathrm{d}B_s$ and let ϕ_j be the value of $u^{(n)}$ on the interval $(t_{j-1}, t_j]$, where $j = 1, \ldots, n$. If $s \leq t_k \leq t_{m-1} \leq t$, then:
$$\begin{aligned} &E\left[M_n(t) - M_n(s)\,|\mathcal{F}_s\right] \\ &= E\left[\phi_k\left(B_{t_k} - B_s\right) + \sum_{j=k+1}^{m-1} \phi_j \Delta B_j + \phi_m\left(B_t - B_{t_{m-1}}\right)|\mathcal{F}_s\right], \end{aligned}$$
hence, by the properties of conditional expectation,
$$\begin{aligned} &= E\left[\phi_k\left(B_{t_k} - B_s\right)|\mathcal{F}_s\right] + \sum_{j=k+1}^{m-1} E\left[E\left[\phi_j \Delta B_j|\mathcal{F}_{j-1}\right]|\mathcal{F}_s\right] + \\ &+ E\left[E\left[\phi_m\left(B_t - B_{t_{m-1}}\right)|\mathcal{F}_{t_{m-1}}\right]|\mathcal{F}_s\right]. \end{aligned}$$

$$= \phi_k E\left[B_{t_k} - B_s|\mathcal{F}_s\right] + \sum_{j=k+1}^{m-1} E\left[\phi_j E\left[\Delta B_j|\mathcal{F}_{j-1}\right]|\mathcal{F}_s\right] +$$
$$+ E\left[\phi_m E\left[B_t - B_{t_{m-1}}|\mathcal{F}_{t_{m-1}}\right]|\mathcal{F}_s\right]$$

using the independence of the Brownian motion increments we get:

$$E\left[M_n(t) - M_n(s)|\mathcal{F}_s\right] = 0.$$

As mean-square convergence implies mean-square convergence of the conditional expectations,

$$E\left[M(t) - M(s)|\mathcal{F}_s\right] = 0$$

and therefore the stochastic integral is a martingale.

4: Using the same notation as in the previous proof, $M_n(t)$ is clearly a process with continuous trajectories, because it is the stochastic integral of a simple process. Then, by Doob's maximal inequality applied to $M_n - M_m$, with $p = 2$, we have:

$$P\left[\sup_{0\le t\le T} |M_n(t) - M_m(t)| > \lambda\right] \le \frac{1}{\lambda^2} E\left[|M_n(T) - M_m(T)|^2\right]$$
$$= \frac{1}{\lambda^2} E\left[\left(\int_0^T \left(u_t^{(n)} - u_t^{(m)}\right) dB_t\right)^2\right]$$
$$= \frac{1}{\lambda^2} E\left[\int_0^T \left|u_t^{(n)} - u^{(m)}\right|^2 dt\right]. \overset{n,m\to\infty}{\longrightarrow} 0,$$

where we used the Itô isometry. We may then choose a increasing sequence of positive integers n_k, $k = 1, 2, \ldots$, such that

$$P\left[\sup_{0\le t\le T} \left|M_{n_{k+1}}(t) - M_{n_k}(t)\right| > 2^{-k}\right] \le 2^{-k}.$$

The events

$$A_k := \left\{\sup_{0\le t\le T} \left|M_{n_{k+1}}(t) - M_{n_k}(t)\right| > 2^{-k}\right\}$$

should then satisfy:

$$\sum_{k=1}^{\infty} P(A_k) < \infty.$$

Hence, by the Borel-Cantelli lemma, $P(\limsup_k A_k) = 0$ or

$$P\left[\sup_{0\le t\le T} \left|M_{n_{k+1}}(t) - M_{n_k}(t)\right| > 2^{-k} \text{ for an infinite \# of } k\right] = 0.$$

Therefore, for almost every $\omega \in \Omega$, there exists $k_1(\omega)$ such that

$$\sup_{0\le t\le T} \left| M_{n_{k+1}}(t) - M_{n_k}(t) \right| \le 2^{-k} \text{ for } k \ge k_1(\omega).$$

Consequently, $M_{n_k}(t,\omega)$ is uniformly convergent inside $[0,T]$ almost surely and hence the limit, which we denote by $J_t(\omega)$, is a continuous function on the variable t. Finally, as $M_{n_k}(t.\cdot) \to M_t(\cdot)$ in mean-square (or in $L^2(P)$) for all t, then we must have

$$M_t = J_t \quad \text{q.c. e para todo o } t \in [0,T].$$

and so we conclude that the undefined stochastic integral has a continuous version. ■

Exercise 4.13 *Prove property 1 of the previous proposition, that is,*

$$\int_a^b u_s \mathrm{d}B_s + \int_b^c u_s \mathrm{d}B_s = \int_a^c u_s \mathrm{d}B_s.$$

Proposition 4.14 *(quadratic variation of the undefined stochastic integral) Let $u \in L^2_{a,T}$. then*

$$\sum_{j=1}^{n} \left(\int_{t_{j-1}}^{t_j} u_s dB_s \right)^2 \overset{L^1(\Omega)}{\longrightarrow} \int_0^t u_s^2 ds,$$

when $n \to \infty$ and with $t_j := \frac{jt}{n}$.

4.5 Extensions of the stochastic integral

When defining the stochastic integral, one may substitute $\{\mathcal{F}_t\}$ (filtration generated by the Brownian motion) by a bigger filtration $\mathcal{H}_t$ such that the Brownian motion B_t is a $\mathcal{H}_t$-martingale.

One may also substitute condition 2) $E\left[\int_0^T u_t^2 dt\right] < \infty$ in the definition of $L^2_{a,T}$ by the (weaker) condition:

$$2')P\left[\int_0^T u_t^2 dt < \infty\right] = 1. \tag{4.11}$$

Let $L_{a,T}$ be the space of processes that satisfy both the condition 1 of the definition of $L^2_{a,T}$ (i.e., u is measurable and adapted) and the condition 2'). The stochastic integral may be defined for processes $u \in L_{a,T}$ but, in this case, the stochastic integral may neither have null expected value nor satisfy the Itô isometry.

Exercise 4.15 *Prove directly, using the definition of stochastic integral, that*

$$\int_0^t s dB_s = tB_t - \int_0^t B_s ds. \tag{4.12}$$

Suggestion: Note that

$$\sum_j \Delta (s_j B_j) = \sum_j s_j \Delta B_j + \sum_j B_{j+1} \Delta s_j. \tag{4.13}$$

Exercise 4.16 *Consider a deterministic function g such that $\int_0^T g(s)^2 ds < \infty$. Show that the stochastic integral $\int_0^T g(s) dB_s$ is a Gaussian random variable and determine its mean and variance.*

For an elementary introduction of the stochastic integral it is recommended the reading of [7]. For a more detailed discussion, reading [5], [9], [10] and [11] are recommended.

Chapter 5

Itô formula

5.1 One-dimensional Itô formula

Let $\Delta B = B_{t+\Delta t} - B_t$. We've seent that

$$E\left[(\Delta B)^2\right] = \Delta t,$$

and using the formula $E\left[(B_t - B_s)^{2k}\right] = \frac{(2k)!}{2^k \cdot k!}(t-s)^k$, we have that

$$\begin{aligned} Var\left[(\Delta B)^2\right] &= E\left[(\Delta B)^4\right] - \left(E\left[(\Delta B)^2\right]\right)^2 \\ &= 3(\Delta t)^2 - (\Delta t)^2 = 2(\Delta t)^2 . \end{aligned}$$

Hence, if Δt is small, the variance of $(\Delta B)^2$ is insignificant when compared with it's expected value. Therefore, when $\Delta t \to 0$ or "$\Delta t = dt$", informally we may conclude:

$$(dB_t)^2 = dt \tag{5.1}$$

The equality (5.1) is the base of the Itô formula (or Itô's Lemma) that we'll discuss throughout this chapter. Itô's formula is, essentially, a stochastic version of the chain rule. Consider the following (equivalent) equalities:

$$\begin{aligned} \int_0^t B_s dB_s &= \frac{1}{2}B_t^2 - \frac{1}{2}t \\ B_t^2 &= 2\int_0^t B_s dB_s + t \\ d\left(B_t^2\right) &= 2B_t dB_t + dt \end{aligned}$$

The last expression represents the Taylor expansion of B_t^2 as a function of B_t and t, with the convention $(dB_t)^2 = dt$ motivated by eq.(5.1).

If f is a function of class C^2, Itô's formula will show we have the following representation of $f(B_t)$:

$$\begin{aligned} f(B_t) &= \text{undefined stoch. integral} + \text{process with differentiable trajectories} \\ &= \text{Itô's process} \end{aligned}$$

Define $L^1_{a,T}$ as the space of processes v such that

1) v is an adapted and measurable process;

2") $P\left[\int_0^T |v_t|\, dt < \infty\right] = 1$.

Definition 5.1 *A continuous and adapted process $X = \{X_t, 0 \le t \le T\}$ is called an Itô's process if it satisfies:*

$$X_t = X_0 + \int_0^t u_s dB_s + \int_0^t v_s ds, \tag{5.2}$$

where $u \in L_{a,T}$ and $v \in L^1_{a,T}$.

Theorem 5.2 *(one-dimensional Itô's Formula) Let $X = \{X_t, 0 \le t \le T\}$ be a Itô process of the form (5.2). Let $f(t,x)$ be a function of class $C^{1,2}$. Then the process $Y_t = f(t, X_t)$ is an Itô process and has the representation*

$$\begin{aligned} f(t, X_t) = f(0, X_0) &+ \int_0^t \frac{\partial f}{\partial t}(s, X_s)\, ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, u_s dB_s \\ &+ \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, v_s ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s)\, u_s^2 ds. \end{aligned}$$

In the differential form, Itô's formula may be written as:

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

where $(dX_t)^2$ is computed using the product table:

$\times$	dB_t	dt
dB_t	dt	0
dt	0	0

Itô's formula for $f(t,x)$ and $X_t = B_t$ that is $Y_t = f(t, B_t)$, is

$$f(t,B_t) = f(0,0) + \int_0^t \frac{\partial f}{\partial t}(s,B_s)\,ds + \int_0^t \frac{\partial f}{\partial x}(s,B_s)\,dB_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,B_s)\,ds.$$

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

Itô's formula when $f(x)$ and $X_t = B_t$, that is $Y_t = f(B_t)$, is simply

$$df(B_t) = \frac{\partial f}{\partial x}(B_t)\,dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(B_t)\,dt.$$

5.2 Multidimensional Itô's formula

Suppose $B_t := (B_t^1, B_t^2, \ldots, B_t^m)$ is a Brownian motion of dimension m, that is, the components B_t^k, $k = 1, ..., m$ are independent one-dimensional Brownian motions. Consider an Itô process of dimension n, defined by

$$\begin{aligned} X_t^1 &= X_0^1 + \int_0^t u_s^{11} dB_s^1 + \cdots + \int_0^t u_s^{1m} dB_s^m + \int_0^t v_s^1 ds, \\ X_t^2 &= X_0^2 + \int_0^t u_s^{21} dB_s^1 + \cdots + \int_0^t u_s^{2m} dB_s^m + \int_0^t v_s^2 ds, \\ &\vdots \\ X_t^n &= X_0^n + \int_0^t u_s^{n1} dB_s^1 + \cdots + \int_0^t u_s^{nm} dB_s^m + \int_0^t v_s^n ds. \end{aligned}$$

In differential notation,

$$dX_t^i = \sum_{j=1}^m u_t^{ij} dB_t^j + v_t^i dt,$$

with $i = 1, 2, \ldots, n$. Or, in a more compact way,

$$dX_t = u_t dB_t + v_t dt,$$

where v_t is a n-dimensional process, u_t is a $n \times m$ matrix of processes. We assume the components of u belong to $L_{a,T}$ and the components of v belong to $L^1_{a,T}$.

Theorem 5.3 *(multidimensional Itô's formula) If $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}^p$ is of class $C^{1,2}$ then $Y_t = f(t, X_t)$ is an Itô process and we have the following Itô formula:*

$$\begin{aligned} dY_t^k &= \frac{\partial f_k}{\partial t}(t, X_t)\,dt + \sum_{i=1}^{n} \frac{\partial f_k}{\partial x_i}(t, X_t)\,dX_t^i \\ &+ \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f_k}{\partial x_i \partial x_j}(t, X_t)\,dX_t^i dX_t^j. \end{aligned}$$

The product of differentials $dX_t^i dX_t^j$ is computed according to the rules:

$$\begin{aligned} dB_t^i dB_t^j &= \begin{cases} 0 & \text{se } i \neq j \\ dt & \text{se } i = j \end{cases}, \\ dB_t^i dt &= 0, \\ (dt)^2 &= 0. \end{aligned}$$

In the particular case where B_t is a n-dimensional Brownian motion and $f : \mathbb{R}^n \to \mathbb{R}$ is of class C^2 with $Y_t = f(B_t)$ then Itô's formula is

$$f(B_t) = f(B_0) + \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}(B_t)\,dB_s^i + \frac{1}{2}\int_0^t \left(\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(B_t) \right) ds$$

Example 5.4 *(integration by parts formula) If X_t^1 and X_t^2 are Itô processes and $Y_t = X_t^1 X_t^2$, then by Itô's formula with $f(x) = f(x_1, x_2) = x_1 x_2$, we get*

$$d\left(X_t^1 X_t^2\right) = X_t^2 dX_t^1 + X_t^1 dX_t^2 + dX_t^1 dX_t^2.$$

That is:

$$X_t^1 X_t^2 = X_0^1 X_0^2 + \int_0^t X_s^2 dX_s^1 + \int_0^t X_s^1 dX_s^2 + \int_0^t dX_s^1 dX_s^2.$$

Example 5.5 *Consider the process*

$$Y_t = \left(B_t^1\right)^2 + \left(B_t^2\right)^2 + \cdots + \left(B_t^n\right)^2.$$

Represent this process in terms of stochastic integrals with respect to the n-dimensional Brownian motion. By the multidimensional Itô's formla, with $f(x) = f(x_1, x_2, \ldots, x_n) = x_1^2 + \cdots + x_n^2$ we get

$$\begin{aligned} dY_t &= 2B_t^1 dB_t^1 + \cdots + 2B_t^n dB_t^n \\ &+ ndt, \end{aligned}$$

that is,

$$Y_t = 2\int_0^t B_s^1 dB_s^1 + \cdots + 2\int_0^t B_s^n dB_s^n + nt.$$

Exercise 5.6 *Let $B_t := (B_t^1, B_t^2)$ be a 2-dimensional Brownian motion. Represent the process*

$$Y_t = \left(B_t^1 t, \left(B_t^2\right)^2 - B_t^1 B_t^2\right)$$

as an Itô process.

Solution 5.7 *Using multidimensional Itô's formula, with $f(t,x) = f(t,x_1,x_2) = (x_1 t, x_2^2 - x_1 x_2)$, we have*

$$dY_t^1 = B_t^1 dt + t dB_t^1,$$
$$dY_t^2 = -B_t^2 dB_t^1 + \left(2B_t^2 - B_t^1\right) dB_t^2 + dt,$$

that is,

$$Y_t^1 = \int_0^t B_s^1 ds + \int_0^t s dB_s^1,$$
$$Y_t^1 = -\int_0^t B_s^2 dB_s^1 + \int_0^t \left(2B_s^2 - B_s^1\right) dB_s^2 + t.$$

We shall now describe how could one prove rigorously one-dimensional Itô's formula. The process

$$Y_t = f(0, X_0) + \int_0^t \frac{\partial f}{\partial t}(s, X_s)\, ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, u_s dB_s + \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, v_s ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s)\, u_s^2 ds.$$

is an Itô process. We assume that f and it's partial derivatives are bounded (the general case may be proven by approximating f and it's partial derivatives by bounded functions). As we know, the stochastic integral may be approximated by a sequence of stochastic integrals of simple processes, and so we may assume u and v are simple processes.

By dividing the interval $[0,t]$ in n equal-sized subintervals, we get

$$f(t, X_t) = f(0, X_0) + \sum_{k=0}^{n-1} \left(f\left(t_{k+1}, X_{t_{k+1}}\right) - f\left(t_k, X_{t_k}\right)\right).$$

Using Taylor's expansion

$$f\left(t_{k+1}, X_{t_{k+1}}\right) - f\left(t_k, X_{t_k}\right) = \frac{\partial f}{\partial t}\left(t_k, X_{t_k}\right)\Delta t + \frac{\partial f}{\partial x}\left(t_k, X_{t_k}\right)\Delta X_k + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\left(t_k, X_{t_k}\right)\left(\Delta X_k\right)^2 + Q_k,$$

where Q_k is the remainder of Taylor's formula. We also have

$$\begin{aligned}\Delta X_k = X_{t_{k+1}} - X_{t_k} = \int_{t_k}^{t_{k+1}} v_s ds + \int_{t_k}^{t_{k+1}} u_s dB_s \\ = v(t_k)\Delta t + u(t_k)\Delta B_k + S_k,\end{aligned}$$

where S_k is the remainder. From here, we obtain

$$\begin{aligned}(\Delta X_k)^2 = (v(t_k))^2 (\Delta t)^2 + (u(t_k))^2 (\Delta B_k)^2 \\ + 2v(t_k) u(t_k) \Delta t \Delta B_k + P_k,\end{aligned}$$

where P_k is the remainder. By substituting these terms we get

$$f(t, X_t) - f(0, X_0) = I_1 + I_2 + I_3 + \frac{1}{2} I_4 + \frac{1}{2} K_1 + K_2 + R,$$

where

$$\begin{aligned}I_1 &= \sum_k \frac{\partial f}{\partial t}(t_k, X_{t_k}) \Delta t, \\ I_2 &= \sum_k \frac{\partial f}{\partial t}(t_k, X_{t_k}) v(t_k) \Delta t, \\ I_3 &= \sum_k \frac{\partial f}{\partial x}(t_k, X_{t_k}) u(t_k) \Delta B_k, \\ I_4 &= \sum_k \frac{\partial^2 f}{\partial x^2}(t_k, X_{t_k}) (u(t_k))^2 (\Delta B_k)^2.\end{aligned}$$

$$\begin{aligned}K_1 &= \sum_k \frac{\partial^2 f}{\partial x^2}(t_k, X_{t_k}) (v(t_k))^2 (\Delta t)^2, \\ K_2 &= \sum_k \frac{\partial^2 f}{\partial x^2}(t_k, X_{t_k}) v(t_k) u(t_k) \Delta t \Delta B_k, \\ R &= \sum_k (Q_k + S_k + P_k).\end{aligned}$$

When $n \to \infty$, it is easily shown that

$$\begin{aligned}I_1 &\to \int_0^t \frac{\partial f}{\partial t}(s, X_s) ds, \\ I_2 &\to \int_0^t \frac{\partial f}{\partial x}(s, X_s) v_s ds, \\ I_3 &\to \int_0^t \frac{\partial f}{\partial x}(s, X_s) u_s dB_s.\end{aligned}$$

As we've seen before (quadratic variation of the Brownian motion),

$$\sum_k (\Delta B_k)^2 \to t,$$

whereby

$$I_4 \to \int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s)\, u_s^2 ds.$$

On the other hand, we also have

$$K_1 \to 0,$$
$$K_2 \to 0.$$

Although harder and more technical, it may also be proved that

$$R \to 0.$$

As a conclusion, we obtain Itô's formula when taking the limit.

5.3 Itô's integral representation theorem

Let $u \in L^2_{a,T}$ (u adapted, measurable and square-integrable) and let

$$M_t = \mathbb{E}[M_0] + \int_0^t u_s dB_s. \tag{5.3}$$

We already know that M_t is a $\mathcal{F}_t$-martingale. We shall now show that any squared-integrable martingale is of the form(5.3).

Theorem 5.8 *(Itô's integral representation): Let $F \in L^2(\Omega, \mathcal{F}_T, P)$. Then there exists one unique process $u \in L^2_{a,T}$ such that*

$$F = \mathbb{E}[F] + \int_0^t u_s dB_s. \tag{5.4}$$

Proof. We shall divide the proof in 3 parts:

1. Consider a random variable F of the form

$$F = \exp\left(\int_0^T h(s)\, dB_s - \frac{1}{2}\int_0^T h(s)^2\, ds\right), \tag{5.5}$$

where h is a deterministic function $\int_0^T h(s)^2 ds < \infty$. Let us apply Itô's formula to $f(x) = e^x$, with $X_t = \int_0^t h(s) dB_s - \frac{1}{2}\int_0^t h(s)^2 ds$ and $Y_t = f(X_t)$. Then

$$\begin{aligned} dY_t &= Y_t \left(h(t) dB_t - \frac{1}{2} h(t)^2 dt \right) + \frac{1}{2} Y_t (h(t) dB_t)^2 \\ &= Y_t h(t) dB_t. \end{aligned}$$

That is,

$$Y_t = 1 + \int_0^t Y_s h(s) dB_s.$$

It is then obtained

$$\begin{aligned} F = Y_T &= 1 + \int_0^T Y_s h(s) dB_s \\ &= \mathbb{E}[F] + \int_0^T Y_s h(s) dB_s \end{aligned}$$

Note that

$$\mathbb{E}\left[\int_0^T (Y_s h(s))^2 ds \right] < \infty,$$

due to $\mathbb{E}[Y_t^2] = \exp\left(\int_0^t h(u)^2 du\right) < \infty$. Therefore

$$\begin{aligned} \mathbb{E}\left[\int_0^T (Y_s h(s))^2 ds \right] &\le \int_0^T \exp\left(\int_0^s h(u)^2 du \right) h(s)^2 ds \\ &\le \exp\left(\int_0^T h(u)^2 du \right) \int_0^T h(s)^2 ds. \end{aligned}$$

2. The representation (5.4) is also valid (by linearity) for linear combinations of the form (5.5). The general case, $F \in L^2(\Omega, \mathcal{F}_T, P)$ may then be approximated (in mean-squares sense) by the sequence $\{F_n\}$ of linear combinations of random variables of the form (5.5). A more detailed approach of this may be found in [10]. Then:

$$F_n = \mathbb{E}[F_n] + \int_0^t u_s^{(n)} dB_s.$$

By Itô's isometry, we have

$$\begin{aligned} \mathbb{E}\left[(F_n - F_m)^2 \right] &= (\mathbb{E}[F_n - F_m])^2 + \mathbb{E}\left[\int_0^t \left(u_s^{(n)} - u_s^{(m)} \right)^2 ds \right] \\ &\ge \mathbb{E}\left[\int_0^t \left(u_s^{(n)} - u_s^{(m)} \right)^2 ds \right] \end{aligned}$$

and $\{F_n\}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}_T, P)$. Hence,

$$\mathbb{E}\left[(F_n - F_m)^2\right] \longrightarrow 0 \text{ when } n, m \to \infty.$$

Thus

$$\mathbb{E}\left[\int_0^t \left(u_s^{(n)} - u_s^{(m)}\right)^2 ds\right] \longrightarrow 0 \text{ qdo. } n, m \to \infty.$$

So $\left\{u^{(n)}\right\}$ is a Cauchy sequence in $L^2([0,T] \times \Omega)$. As this is a complete space, $u^{(n)} \to u$ in $L^2([0,T] \times \Omega)$. The process u is adapted because $u^{(n)} \in L^2_{a,T}$ and exists a subsequence of $\left\{u^{(n)}(t,\omega)\right\}$ converging to $u(t,\omega)$ a.s. in $(t,\omega) \in [0,T] \times \Omega$. Then, $u(t,\cdot)$ is $\mathcal{F}_t$-measurable for almost every t. Modifying the process u in a null measure set on the variable t, we obtain a process u that is adapted to $\{\mathcal{F}_t\}$. Therefore

$$\lim_{n\to\infty} \mathbb{E}\left[(F_n - F)^2\right] = \lim_{n\to\infty} \mathbb{E}\left(\mathbb{E}[F_n] + \int_0^T u_s^{(n)} dB_s - F\right)^2 = 0.$$

On the other hand, by Itô's isometry,

$$\lim_{n\to\infty} \mathbb{E}\left(\mathbb{E}[F_n] - \mathbb{E}[F]\right)^2 = 0$$

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T \left(u_s^{(n)} - u_s\right) dB_s\right)^2 = \lim_{n\to\infty} \mathbb{E}\int_0^T \left(u_s^{(n)} - u_s\right)^2 ds = 0.$$

and thus $F = \mathbb{E}[F] + \int_0^T u_s dB_s$.

3. Uniqueness: Suppose $u^{(1)}$ and $u^{(2)} \in L^2_{a,T}$ and

$$F = \mathbb{E}[F] + \int_0^T u_s^{(1)} dB_s = \mathbb{E}[F] + \int_0^T u_s^{(2)} dB_s.$$

By Itô's isometry,

$$\mathbb{E}\left[\left(\int_0^T \left(u_s^{(1)} - u_s^{(2)}\right) dB_s\right)^2\right] = \mathbb{E}\left[\int_0^T \left(u_s^{(1)} - u_s^{(2)}\right)^2 ds\right] = 0$$

hence

$$u^{(1)}(t,\omega) = u(t,\omega)^{(2)} \quad \text{p.q.t. } (t,\omega) \in [0,T] \times \Omega.$$

■

5.4 Martingale representation theorem

Theorem 5.9 *(Martingale representation theorem) Suppose* $\{M_t, t \in [0,T]\}$ *is a* $\{\mathcal{F}_t\}$*-martingale and* $\mathbb{E}[M_T^2] < \infty$. *Then there exists a unique process* $u \in L^2_{a,T}$ *such that*

$$M_t = \mathbb{E}[M_0] + \int_0^t u_s dB_s \qquad \forall t \in [0,T].$$

Proof. Itô's representation theorem may be applied to $F = M_T$. So $\exists^1 u \in L^2_{a,T}$ such that

$$M_T = \mathbb{E}[M_T] + \int_0^T u_s dB_s.$$

As $\{M_t, t \in [0,T]\}$ is a martingale, $\mathbb{E}[M_T] = \mathbb{E}[M_0]$ and

$$\begin{aligned} M_t &= \mathbb{E}[M_T|\mathcal{F}_t] = \mathbb{E}[\mathbb{E}[M_T]|\mathcal{F}_t] + \mathbb{E}\left[\int_0^T u_s dB_s|\mathcal{F}_t\right] \\ &= \mathbb{E}[M_0] + \int_0^t u_s dB_s. \end{aligned}$$

where it was used the martingale property of the undefined stochastic integral. ■

Example 5.10 *Let* $F = B_T^3$. *What is the Itô integral representation of this r.v.? By Itô's formula (applied to* $f(x) = x^3$ *and* $B_T^3 = f(B_t)$*), we have:*

$$B_T^3 = \int_0^T 3B_t^2 dB_t + 3\int_0^T B_t dt.$$

Integrating by parts,

$$\int_0^T B_t dt = TB_T - \int_0^T t dB_t = \int_0^T (T-t)\, dB_t.$$

Hence

$$F = B_T^3 = \int_0^T 3\left[B_t^2 + (T-t)\right] dB_t. \tag{5.6}$$

And as $\mathbb{E}[B_T^3] = 0$ *(because* $B_T \sim N(0,T)$*), Itô's integral representation is given by (5.6).*

Example 5.11 *Which process u satisfies $\int_0^T tB_t^2 dt - \frac{T^2}{2}B_T^2 = -\frac{T^3}{6} + \int_0^T u_t dB_t$? Applying Itô's formula to $X_t = f(t, B_t) = t^2 B_t^2$, with $f(t,x) = t^2 x^2$, we get:*

$$T^2 B_T^2 = \int_0^T 2tB_t^2 dt + \int_0^T 2t^2 B_t dB_t + \int_0^T t^2 dt.$$

From here we obtain

$$\int_0^T tB_t^2 dt - \frac{T^2}{2}B_T^2 = -\frac{T^3}{6} - \int_0^T t^2 B_t dB_t$$

and so

$$u_t = -t^2 B_t.$$

Note that $\mathbb{E}\left[\int_0^T tB_t^2 dt - \frac{T^2}{2}B_T^2\right] = -\frac{T^3}{6}$.

The integration by parts formula for a general case is given next.

Theorem 5.12 *(integration by parts) Suppose $f(s)$ is a deterministic function of class C^1. Then,*

$$\int_0^t f(s)\, dB_s = f(t)\, B_t - \int_0^t f'(s)\, B_s ds.$$

Proof. To prove this formula, one just has to apply Itô's formula to $g(t,x) = f(t)\, x$, obtaining

$$f(t)\, B_t = \int_0^t f'(s)\, B_s ds + \int_0^t f(s)\, dB_s.$$

■

Chapter 6

Stochastic Differential Equations

6.1 Motivation and Examples

A deterministic ordinary differential equation (ODE) of order n has the general form

$$f\left(t, x\left(t\right), x'\left(t\right), x''\left(t\right), \ldots, x^{(n)}(t)\right) = 0, \qquad 0 \leq t \leq T,$$

where $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ is a function and $x(t)$ is the unknown function. A differential equation of order 1 may be represented by

$$\frac{dx\left(t\right)}{dt} = b\left(t, x\left(t\right)\right)$$

or

$$dx\left(t\right) = b\left(t, x\left(t\right)\right) dt$$

The discrete version is the difference equation

$$\Delta x\left(t\right) = x\left(t + \Delta t\right) - x\left(t\right) \approx b\left(t, x\left(t\right)\right) \Delta t$$

Example 6.1 *The first order linear ODE:*

$$\frac{dx\left(t\right)}{dt} = cx\left(t\right),$$

where c is constant, has the solution:

$$x\left(t\right) = x\left(0\right) e^{ct}.$$

A stochastic differential equation (SDE) may be generally written in differential form as

$$\begin{aligned} dX_t &= b\left(t, X_t\right) dt + \sigma\left(t, X_t\right) dB_t, \qquad (6.1) \\ X_0 &= X_0, \end{aligned}$$

where $b(t, X_t)$ is the drift coefficient and $\sigma(t, X_t)$ is the diffusion coefficient. The same SDE may be expressed in integral form as

$$X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dB_s. \tag{6.2}$$

A "naïve" interpretation of the SDE may be done by considering the discrete version $\Delta X_t \approx b(t, X_t)\,\Delta t + \sigma(t, X_t)\,\Delta B_t$ and so the random variable ΔX_t has a distribution "close" to a normal distribution $N\left(b(t, X_t)\,\Delta t, (\sigma(t, X_t))^2\,\Delta t\right)$.

A more rigorous definition of the solution of a stochastic differential equation is presented next.

Definition 6.2 *A solution of the SDE (6.1) or (6.2) is a stochastic process* $\{X_t\}$ *that satisfies:*

1. $\{X_t\}$ *is adapted to the Brownian motion with continuous trajectories.*
2. $\mathbb{E}\left[\int_0^T (\sigma(s, X_s))^2\, ds\right] < \infty.$
3. $\{X_t\}$ *satisfies the SDE (6.1) or (6.2)*

The solutions of a stochastic differential equation are also known as "diffusions" or "diffusion processes".

6.2 The SDE of the geometric Brownian motion and the Langevin equation

Consider the SDE:

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \tag{6.3}$$

where both μ and σ are constant, or

$$X_t = X_0 + \mu \int_0^t X_s ds + \sigma \int_0^t X_s dB_s. \tag{6.4}$$

How may this equation be solved? Suppose $X_t = f(t, B_t)$, where f is a $C^{1,2}$ function. By Itô's formula,

$$X_t = f(t, B_t) = X_0 + \int_0^t \left(\frac{\partial f}{\partial t}(s, B_s) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s, B_s) \right) ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s)\, dB_s. \tag{6.5}$$

Comparing (6.4) with (6.5) we have (because the representation as an Itô process is unique)

$$\frac{\partial f}{\partial s}(s,B_s)+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s,B_s)=\mu f(s,B_s), \tag{6.6}$$

$$\frac{\partial f}{\partial x}(s,B_s)=\sigma f(s,B_s). \tag{6.7}$$

Deriving eq. (6.7), we obtain

$$\frac{\partial^2 f}{\partial x^2}(s,x)=\sigma\frac{\partial f}{\partial x}(s,x)=\sigma^2 f(s,x)$$

and substituting in (6.6), we have

$$\left(\mu-\frac{1}{2}\sigma^2\right)f(s,x)=\frac{\partial f}{\partial s}(s,x).$$

By separation of variables $f(s,x)=g(s)\,h(x)$, we obtain

$$\frac{\partial f}{\partial s}(s,x)=g'(s)\,h(x)$$

and

$$g'(s)=\left(\mu-\frac{1}{2}\sigma^2\right)g(s)$$

that is a linear ODE with solution given by

$$g(s)=g(0)\exp\left[\left(\mu-\frac{1}{2}\sigma^2\right)s\right]$$

Using eq. (6.7), we get $h'(x)=\sigma h(x)$ and therefore

$$f(s,x)=f(0,0)\exp\left[\left(\mu-\frac{1}{2}\sigma^2\right)s+\sigma x\right].$$

We hence conclude the solution of the SDE (6.3) is the process

$$X_t=f(t,B_t)=X_0\exp\left[\left(\mu-\frac{1}{2}\sigma^2\right)t+\sigma B_t\right], \tag{6.8}$$

that is precisely the geometric Brownian motion. Note that this solution was obtained by solving a deterministic partial differential (PDE).

In order to verify that (6.8) satisfies the SDE (6.3) or (6.4), it is sufficient to apply Itô's formula to $X_t=f(t,B_t)$, with

$$f(t,x)=X_0\exp\left[\left(\mu-\frac{1}{2}\sigma^2\right)t+\sigma x\right].$$

We get

$$\begin{aligned} X_t &= X_0 + \int_0^t \left[\left(\mu - \frac{1}{2}\sigma^2 \right) X_s + \frac{1}{2}\sigma^2 X_s \right] ds + \int_0^t \sigma X_s dB_s \\ &= X_0 + \mu \int_0^t X_s ds + \sigma \int_0^t X_s dB_s, \end{aligned}$$

or

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

and hence the SDE is satisfied by the geometric Brownian motion.

Let us now consider the Langevin equation

$$dX_t = \mu X_t dt + \sigma dB_t, \tag{6.9}$$

where μ and σ are constant, or

$$X_t = X_0 + \mu \int_0^t X_s ds + \sigma \int_0^t dB_s.$$

The discrete version of this SDE is

$$X_{t+1} = (1 + \mu) X_t + \sigma (B_{t+1} - B_t),$$

or

$$X_{t+1} = \phi X_t + Z_t,$$

with $\phi = 1 + \mu$ and $Z_t \sim N(0, \sigma^2)$, that is the equation for an autoregressive time series of order 1.

Consider the process

$$Y_t = e^{-\mu t} X_t$$

or $Y_t = f(t, X_t)$ with $f(t, x) = e^{-\mu t} x$. By Itô's formula,

$$\begin{aligned} Y_t = Y_0 &+ \int_0^t \left(-ce^{-\mu s} X_s + ce^{-\mu s} X_s + \frac{1}{2}\sigma^2 \times 0 \right) ds \\ &+ \int_0^t \sigma e^{-\mu s} dB_s. \end{aligned}$$

Therefore the solution of the equation (6.9) is the process

$$X_t = e^{\mu t} X_0 + e^{\mu t} \int_0^t \sigma e^{-\mu s} dB_s. \tag{6.10}$$

If X_0 is constant, the process (6.10) is called an Ornstein-Uhlenbeck process.

Example 6.3 *The geometric Brownian motion (again). Consider the SDE*

$$dX_t = \mu X_t dt + \sigma X_t dB_t \tag{6.11}$$

or

$$X_t = X_0 + \mu \int_0^t X_s ds + \sigma \int_0^t X_s dB_s. \tag{6.12}$$

Let us start by supposing that the solution to this equation may be expressed as

$$X_t = e^{Z_t},$$

where Z_t is a stochastic process. Equivalently,

$$Z_t = \ln (X_t).$$

By applying Itô's formula to $f(X_t) = \ln (X_t)$, we obtain

$$\begin{aligned} dZ_t &= \frac{1}{X_t} dX_t + \frac{1}{2}\left(\frac{-1}{X_t^2}\right)(dX_t)^2 \\ &= \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dB_t. \end{aligned}$$

that is,

$$Z_t = Z_0 + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma B_t$$

and

$$X_t = X_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma B_t\right].$$

So we once more obtain the geometric Brownian motion as a solution of the SDE (6.11).

The solution of the linear homogeneous SDE

$$dX_t = b(t) X_t dt + \sigma(t) X_t dB_t$$

is given by

$$X_t = X_0 \exp\left[\int_0^t \left(b(s) - \frac{1}{2}\sigma(s)^2\right) ds + \int_0^t \sigma(s) dB_s\right].$$

To obtain this solution, the same technique used in Example 6.3 may be applied.

Exercise 6.4 *Determine the solution of the SDE*

$$dX_t = a\,(m - X_t)\,dt + \sigma dB_t,$$
$$X_0 = x,$$

where $a, \sigma > 0$ and $m \in \mathbb{R}$. Compute the mean and variance of X_t and determine the distribution of X_t when $t \to \infty$ (invariant or stationary distribution).

Exercise 6.5 *Consider the SDE*

$$dX_t = \mu X_t dt + \sigma X_t dB_t,$$
$$X_0 = X_0.$$

a)Supposing $X_t = f(t, B_t)$, where f is a $C^{1,2}$ function, apply Itô's formula and determine the partial differential equation (PDE) satisfied by function f.

b) Using the method of separation of variables ($f(s,x) = g(s)h(x)$), and considering the PDE obtained in a), determine the ordinary differential equations (ODE's) satisfied by g and h.

c) Determine the process X_t.

6.3 Existence and Uniqueness Theorem for SDE's

Theorem 6.6 *(existence and uniqueness of solutions) Let $T > 0$, $b(\cdot,\cdot) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma(\cdot,\cdot) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ measurable functions such that the following conditions are satisfied.*

1. *Linear growth property:*

$$|b\,(t,x)| + |\sigma\,(t,x)| \le C\,(1 + |x|)\,, \quad \forall x \in \mathbb{R}^n, \forall t \in [0,T]\,.$$

2. *Lipschitz property:*

$$|b\,(t,x) - b\,(t,y)| + |\sigma\,(t,x) - \sigma\,(t,y)| \le D\,|x - y|\,,\ \forall x, y \in \mathbb{R}^n,\ \forall t \in [0,T]\,.$$

Also, assume Z is a random variable independent from the Brownian motion B and $\mathbb{E}\left[|Z|^2\right] < \infty$. Then, the SDE

$$X_t = Z + \int_0^t b\,(s, X_s)\,ds + \int_0^t \sigma\,(s, X_s)\,dB_s \qquad (6.13)$$

has a unique solution. This means that there is an unique stochastic process $X = \{X_t, 0 \le t \le T\}$ that is continuous, adapted, satisfies (6.13) and

$$E\left[\int_0^T |X_s|^2\,ds\right] < \infty.$$

Proof. Consider the space $L^2_{a,T}$ of the processes adapted to the filtration $\mathcal{F}^Z_t := \sigma(Z) \cup \mathcal{F}_t$ such that $E\left[\int_0^T |X_s|^2 ds\right] < \infty$. On this space, consider the norm:

$$\|X\| = \left(\int_0^T e^{-\lambda s} E\left[|X_s|^2\right] ds\right)^{\frac{1}{2}},$$

where $\lambda > 2D^2 (T+1)$.

Define the operator $\mathcal{L} : L^2_{a,T} \to L^2_{a,T}$ by:

$$(\mathcal{L}X)_t = Z + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dB_s$$

The linear growth property on b and σ ensures the operator $\mathcal{L}$ is well defined. Using Cauchy-Schwarz inequality and Itô's isometry,

$$E\left[|(\mathcal{L}X)_t - (\mathcal{L}Y)_t|^2\right] \le 2E\left[\left(\int_0^t (b(s, X_s) - b(s, Y_s))\, ds\right)^2\right]$$
$$+ 2E\left[\left(\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s))\, dB_s\right)^2\right]$$
$$\le 2TE\left[\int_0^t (b(s, X_s) - b(s, Y_s))^2\, ds\right] +$$
$$+ 2E\left[\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s))^2\, ds\right]$$

By the Lipschitz property,

$$E\left[|(\mathcal{L}X)_t - (\mathcal{L}Y)_t|^2\right] \le 2D^2 (T+1)\, E\left[\int_0^t (X_s - Y_s)^2\, ds\right].$$

Define $K = 2D^2 (T+1)$. Multiplying the above inequality by $e^{-\lambda t}$ and integrating in $[0, T]$, we get

$$\int_0^T e^{-\lambda t} E\left[|(\mathcal{L}X)_t - (\mathcal{L}Y)_t|^2\right] dt$$
$$\le K \int_0^T e^{-\lambda t} E\left[\int_0^t (X_s - Y_s)^2\, ds\right] dt.$$

Changing the integration order,

$$= K \int_0^T \left[\int_s^T e^{-\lambda t} dt\right] E\left[(X_s - Y_s)^2\right] ds$$
$$\le \frac{K}{\lambda} \int_0^T e^{-\lambda s} E\left[(X_s - Y_s)^2\right] ds.$$

Therefore:

$$\|(\mathcal{L}X) - (\mathcal{L}Y)\| \leq \sqrt{\frac{K}{\lambda}} \|X - Y\|.$$

As $\lambda > K$, we have that $\sqrt{\frac{K}{\lambda}} < 1$, hence the operator $\mathcal{L}$ is a contraction on the space $L^2_{a,T}$. Then, by Banach's Fixed Point Theorem, there exists a unique fixed point for $\mathcal{L}$, which is precisely the solution of the SDE:

$$(\mathcal{L}X)_t = X_t.$$

■

[10] provides a different approach to this same proof, based on Picard approximations and Grownwall's inequality.

6.4 Examples

Example 6.7 *The geometric Brownian motion*

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B_t\right]$$

is the solution of the SDE

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dB_t, \\ S_0 &= S_0. \end{aligned}$$

This SDE describes the evolution of the price of a risky financial asset on the Black-Scholes model. Consider the Black-Scholes model with coefficients $\mu(t)$ and $\sigma(t) > 0$ time dependent.

$$\begin{aligned} dS_t &= S_t (\mu(t)\, dt + \sigma(t)\, dB_t), \\ S_0 &= S_0. \end{aligned}$$

Let $S_t = \exp(Z_t)$ and $Z_t = \ln(S_t)$. By Itô's formula taking $f(x) = \ln(x)$,

$$\begin{aligned} dZ_t &= \frac{1}{S_t} (S_t (\mu(t)\, dt + \sigma(t)\, dB_t)) - \frac{1}{2S_t^2} \left(S_t^2 \sigma^2(t)\, dt\right) \\ &= \left(\mu(t) - \frac{1}{2}\sigma^2(t)\right) dt + \sigma(t)\, dB_t. \end{aligned}$$

So,

$$Z_t = Z_0 + \int_0^t \left(\mu(s) - \frac{1}{2}\sigma^2(s)\right) ds + \int_0^t \sigma(s)\, dB_s$$

and therefore

$$S_t = S_0 \exp\left(\int_0^t \left(\mu(s) - \frac{1}{2}\sigma^2(s)\right) ds + \int_0^t \sigma(s)\, dB_s\right).$$

Example 6.8 *(Orsntein-Uhlenbeck process with mean-reversion). Consider the SDE with mean reversion*

$$\begin{aligned} dX_t &= a(m - X_t)\, dt + \sigma dB_t, \\ X_0 &= x, \end{aligned}$$

where $a, \sigma > 0$ and $m \in \mathbb{R}$.

The solution of the corresponding homogeneous ODE $dx_t = -ax_t dt$ is $x_t = xe^{-at}$. Consider the change of variables $X_t = Y_t e^{-at}$ or $Y_t = X_t e^{at}$. Applying Itô's formula to $f(t,x) = xe^{at}$,

$$Y_t = x + m\left(e^{at} - 1\right) + \sigma \int_0^t e^{as} dB_s.$$

Hence,

$$X_t = m + (x - m)\, e^{-at} + \sigma e^{-at} \int_0^t e^{as} dB_s. \tag{6.14}$$

This process is known as an Orsntein-Uhlenbeck process with mean reversion. It is a Gaussian process, as a stochastic integral of the form $\int_0^t f(s)\, dB_s$, where f is a deterministic function, is a Gaussian process. It's expected value is

$$E[X_t] = m + (x - m)\, e^{-at}$$

and the covariance function is obtained by applying Itô's isometry:

$$\begin{aligned} Cov[X_t, X_s] &= \sigma^2 e^{-a(t+s)} E\left[\left(\int_0^t e^{ar} dB_r\right)\left(\int_0^s e^{ar} dB_r\right)\right] \\ &= \sigma^2 e^{-a(t+s)} \int_0^{t \wedge s} e^{2ar} dr \\ &= \frac{\sigma^2}{2a}\left(e^{-a|t-s|} - e^{-a(t+s)}\right). \end{aligned}$$

Note that

$$X_t \sim N\left[m + (x - m)\, e^{-at}, \frac{\sigma^2}{2a}\left(1 - e^{-2at}\right)\right].$$

When $t \to \infty$, the distribution of X_t converges to

$$\nu := N\left[m, \frac{\sigma^2}{2a}\right],$$

which is it's invariant or stationary distribution. If X_0 has distribution ν then the distribution of X_t will be ν for all t.

Some financial applications of the Ornstein-Uhlenbeck process with mean reversion are shown below:

- *Vasicek's model for interest rates:*

$$dr_t = a(b - r_t)\,dt + \sigma dB_t,$$

  with a, b, σ constants. The solution of the SDE is

$$r_t = b + (r_0 - b)\,e^{-at} + \sigma e^{-at} \int_0^t e^{as} dB_s.$$

- *Black-Scholes model with stochastic volatility: assume $\sigma(t) = f(Y_t)$ is a function of an Ornstein-Uhlenbeck process with mean reversion*

$$dY_t = a(m - Y_t)\,dt + \beta dW_t,$$

  with a, m, β constants and where $\{W_t, 0 \le t \le T\}$ is a Brownian motion. The SDE that models the evolution of the price of a risky asset is

$$dS_t = \mu S_t dt + f(Y_t)\,S_t dB_t$$

  where $\{B_t, 0 \le t \le T\}$ is a Brownian motion. These Brownian motions W_t e B_t may be correlated, i.e.,

$$E[B_t W_s] = \rho(s \wedge t).$$

6.5 Linear SDE's

Consider the SDE

$$X_t = x + \int_0^t f(s, X_s)\,ds + \int_0^t c(s)\,X_s dB_s,$$

where f and c are continuous, deterministic functions. Suppose f satisfies the Lipschitz and linear growth conditions on x. Then, by the existence and uniqueness of solutions theorem, there is a unique solution for the SDE. In order to obtain this solution, consider the "integrating factor"

$$F_t = \exp\left(\int_0^t c(s)\,dB_s - \frac{1}{2}\int_0^t c(s)^2\,ds\right).$$

Clearly, F_t is a solution for the SDE if $f = 0$ and $x = 1$. Suppose $X_t = F_t Y_t$ or $Y_t = (F_t)^{-1} X_t$. Then, by Itô's formula,

$$dY_t = (F_t)^{-1} f(t, F_t Y_t)\, dt$$

and $Y_0 = x$. This equation for Y is an ODE with random coefficients (it is a deterministic differential equation parameterized by $\omega \in \Omega$).

Example 6.9 *As an example, if we take $f(t,x) = f(t)x$, we get the ODE*

$$\frac{dY_t}{dt} = f(t) Y_t$$

hence

$$Y_t = x \exp\left(\int_0^t f(s)\, ds \right).$$

and therefore,

$$X_t = x \exp\left(\int_0^t f(s)\, ds + \int_0^t c(s)\, dB_s - \frac{1}{2} \int_0^t c(s)^2\, ds \right).$$

Let us now consider a linear SDE of the form

$$\begin{aligned} dX_t &= (a(t) + b(t) X_t)\, dt + (c(t) + d(t) X_t)\, dB_t, \\ X_0 &= x, \end{aligned}$$

where a, b, c, d are continuous, deterministic functions. Suppose the solution may be expressed as

$$X_t = U_t V_t, \tag{6.15}$$

where

$$\begin{cases} dU_t = b(t) U_t dt + d(t) U_t dB_t, \\ dV_t = \alpha(t)\, dt + \beta(t)\, dB_t. \end{cases}$$

and $U_0 = 1$, $V_0 = x$. From the previous example, we already know that

$$U_t = \exp\left(\int_0^t b(s)\, ds + \int_0^t d(s)\, dB_s - \frac{1}{2} \int_0^t d(s)^2\, ds \right) \tag{6.16}$$

On the other hand, by computing the differential of (6.15) and using Itô's formula with $f(u,v) = uv$, we obtain

$$\begin{aligned} dX_t &= V_t dU_t + U_t dV_t + \frac{1}{2}(dU_t)(dV_t) + \frac{1}{2}(dV_t)(dU_t) \\ &= (b(t) X_t + \alpha(t) U_t + \beta(t) d(t) U_t)\, dt + (d(t) X_t + \beta(t) U_t)\, dB_t. \end{aligned}$$

Comparing with the initial SDE for X, we have that

$$\begin{aligned} a(t) &= \alpha(t) U_t + \beta(t) d(t) U_t, \\ c(t) &= \beta(t) U_t. \end{aligned}$$

So,

$$\begin{aligned} \beta(t) &= c(t) U_t^{-1}, \\ \alpha(t) &= [a(t) - c(t) d(t)] U_t^{-1}. \end{aligned}$$

Therefore:

$$X_t = U_t \left(x + \int_0^t [a(s) - c(s) d(s)] U_s^{-1} ds + \int_0^t c(s) U_s^{-1} dB_s \right),$$

where U_t is given by (6.16).

In the one-dimensional case ($n = 1$), the Lipschitz condition for the coefficient σ in the existence and uniqueness of solutions theorem may be weakened if $\sigma(t, x) = \sigma(x)$ (coefficient of diffusion is independent from time). Suppose the coefficient b satisfies the Lipschitz condition and σ satisfies

$$|\sigma(t, x) - \sigma(t, y)| \leq D |x - y|^{\alpha}, \ x, y \in \mathbb{R}, \ t \in [0, T]$$

with $\alpha \geq \frac{1}{2}$. Then, there exists a unique solution for the SDE. Further reading on this case may be found on [5].

Example 6.10 *The SDE of the Cox-Igersoll-Ross interest rate model:*

$$\begin{aligned} dr_t &= a(b - r_t) dt + \sigma \sqrt{r_t} dB_t \\ r_0 &= x, \end{aligned}$$

has a unique solution.

6.6 Strong and weak solutions

The problem of obtaining a solution of an SDE may be defined in a different way. Suppose the only given data is the coefficients $b(t, x)$ and $\sigma(t, x)$, so that we are to find a pair of stochastic processes, $\{X_t\}$ and $\{B_t\}$, defined on a probability space $(\Omega, \mathcal{F}, P)$, and a filtration $\{\mathcal{H}_t\}$ such that $\{B_t\}$ is a $\{\mathcal{H}_t\}$-Brownian motion, and such that $\{X_t\}$ and $\{B_t\}$ satisfy the SDE

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s \tag{6.17}$$

in that probability space.

In this case, it is said that $(\Omega, \mathcal{F}, P, \{\mathcal{H}_t\}, \{X_t\}, \{B_t\})$ is a weak solution of (6.17).

The proofs of the following results are of no interest for us and therefore omitted. The interested reader may find them in [5].

- Every strong solution is also a weak solution.
- It is said that a SDE satisfies the weak uniqueness property if two weak solutions have the same distribution (the same finite dimension distributions, or fidis). If the coefficients satisfy the existence and uniqueness theorem conditions, then the weak uniqueness property is hold for the SDE.
- The existence of weak solutions is guaranteed if the coefficients $b(t, x)$ and $\sigma(t, x)$ are continuous and bounded functions.

Example 6.11 *Consider the Tanaka's SDE*

$$\begin{aligned} dX_t &= \operatorname{sign}(X_t)\, dB_t, \\ X_0 &= 0, \end{aligned}$$

where

$$\operatorname{sign}(x) := \begin{cases} +1 & se\ x \geq 0 \\ -1 & se\ x < 0. \end{cases}$$

Note that $\operatorname{sign}(x)$ *does not satisfy the Lipschitz condition (it is not continuous* 0*). Hence, we may not apply the existence and uniqueness theorem in this case. It can be shown that there is no solution (in the strong sense) for this SDE, but there is a weak (unique) solution - this example is explained in detail in [10].*

Exercise 6.12 *Consider the following system of stochastic differential equations:*

$$dX(t) = 4e^{2t}\, dt + t\, dW(t)\ , \quad X(0) = 10.$$

$$dZ(t) = \left(t^2 + 3\sin t\right)\, dt + 4\, t\, dW(t), \quad Z(0) = 5.$$

a) Write the equation in integral form.
b) Deduce the respective solution.

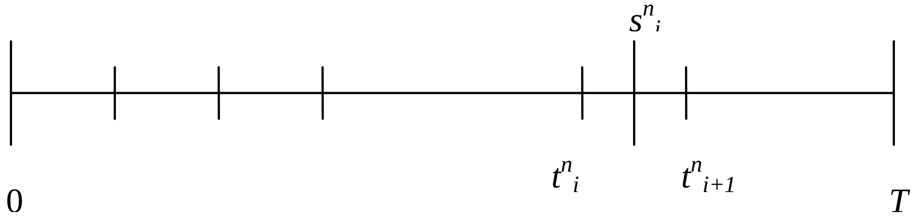

Figure 6.1: Partição

Exercise 6.13 *The Cox-Ingersoll-Ross (CIR) interest rates model $R(t)$ is*

$$dR(t) = (\alpha - \beta R(t))\, dt + \sigma \sqrt{R(t)}\, dW(t),$$

where α, β and σ are positive constants. This equation does not have a closed form solution. However, the expectation and variance of $R(t)$ may be determined.

a) Compute the expected value of $R(t)$. (Hint: Let $X(t) = e^{\beta t} R(t)$. Use the function $f(t, x) = e^{\beta t} x$, apply Itô's formula in differential form, integrate and apply the expectancy operator.)

b) Compute the variance of $R(t)$. (Hint: Compute $d(X^2(t))$ by applying Itô's formula in differential form, integrate and apply the expectancy operator.)

c) Compute $\lim_{t \to +\infty} Var(R(t))$.

6.7 Numerical approximations

Like deterministic differential equations, many SDE's cannot be solved explicitly, so numerical methods are necessary to approximate these solutions. Consider the SDE

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dB_t,$$

with initial condition $X_0 = x$. Let us consider a sequence of partitions of the interval $[0, T]$, where each partition is defined by the points $t_i = \frac{iT}{n}$, $i = 0, 1, \ldots, n$ and the length of each subinterval of the partition is $\delta_n = \frac{T}{n}$

We'll now see, in a summary way, the Euler's finite differences method. The exact values of the solution are:

$$X(t_i) = X(t_{i-1}) + \int_{t_{i-1}}^{t_i} b(X_s)\, ds + \int_{t_{i-1}}^{t_i} \sigma(X_s)\, dB_s. \tag{6.18}$$

Euler approximated is defined by:

$$\int_{t_{i-1}}^{t_i} b(X_s)\,ds \approx b(X(t_{i-1}))\,\delta_n,$$
$$\int_{t_{i-1}}^{t_i} \sigma(X_s)\,dB_s \approx \sigma(X(t_{i-1}))\,\Delta B_i,$$

where $\Delta B_i := B(t_i) - B(t_{i-1})$. Finally, the Euler scheme is defined by

$$X^{(n)}(t_i) = X^{(n)}(t_{i-1}) + b\left(X^{(n)}(t_{i-1})\right)\delta_n + \sigma(X(t_{i-1}))\,\Delta B_i, \tag{6.19}$$

$i = 1, 2, \ldots, n$. In each interval (t_{i-1}, t_i), the value of $X^{(n)}$ is obtained by linear interpolation.

The approximation error is defined by

$$e_n := \sqrt{E\left[\left(X_T - X_T^{(n)}\right)^2\right]}. \tag{6.20}$$

It may be shown, for this scheme, that

$$e_n^{Eul} \leq c\sqrt{\delta_n},$$

where c is a constant.

To simulate a trajectory of the solution using this method, one just needs to follow the procedure:

1. Simulate the values of n random variables with normal distribution $N(0,1)$: $\xi_1, \xi_2, \ldots, \xi_n$.

2. Substitute ΔB_i in (6.19) by $\xi_i\sqrt{\delta_n}$ and determine the values of $X^{(n)}(t_i)$ using the recurrence scheme (6.19).

3. In each interval (t_{i-1}, t_i) determine $X^{(n)}$ by linear interpolation between $X^{(n)}(t_{i-1})$ and $X^{(n)}(t_i)$.

We'll now discuss one other finite differences method - the Milsein method. In order to do so, it is necessary to apply the Itô formula to $b(X_t)$ and $\sigma(X_t)$, considering $t_{i-1} \leq t \leq t_i$. We obtain

$$\begin{aligned}
&\int_{t_{i-1}}^{t_i} b\left(X_s\right) ds = \int_{t_{i-1}}^{t_i} \left[b\left(X\left(t_{i-1}\right)\right) + \right. \\
&\left. + \int_{t_{i-1}}^{s} \left(bb' + \frac{1}{2}b''\sigma^2\right)\left(X_r\right) dr + \int_{t_{i-1}}^{s} \left(\sigma b'\right)\left(X_r\right) dB_r\right] ds, \\
&\int_{t_{i-1}}^{t_i} \sigma\left(X_s\right) dB_s = \int_{t_{i-1}}^{t_i} \left[\sigma\left(X\left(t_{i-1}\right)\right) + \right. \\
&\left. + \int_{t_{i-1}}^{s} \left(b\sigma' + \frac{1}{2}\sigma''\sigma^2\right)\left(X_r\right) dr + \int_{t_{i-1}}^{s} \left(\sigma\sigma'\right)\left(X_r\right) dB_r\right] dB_s.
\end{aligned}$$

Exercise 6.14 *Prove this equality.*

From eq. (6.18), we get

$$X^{(n)}\left(t_i\right) - X^{(n)}\left(t_{i-1}\right) = b\left(X\left(t_{i-1}\right)\right)\delta_n + \sigma\left(X\left(t_{i-1}\right)\right)\Delta B_i + R_i.$$

It can be shown that the dominant term of R_i is the double stochastic integral

$$\int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^{s} \left(\sigma\sigma'\right)\left(X_r\right) dB_r\right) dB_s,$$

being all the lower order terms insignificant. The Milstein approximation is

$$\begin{aligned}
R_i &\approx \int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^{s} \left(\sigma\sigma'\right)\left(X_r\right) dB_r\right) dB_s \\
&\approx \left(\sigma\sigma'\right)\left(X\left(t_{i-1}\right)\right) \int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^{s} dB_r\right) dB_s
\end{aligned}$$

and

$$\begin{aligned}
&\int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^{s} dB_r\right) dB_s = \int_{t_{i-1}}^{t_i} \left(B_s - B\left(t_{i-1}\right)\right) dB_s \\
&= \int_{t_{i-1}}^{t_i} B_s dB_s - B\left(t_{i-1}\right)\left(B\left(t_i\right) - B\left(t_{i-1}\right)\right) \\
&= \frac{1}{2}\left[B_{t_i}^2 - B_{t_{i-1}}^2 - \delta_n\right] - B\left(t_{i-1}\right)\left(B\left(t_i\right) - B\left(t_{i-1}\right)\right) \\
&= \frac{1}{2}\left[\left(\Delta B_i\right)^2 - \delta_n\right],
\end{aligned}$$

where, to compute $\int_{t_{i-1}}^{t_i} B_s dB_s$, one can apply Itô's formula to $f(B_t) = B_t^2$.

Milstein's scheme is

$$X^{(n)}(t_i) = X^{(n)}(t_{i-1}) + b\left(X^{(n)}(t_{i-1})\right)\delta_n + \sigma\left(X(t_{i-1})\right)\Delta B_i$$
$$+\frac{1}{2}\left(\sigma\sigma'\right)\left(X(t_{i-1})\right)\left[\left(\Delta B_i\right)^2 - \delta_n\right].$$

It can be shown that Milstein's approximation error is given by

$$e_n^{Mil} \leq c\delta_n.$$

6.8 The Markov property

The solutions of SDE's are called diffusion processes. Let $X = \{X_t, t \geq 0\}$ be a (n-dimensional) diffusion process that satisfies the SDE:

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dB_t, \tag{6.21}$$

where B is an m-dimensional Brownian motion and both b and σ satisfy the existence and uniqueness of solutions theorem conditions.

Definition 6.15 *A stochastic process $X = \{X_t, t \geq 0\}$ is called a Markov process if $\forall s < t$,*

$$E\left[f(X_t)\,|X_r, r \leq s\right] = E\left[f(X_t)\,|X_s\right].$$

for any bounded and measurable function f defined in $\mathbb{R}^n$.

In particular, if $C \subset \mathbb{R}^n$ and it is measurable,

$$P\left[X_t \in C|X_r, r \leq s\right] = P\left[X_t \in C|X_s\right].$$

Essentially, the Markov property states that "the future values of a process only depend from its present value and not from the past values (given that the present value is known)". The probability law of Markov processes is described by the transition probabilities

$$P(C, t, x, s) := P(X_t \in C|X_s = x), \quad 0 \leq s < t.$$

$P(\cdot, t, x, s)$ is the probability law of X_t conditional to $X_s = x$. If this conditional probability has a density, we represent it by:

$$p(y, t, x, s).$$

Example 6.16 *The Brownian motion is a Markov process with transition probabilities:*

$$p(y,t,x,s) = \frac{1}{\sqrt{2\pi (t-s)}} \exp\left(-\frac{(x-y)^2}{2(t-s)} \right). \tag{6.22}$$

In fact, using the conditional expectation properties, $B_t - B_s$ is independent from $\mathcal{F}_s$ and the fact that B_s is $\mathcal{F}_s$-measurable,

$$\begin{aligned} &P[B_t \in C|\mathcal{F}_s] = P[B_t - B_s + B_s \in C|\mathcal{F}_s] \\ &= P[B_t - B_s + x \in C]\,|_{x=B_s} \\ &= P[B_t \in C|B_s = x], \end{aligned}$$

As $B_t - B_s + x$ is normally distributed with mean x and variance $t-s$, the transition probability density function may be expressed by (6.22).

We shall now introduce some useful notation. We will represent by $\{X_t^{s,x}, t \geq s\}$ the solution of the SDE (6.21) defined in $[s, +\infty)$ and with initial condition $X_s^{s,x} = x$. If $s = 0$, we simplify the notation to $X_t^{0,x} = X_t^x$.

Proposition 6.17 *The following properties hold:*

- 1. There exists a continuous version (with respect to all the parameters s, t, x) of the process $\{X_t^{s,x}, 0 \leq s \leq t, x \in \mathbb{R}^n\}$.
  2. For all $t \geq s$, we have
$$X_t^x = X_t^{s,X_s^x}. \tag{6.23}$$

Proof. Proof of (2): X_t^x satisfies the SDE

$$X_t^x = X_s^x + \int_s^t b(u, X_u^x)\,du + \int_s^t \sigma(u, X_u^x)\,dB_u.$$

On the other hand, $X_t^{s,y}$ satisfies

$$X_t^{s,y} = y + \int_s^t b(u, X_u^{s,y})\,du + \int_s^t \sigma(u, X_u^{s,y})\,dB_u.$$

Substituting y by X_s^x we obtain that X_t^x and X_t^{s,X_s^x} are solutions of the same SDE in $[s, +\infty)$ with the same initial condition X_s^x. Hence, by the existence and uniqueness theorem, $X_t^x = X_t^{s,X_s^x}$. ∎

Theorem 6.18 *(Markov property of diffusion processes) Let f be a bounded and measurable function in $\mathbb{R}^n$. Then, for any $0 \leq s \leq t$,*

$$E\left[f\left(X_t\right)|\mathcal{F}_s\right] = E\left[f\left(X_t^{s,x}\right)\right]|_{x=X_s} \tag{6.24}$$

Proof. By (6.23) and using the properties of the conditional expectation,

$$E\left[f\left(X_t\right)|\mathcal{F}_s\right] = E\left[f\left(X_t^{s,x}\right)|\mathcal{F}_s\right] = E\left[f\left(X_t^{s,x}\right)\right]|_{x=X_s},$$

because $X_t^{s,x}$ is independent from $\mathcal{F}_s$ and X_s is known form the "information" $\mathcal{F}_s$ (i.e., X_s is $\mathcal{F}_s$-measurable). Property 7. of the conditional expectation is used. ■

Diffusion processes are Markov process. The transition probabilities of a diffusion process are given by

$$P\left(C, t, x, s\right) = P(X_t^{s,x} \in C).$$

If a diffusion process is homogeneous with respect to time (the coefficients b and σ do not depend from time) then the Markov property (6.24) may be written as:

$$E\left[f\left(X_t\right)|\mathcal{F}_s\right] = E\left[f\left(X_{t-s}^x\right)\right]|_{x=X_s}.$$

Exercise 6.19 *Compute the transition probabilities for the Ornstein-Uhlenbeck process with mean reversion.*

Solution 6.20 *The mean-reverted Langevin equation is*

$$dX_t = a\left(m - X_t\right)dt + \sigma dB_t.$$

The solution in $[s, +\infty)$, with initial condition $X_s = x$, is given by

$$X_t^{s,x} = m + \left(x - m\right)e^{-a(t-s)} + \sigma e^{-at}\int_s^t e^{ar}dB_r.$$

Thus, as $\left\{\int_s^t e^{ar}dB_r, t \geq s\right\}$ is a Gaussian process with mean 0 and variance $\frac{1}{2a}\left(e^{2at} - e^{2as}\right)$, we have that

$$E\left[X_t^{s,x}\right] = m + \left(x - m\right)e^{-a(t-s)},$$
$$\mathrm{Var}\left[X_t^{s,x}\right] = \frac{\sigma^2}{2a}\left(1 - e^{-2a(t-s)}\right).$$

The transition probability is

$$P\left(\cdot, t, x, s\right) = \textit{Distribution of } X_t^{s,x}.$$

Therefore it is a normal distribution with mean and variance given above.

6.9 Stratovonich Calculus and Stratonovich PDE's

When we defined the Itô integral $\int_0^t u_s dB_s$ for continuous processes using Riemann-Stieltjes-like sums, we have always considered the values of the process u at the point t_{j-1} and assumed the process is constant in $[t_{j-1}, t_j)$. As a consequence of this, the expected value of the Itô integral is null and it's variance may be calculated using Itô's isometry property. Moreover, the Itô integral is a martingale. The downside of the Itô integral is that the "chain rule" (Itô's formula) has now a term of order 2 (which does not appear in classic calculus).

The Stratonovich integral $\int_0^T u_s \circ dB_s$ is defined as the limit (in probability) of the sequence:

$$\sum_{i=1}^{n} \frac{1}{2}\left(u_{t_{i-1}} + u_{t_i}\right)\Delta B_i,$$

with $t_i = \frac{iT}{n}$. We can now impose the question: what is the relationship between the Stratonovich integral and the Itô integral?

If u is an Itô process of the form

$$u_t = u_0 + \int_0^t \beta_s ds + \int_0^t \alpha_s dB_s. \tag{6.25}$$

then it is possible to prove that

$$\int_0^T u_s \circ dB_s = \int_0^T u_s dB_s + \frac{1}{2}\int_0^T \alpha_s ds.$$

The "Itô formula" for the Stratonovich stochastic integral coincides with the classic calculus chain rule. In fact, if u is a process of the form (6.25) and

$$X_t = X_0 + \int_0^t v_s ds + \int_0^t u_s \circ dB_s$$

it can be shown that

$$df\left(X_t\right) = f'\left(X_t\right) \circ dX_t$$

An SDE in the Itô sense may be transformed in a SDE in the Stratonovich sense, by using the formula that relates both integrals:

- Itô form SDE:

$$X_t = X_0 + \int_0^t b\left(s, X_s\right) ds + \int_0^t \sigma\left(s, X_s\right) dB_s.$$

- Equivalent Stratonovich form SDE:

$$X_t = X_0 + \int_0^t b\,(s, X_s)\, ds - \frac{1}{2}\int_0^t (\sigma\sigma')\,(s, X_s)\, ds + \int_0^t \sigma\,(s, X_s) \circ dB_s.$$

- This is because the Itô decomposition of $\sigma\,(t, X_t)$ is

$$\sigma\,(t, X_t) = \sigma\,(0, X_0) + \int_0^t \left(\sigma' b - \frac{1}{2}\sigma''\sigma^2\right)(s, X_s)\, ds + \int_0^t (\sigma\sigma')\,(s, X_s)\, dB_s.$$

Chapter 7

Relationship between PDE's and SDE's

7.1 Infinitesimal operator of a diffusion

Consider an n-dimensional diffusion X that satisfies the SDE

$$\begin{aligned} dX_t &= b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t, \\ X_0 &= x_0 \end{aligned}$$

where B is an m-dimensional Brownian motion. Assume that both b and σ satisfy the conditions of the existence and uniqueness theorem for solutions of SDE's. Consider that $b : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^+ \times \mathbb{R}^n \to M(n, m)$, where $M(n, m)$ is the set of $n \times m$ matrices and $x_0 \in \mathbb{R}^n$.

Definition 7.1 *The generator or infinitesimal operator associated to the diffusion X is the second order differential operator A defined by*

$$Ah(t, x) := \sum_{i=1}^{n} b_i(t, x) \frac{\partial h}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \left(\sigma\sigma^T\right)_{i,j}(t, x) \frac{\partial^2 h}{\partial x_i \partial x_j},$$

where h is a class $C^{1,2}$ function defined in $\mathbb{R}^+ \times \mathbb{R}^n$.

The infinitesimal operator is also known as Dynkin's operator, Itô's operator or "Kolmogorov backward operator". We will now see the relationship between the diffusion X and the operator A. By Itô's formula, if $f(t, x)$ is a class $C^{1,2}$ function, then $f(t, X_t)$ is an Itô process with "differential":

$$df(t, X_t) = \left\{ \frac{\partial f}{\partial t}(t, X_t) + Af(t, X_t) \right\} dt + \left[\nabla_x f(t, X_t)\right] \sigma(t, X_t)\, dB_t, \quad (7.1)$$

where the gradient is defined as:

$$\nabla_x f = \left[\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right].$$

Note that if

$$E\int_0^t \left(\frac{\partial f}{\partial x_i}(t, X_t)\,\sigma_{i,j}(t, X_t)\right)^2 ds < \infty, \tag{7.2}$$

for all $t > 0$ and every i, j, then the stochastic integrals in (7.1) are well defined and are martingales, so that

$$M_t = f(t, X_t) - \int_0^t \left(\frac{\partial f}{\partial s}(s, X_s) + Af(s, X_s)\right) ds$$

is a martingale. A sufficient condition for (7.2) to be satisfied is for the partial derivatives $\frac{\partial f}{\partial s}(s, X_s)$ to show linear growth, i.e.

$$\left|\frac{\partial f}{\partial x_i}(t, x)\right| \leq C(1 + |x|).$$

7.2 Feynman-Kac formulae

The partial differential equation

$$\begin{aligned}\frac{\partial F}{\partial t}(t, x) + AF(t, x) = 0,\\ F(T, x) = \Phi(x)\end{aligned} \tag{7.3}$$

is a parabolic PDE with terminal condition (in T). The previous PDE may also be written as (supposing $n = 1$, to simplify the notation)

$$\begin{aligned}\frac{\partial F}{\partial t}(t, x) + b(t, x)\frac{\partial F}{\partial x} + \frac{1}{2}\sigma^2(t, x)\frac{\partial^2 F}{\partial x^2} = 0,\\ F(T, x) = \Phi(x).\end{aligned} \tag{7.4}$$

Instead of solving the PDE analitically we will try to obtain a solution using a "stochastic representation formula". Suppose there exists a solution F. Fix t and x and define the process X_s in $[t, T]$ as the solution of the SDE

$$\begin{aligned}dX_s &= b(s, X_s)\,ds + \sigma(s, X_s)\,dB_s,\\ X_t &= x.\end{aligned}$$

The infinitesimal operator associated to X_s is

$$A = b(t,x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2}{\partial x^2},$$

which is exactly the operator present in the PDE (7.3) or (7.4). By applying Itô's formula to F, we have (see (7.1))

$$F(T,X_T) = F(t,X_t) + \int_t^T \left(\frac{\partial F}{\partial s}(s,X_s) + AF(s,X_s)\right) ds + \int_t^T \sigma(s,X_s)\frac{\partial F}{\partial x}(s,X_s)\,dB_s.$$

But $\frac{\partial F}{\partial s}(s,X_s) + AF(s,X_s) = 0$ and applying expectation (considering the initial value $X_t = x$), we obtain

$$E_{t,x}[F(T,X_T)] = E_{t,x}[F(t,X_t)],$$

supposing that the stochastic integral is well defined, it's expected value is zero. The boundary values ensure that $E_{t,x}[F(T,X_T)] = E_{t,x}\left[\Phi(X_T^{t,x})\right]$ and $E_{t,x}\left[F\left(t,X_t^{t,x}\right)\right] = F(t,x)$, so

$$F(t,x) = E_{t,x}\left[\Phi(X_T^{t,x})\right],$$

being this the stochastic representation of the PDE (7.4).

Proposition 7.2 *(Feynman-Kac formula) Suppose that F is a solution of the boundary value problem (7.4). Also, suppose $\sigma(s,X_s)\frac{\partial F}{\partial x}(s,X_s)$ is a process in L^2 (i.e. $E\int_0^t \left(\frac{\partial f}{\partial x_i}(t,X_t)\sigma_{i,j}(t,X_t)\right)^2 ds < \infty$). Then*

$$F(t,x) = E_{t,x}\left[\Phi(X_T^{t,x})\right],$$

where $X_s^{t,x}$ satisfies

$$dX_s = b(s,X_s)\,ds + \sigma(s,X_s)\,dB_s,$$
$$X_t = x.$$

Proposition 7.3 *(Feynman-Kac formula) Suppose that F is a solution of the problem (7.3). Also, suppose $E\int_0^t \left(\frac{\partial f}{\partial x_i}(t,X_t)\sigma_{i,j}(t,X_t)\right)^2 ds < \infty$, for all $t > 0$ and every i,j. Then,*

$$F(t,x) = E_{t,x}\left[\Phi(X_T^{t,x})\right],$$

where $X_s^{t,x}$ satisfies

$$dX_s = b(s,X_s)\,ds + \sigma(s,X_s)\,dB_s,$$
$$X_t = x.$$

Consider now a continuous and lower bounded function $q(x)$ and the partial differential equation

$$\begin{aligned} \frac{\partial F}{\partial t}(t,x) + AF(t,x) - q(x)F(t,x) &= 0, \\ F(T,x) &= \Phi(x) \end{aligned} \tag{7.5}$$

with terminal condition (in T). The previous PDE may also be written as (supposing $n = 1$ to simplify notation)

$$\begin{aligned} \frac{\partial F}{\partial t}(t,x) + b(t,x)\frac{\partial F}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F}{\partial x^2} - q(x)F(t,x) &= 0, \\ F(T,x) &= \Phi(x). \end{aligned} \tag{7.6}$$

Again, instead of solving the PDE analytically, we are going to try to solve this problem using a stochastic representation formula. Suppose that there exists a solution F, fix t and x, and define the process X in $[t,T]$ as the solution of the SDE

$$\begin{aligned} dX_s &= b(s,X_s)\,ds + \sigma(s,X_s)\,dB_s, \\ X_t &= x. \end{aligned}$$

The infinitesimal operator associated to X is

$$A = b(t,x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2}{\partial x^2},$$

which is exactly the same operator present in the PDE (7.5) or (7.6). By applying Itô's formula to $g(t,X_t) = \exp\left(-\int_0^t q(X_s)\,ds\right)F(t,X_t)$ and integrating between t and T,

$$\begin{aligned} \exp\left(-\int_0^T q(X_s)\,ds\right)F(T,X_T) &= \exp\left(-\int_0^t q(X_s)\,ds\right)F(t,X_t) + \\ &+ \int_t^T e^{-\int_0^s q(X_r)dr}\left(\frac{\partial F}{\partial s}(s,X_s) + AF(s,X_s) - q(X_s)F(s,X_s)\right)ds \\ &+ \int_t^T \exp\left(-\int_0^s q(X_r)\,dr\right)\sigma(s,X_s)\frac{\partial F}{\partial x}(s,X_s)\,dB_s. \end{aligned}$$

But $\frac{\partial F}{\partial s}(s,X_s) + AF(s,X_s) - q(X_s)F(s,X_s) = 0$ and applying the expected value (conditional to the initial condition $X_t = x$), we obtain

$$E_{t,x}\left[\exp\left(-\int_t^T q(X_s)\,ds\right)F(T,X_T)\right] = E_{t,x}\left[F(t,X_t)\right],$$

supposing that the stochastic integral is well defined and so it's expected value is zero. As

$$E_{t,x}\left[\exp\left(-\int_t^T q\left(X_s\right)ds\right)F\left(T,X_T\right)\right]=E_{t,x}\left[\exp\left(-\int_t^T q\left(X_s\right)ds\right)\Phi(X_T^{t,x})\right]$$

and $E_{t,x}\left[F\left(t,X_t^{t,x}\right)\right]=F\left(t,x\right)$, we have that

$$F\left(t,x\right)=E_{t,x}\left[\exp\left(-\int_t^T q\left(X_s^{t,x}\right)ds\right)\Phi(X_T^{t,x})\right],$$

being this the stochastic representation for the solution of the PDE (7.5) or (7.6).

Proposition 7.4 *(Feynman-Kac formula 2) Let F be a solution of the problem (7.5) or (7.6). Consider that $\sigma\left(s,X_s\right)\frac{\partial F}{\partial x}\left(s,X_s\right)$ is a process in $L^2_{a,T}$ (i.e. $E\int_0^T\left[\frac{\partial F}{\partial x}\left(s,X_s\right)\sigma\left(s,X_s\right)\right]^2ds<\infty$). Then,*

$$F\left(t,x\right)=E_{t,x}\left[\exp\left(-\int_t^T q\left(X_s^{t,x}\right)ds\right)\Phi(X_T^{t,x})\right],$$

where $X_s^{t,x}$ satisfies

$$\begin{aligned}dX_s&=b\left(s,X_s\right)ds+\sigma\left(s,X_s\right)dB_s,\\ X_t&=x.\end{aligned}$$

Remark: Considering $q(x)$ as a lower bounded continuous function, a sufficient condition for $E\int_0^T\left[\exp\left(-\int_0^s q\left(X_r\right)dr\right)\frac{\partial F}{\partial x}\left(s,X_s\right)\sigma\left(s,X_s\right)\right]^2ds<\infty$ is that the derivative $\frac{\partial F}{\partial x}\left(s,x\right)$ has linear growth, i.e.

$$\left|\frac{\partial F}{\partial x}\left(s,x\right)\right|\leq C\left(1+|x|\right).$$

7.3 Relationship between the heat equation and the Brownian motion

Let f be a continuous function with polynomial growth. The function

$$u\left(t,x\right)=E\left[f\left(B_t+x\right)\right]$$

satisfies the heat equation

$$\begin{aligned}\frac{\partial u}{\partial t}&=\frac{1}{2}\frac{\partial^2 u}{\partial x^2},\\ u\left(0,x\right)&=f\left(x\right).\end{aligned}$$

In fact, as B_t has distribution $N(0,t)$, we have that

$$E[f(B_t+x)] = \int_{-\infty}^{+\infty} f(y) \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}} dy,$$

and the function $\frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}$, for each fixed y, satisfies the heat equation $\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}$.

The function $x \to u(t,x)$ represents the temperature distribution on a bar of infinite length, supposing that the initial temperature profile is given by the function $f(x)$.

7.4 Kolmogorov's Backward Equation

Consider a diffusion that is homogeneous in time X that satisfies the SDE

$$\begin{aligned} dX_t &= b(X_t)\,dt + \sigma(X_t)\,dB_t, \\ X_0 &= x. \end{aligned}$$

The infinitesimal generator associated to X does not depend on time and is given by

$$Af(x) := \sum_{i=1}^{n} b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \left(\sigma\sigma^T\right)_{i,j}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}, \tag{7.7}$$

Applying Itô's formula to $f(X_s)$, we obtain

$$df(X_s) = Af(X_s)\,ds + [\nabla_x f(X_s)]\,\sigma(X_s)\,dB_s.$$

and applying the expected value,

$$E[f(X_t^x)] = f(x) + \int_0^t E[Af(X_s)]\,ds. \tag{7.8}$$

Consider the function

$$u(t,x) = E[f(X_t^x)].$$

By (7.8), u is differentiable with respect to t and satisfies the equation:

$$\frac{\partial u}{\partial t} = E[Af(X_t^x)].$$

The expression $E[Af(X_t^x)]$ may be represented as a function of u. In order to do so, we need to introduce the domain of the infinitesimal operator.

Definition 7.5 *The domain D_A of the infinitesimal generator A is the set of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that the following limit exists for all $x \in \mathbb{R}^n$:*

$$Af(x) = \lim_{t \searrow 0} \frac{E[f(X_t^x)] - f(x)}{t}. \tag{7.9}$$

By (7.8), we have that $C_0^2(\mathbb{R}^n) \subset D_A$ and if $f \in C_0^2(\mathbb{R}^n)$, the limit (7.9) is equal to Af given by (7.7). The function $u(t,x)$ satisfies the PDE known as Kolmogorov "backward" equation:

Theorem 7.6 *Let $f \in C_0^2(\mathbb{R}^n)$.*

a) Let $u(t,x) = E[f(X_t^x)]$. Then $u(t,\cdot) \in D_A$ and satisfies the PDE (Kolmogorov "backward" PDE)

$$\begin{aligned} \frac{\partial u}{\partial t} &= Au, \\ u(0,x) &= f(x). \end{aligned} \tag{7.10}$$

b) If $w \in C^{1,2}([0,\infty) \times \mathbb{R}^n)$ is a bounded function that satisfies the PDE (7.10), then

$$w(t,x) = E[f(X_t^x)].$$

Proof. a) One just needs to compute the limit

$$Au = \lim_{r \searrow 0} \frac{E[u(t, X_r^x)] - u(t,x)}{r}$$

By the Markov property, we have that

$$\begin{aligned} E[u(t, X_r^x)] &= E\left[E[f(X_t^y)]\,|_{y=X_r^x}\right] \\ &= E\left[f\left(X_{t+r}^x\right)\right] = u(t+r, x). \end{aligned}$$

Then, as $t \to u(t,x)$ is differentiable,

$$\begin{aligned} \lim_{r \searrow 0} \frac{E[u(t, X_r^x)] - u(t,x)}{r} &= \lim_{r \searrow 0} \frac{u(t+r,x) - u(t,x)}{r} \\ &= \frac{\partial u}{\partial t}. \end{aligned}$$

b) Consider the $n+1$-dimensional process:

$$Y_t = (s - t, X_t^x).$$

Itô's formula applied to $w(Y_t)$, yields

$$w(Y_t) = w(s,x) + \int_0^t \left(Aw - \frac{\partial w}{\partial r} \right)(s-r, X_r^x)\, dr$$
$$+ \int_0^t \sum_{i=1}^n \sum_{j=1}^m \frac{\partial w}{\partial x_i}(s-r, X_r^x)\, \sigma_{i,j}(X_r^x)\, dB_r^j$$

As $Aw = \frac{\partial w}{\partial t}$, we obtain

$$w(Y_t) = w(s,x) + \int_0^t \sum_{i=1}^n \sum_{j=1}^m \frac{\partial w}{\partial x_i}(s-r, X_r^x)\, \sigma_{i,j}(X_r^x)\, dB_r^j$$

We are now interested in applying the expected value. However, as no condition on the growth of the partial derivatives of w was imposed, we are not certain wether the expectation of the stochastic integrals is null.

To resolve this issue, we need to introduce a stopping time τ_R for $R > 0$, given by

$$\tau_R := \inf\{t > 0 : |X_t^x| \geq R\}.$$

If $r \leq \tau_R$, the process $\frac{\partial w}{\partial x_i}(s-r, X_r^x)\, \sigma_{i,j}(X_r^x)$ is bounded, so the stochastic integrals are well defined and their expectation is zero. Thus,

$$E[w(Y_{t\wedge\tau_R})] = w(s,x)$$

and taking $R \to \infty$, we have, for all $t \geq 0$:

$$E[w(Y_t)] = w(s,x).$$

Finally, with $s = t$ and using $w(0,x) = f(x)$, we obtain

$$w(s,x) = E[w(Y_s)] = E[w(0, X_s^x)] = E[f(X_s^x)].$$

■

The following theorem may be proven in a similar fashion (see [10]).

Theorem 7.7 *Let $f \in C_0^2(\mathbb{R}^n)$ and $q \in C(\mathbb{R}^n)$, with q lower bounded.*

a) Let

$$v(t,x) = E\left[\exp\left(-\int_0^t q(X_s^x)\, ds\right) f(X_t^x)\right]$$

. Then $v(t,\cdot) \in D_A$ for every t and it satisfies the PDE

$$\frac{\partial v}{\partial t} = Av - qv, \tag{7.11}$$
$$v(0,x) = f(x).$$

b) If $w \in C^{1,2}([0,\infty) \times \mathbb{R}^n)$ is a bounded function in $[0,T] \times \mathbb{R}^n$ that satisfies the PDE (7.11), then

$$w(t,x) = v(t,x).$$

In the proof of the previous theorem it was necessary to use te concept of stopping time relative to a filtration $\{\mathcal{F}_t, t \geq 0\}$, which is a random variable

$$\tau : \Omega \to [0,+\infty]$$

such that, for all $t \geq 0$, we have that $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$. Intuitively, this means that we may decide if we should or shouldn't "stop" before an instant t, from the information contained in $\mathcal{F}_t$.

Example 7.8 *The arrival time of a continuous and adapted process $X = \{X_t, t \geq 0\}$ at a certain level a, i.e.*

$$\tau_a := \inf\{t > 0 : X_t = a\}$$

is a stopping time. In fact, we have that

$$\{\tau \leq t\} = \left\{\sup_{0 \leq s \leq t} X_s \geq a\right\} = \left\{\sup_{0 \leq s \leq t, s \in \mathbb{Q}} X_s \geq a\right\} \in \mathcal{F}_t$$

We may associate to a stopping time τ the σ-algebra $\mathcal{F}_\tau$ formed by the sets G such that

$$G \cap \{\tau \leq t\} \in \mathcal{F}_t$$

Stopping times satisfy the following properties (see [9]):

1. If $\{M_t, t \in [0,T]\}$ is a continuous martingale and τ is a stopping time bounded by T, then
$$E[M_T | \mathcal{F}_\tau] = M_\tau.$$

2. If $u \in L^2_{a,T}$ and τ is a stopping time bounded by T, the process $u\mathbf{1}_{[0,\tau]}$ also belongs to $L^2_{a,T}$. Moreover,
$$\int_0^T u\mathbf{1}_{[0,\tau]}(t)\, dB_t = \int_0^\tau u(t)\, dB_t$$

7.5 Kolmogorov's Forward Equation

The Kolmogorov equations are partial differential equations for the transition probabilities of the solution of a stochastic differential equation (diffusion). The forward Kolmogorov equation, that we'll discuss throughout this section, is also known as the Fokker–Planck equation, or "master equation" (in natural sciences).

Assume that the process X is a solution of the stochastic differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)\, dB_t, \tag{7.12}$$

with associated infinitesimal generator

$$Af(s,y) = \sum_{i=1}^{n} b_i(s,y) \frac{\partial f}{\partial y_i}(s,y) \tag{7.13}$$

$$+ \frac{1}{2} \sum_{i.j=1}^{n} \left[\sigma(s,y)\, \sigma^T(s,y)\right]_{i,j} \frac{\partial^2 f}{\partial y_i \partial y_j}(s,y), \tag{7.14}$$

or, in the one-dimensional case:

$$Af(s,y) = b(s,y) \frac{\partial f}{\partial y}(s,y) + \frac{1}{2}\sigma^2(s,y) \frac{\partial^2 f}{\partial y^2}(s,y) \tag{7.15}$$

Consider the boundary value problem :

$$\left(\frac{\partial u}{\partial s} + Au\right)(s,y) = 0 \quad \text{if } (s,y) \in \,]0,T[\, \times \mathbb{R}^n, \tag{7.16}$$

$$u(T,y) = \mathbf{1}_C(y) \quad \text{if } y \in \mathbb{R}^n.$$

By the Feynman-Kac formula, we know that

$$u(s,y) = \mathbb{E}_{s,y}\left[\mathbf{1}_C(X_T)\right] = \mathbb{P}\left[X_T \in C | X_s = y\right] = P(C,T,y,s),$$

where

$$\begin{cases} dX_t = b(t,X_t)\,dt + \sigma(t,X_t)\,dB_t \\ X_s = y \end{cases}$$

and $P(C,T,y,s)$ are the transition probabilities associated to the Markov process X from time s to time T.

Theorem 7.9 *(Kolmogorov Backward Equation) Let X be a solution of (7.12). Then, the transition probabilities $P(C,t,y,s) = \mathbb{P}[X_t \in C | X_s = y]$ are solutions of*

$$\begin{cases} \left(\frac{\partial P}{\partial s} + AP\right)(C,t,s,y) = 0 \quad \text{if } (s,y) \in \,]0,t[\, \times \mathbb{R}^n, \\ P(C,t,y,t) = \mathbf{1}_C(y) \quad \text{if } y \in \mathbb{R}^n. \end{cases} \tag{7.17}$$

If the transition measure $P(dx, t, y, s)$ has a probability density function $\widetilde{f}(x, t, y, s)\,dx$, then $\widetilde{f}(x, t, y, s)$ is a solution of

$$\begin{cases} \left(\frac{\partial \widetilde{f}}{\partial s} + A\widetilde{f}\right)(x, t, s, y) = 0 \quad \text{if } (s, y) \in \left]0, t\right[\times \mathbb{R}^n, \\ \widetilde{f}(x, t, y, s) \longrightarrow \delta_x \quad \text{when } s \nearrow t. \end{cases} \tag{7.18}$$

These equations are called "backward" because the differential operator A applies to the "backward" variables (s, y) and not to the forward variables (x, t).

Consider the one dimensional case in order to have a simple notation. Let $s < T$ and let $h(t, x) \in C_c^\infty(\left]s, T\right[\times \mathbb{R})$ be a smooth function (test function) of compact support in $\left]s, T\right[\times \mathbb{R}$.

By the Itô formula, we have

$$h(T, X_T) = h(s, X_s) + \int_s^T \left(\frac{\partial h}{\partial t} + Ah\right)(t, X_t)\,dt + \int_s^T \frac{\partial h}{\partial x}(t, X_t)\,dB_t.$$

Applying the conditional expectation $\mathbb{E}_{s,y}[\cdot] = \mathbb{E}[\cdot | X_s = y]$, and using the fact that $h(T, x) = h(s, x) = 0$ (because $h(t, x)$ has compact support in $\left]s, T\right[\times \mathbb{R}$) and the zero mean property of the stochastic integral, we obtain

$$\int_{-\infty}^{+\infty} \int_s^T \left(\frac{\partial}{\partial t} + b(t, x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t, x)\frac{\partial^2}{\partial x^2}\right) \times h(t, x)\,\widetilde{f}(x, t, y, s)\,dtdx = 0.$$

If we integrate by parts with respect to t (for the $\frac{\partial}{\partial t}$ part) and by parts with respect to x (for the $\frac{\partial}{\partial x}$ and $\frac{\partial^2}{\partial x^2}$ parts), we obtain:

$$\int_{-\infty}^{+\infty} \int_s^T h(t, x)\left(-\frac{\partial}{\partial t}\widetilde{f}(x, t, y, s) - \frac{\partial}{\partial x}\left[b(t, x)\,\widetilde{f}(x, t, y, s)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma^2(t, x)\,\widetilde{f}(x, t, y, s)\right]\right)dtdx = 0.$$

This equation must hold for all test functions $h(t, x) \in C_c^\infty(\left]s, T\right[\times \mathbb{R})$, and therefore:

$$-\frac{\partial}{\partial t}\widetilde{f}(x, t, y, s) - \frac{\partial}{\partial x}\left[b(t, x)\,\widetilde{f}(x, t, y, s)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma^2(t, x)\,\widetilde{f}(x, t, y, s)\right] = 0.$$

Theorem 7.10 *(Kolmogorov Forward Equation): Let X be a solution of (7.12) with transition probability density function $\widetilde{f}(x,t,y,s)$. Then $\widetilde{f}$ satisfies the equation*

$$\begin{cases} \frac{\partial \widetilde{f}}{\partial t}(x,t,s,y) = A^{*}\widetilde{f}(x,t,y,s) \ \ if \ \ (t,x) \in \,]s,T[\times \mathbb{R}, \\ \widetilde{f}(x,t,y,s) \longrightarrow \delta_y \ \ when \ t \searrow s, \end{cases} \tag{7.19}$$

where the operator A^{*} is the adjoint operator of A and is defined by

$$(A^{*}f)(t,x) = -\frac{\partial}{\partial x}\left[b(t,x) f(t,x)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma^2(t,x) f(t,x)\right]. \tag{7.20}$$

In the multidimensional case, the Kolmogorov forward equation is

$$\frac{\partial \widetilde{f}}{\partial t}(x,t,s,y) = A^{*}\widetilde{f}(x,t,y,s) \ \text{ if } \ (t,x) \in \,]s,T[\times \mathbb{R}^n,$$

where the adjoint operator A^{*} is defined by

$$\begin{aligned} (A^{*}f)(t,x) = & -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left[b_i(t,x) f(t,x)\right] \\ & + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_i}\left[\left[\left[\sigma(t,x)\sigma^T(t,x)\right]_{i,j}\right] f(t,x)\right]. \end{aligned}$$

Note that in the forward equation, the adjoint operator applies to the "forward" variables (x,t).

Consider the stochastic differential equation

$$\begin{aligned} dX_t &= \sigma dB_t, \\ X_s &= y, \end{aligned}$$

where σ is a constant. The Fokker-Planck equation for this process is

$$\frac{\partial \widetilde{f}}{\partial t}(x,t,s,y) = \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}\left[\widetilde{f}(x,t,s,y)\right],$$

and the solution is given by the Gaussian probability density function

$$\widetilde{f}(x,t,s,y) = \frac{1}{\sigma\sqrt{2\pi(t-s)}} \exp\left[-\frac{(x-y)}{2\sigma^2(t-s)}\right].$$

Consider the stochastic differential equation for the geometric Brownian motion

$$\begin{aligned} dX_t &= \alpha X_t dt + \sigma X_t dB_t, \\ X_s &= y. \end{aligned}$$

The Fokker-Planck equation for this process is

$$\frac{\partial \widetilde{f}}{\partial t}(x,t,s,y) = \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}\left[x^2 \widetilde{f}(x,t,s,y)\right] - \alpha \frac{\partial}{\partial x}\left[x \widetilde{f}(t,x)\right],$$

or

$$\frac{\partial \widetilde{f}}{\partial t} = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 \widetilde{f}}{\partial x^2} + \left(2\sigma^2 - \alpha\right) x \frac{\partial \widetilde{f}}{\partial x} + \left(\sigma^2 - \alpha\right) \widetilde{f}.$$

Chapter 8

The Girsanov Theorem

The Girsanov theorem states, in it's simpler version, that the Brownian motion with drift: $\widetilde{B}_t = B_t + \lambda t$, may be seen as a standard Brownian motion if we change the probability measure. In a broader way, the theorem states that if we change the drift coefficient of an Itô process then the law of the process does not radically change. The law of the new Itô process will be absolutely continuous with respect to the law of the original process, and we may explicitly compute the Radon-Nikodym derivative.

8.1 Changing the probability measure

Suppose that $L \geq 0$ is a random variable of mean 1 defined in the probability space. $(\Omega, \mathcal{F}, P)$. Then

$$Q(A) = E[\mathbf{1}_A L]$$

defines a new probability measure. Clearly, $Q(\Omega) = E[L] = 1$. Moreover, $Q(A) = E[\mathbf{1}_A L]$ is equivalent to

$$\int_\Omega \mathbf{1}_A dQ = \int_\Omega \mathbf{1}_A L dP.$$

L is called the density of Q with respect to P and is written as

$$\frac{dQ}{dP} = L.$$

L is also said to be the Radon-Nikodym derivative of Q with respect to P.

The expected value of a r.v. X defined in the probability space $(\Omega, \mathcal{F}, P)$ is calculated by the formula

$$E_Q[X] = E[XL].$$

The probability measure Q is absolutely continuous with respect to P, which means that

$$P(A) = 0 \Longrightarrow Q(A) = 0.$$

If the random variable L is strictly prositive ($L > 0$), the probabilities P and Q are equivalent (that is, they're mutually absolutely continuous), which means that

$$P(A) = 0 \Longleftrightarrow Q(A) = 0.$$

8.2 Girsanov Theorem

Let X be a r.v. with distribution $N(m, \sigma^2)$. Is there a probability measure Q with respect to which X has distribution $N(0, \sigma^2)$?

Consider the r.v.

$$L = \exp\left(-\frac{m}{\sigma^2}X + \frac{m^2}{2\sigma^2}\right).$$

It is easily verified that $E[L] = 1$. It is enough to consider the density of the normal distribution $N(m, \sigma^2)$ and it follows that

$$\begin{aligned} E[L] &= \int_{-\infty}^{+\infty} \exp\left(-\frac{m}{\sigma^2}x + \frac{m^2}{2\sigma^2}\right) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = 1. \end{aligned}$$

Suppose that the probability measure Q has density L with respect to P. Then, in the probability space $(\Omega, \mathcal{F}, Q)$, the r.v. X has the characteristic function:

$$\begin{aligned} E_Q\left[e^{itX}\right] &= E\left[e^{itX}L\right] \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(itx - \frac{m}{\sigma^2}x + \frac{m^2}{2\sigma^2}\right) \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(itx - \frac{x^2}{2\sigma^2}\right) dx = e^{-\frac{\sigma^2 t^2}{2}}. \end{aligned}$$

Conclusion: X has distribution $N(0, \sigma^2)$.

The more general form for the characteristic function of a normal distribution may be found in the appendix of [10] on normal distributions.

Let $\{B_t, t \in [0,T]\}$ be a Brownian motion. Fix a real number λ and consider the martingale:

$$L_t = \exp\left(-\lambda B_t - \frac{\lambda^2}{2}t\right). \tag{8.1}$$

Example 8.1 *Prove that the stochastic process $\{L_t, t \in [0,T]\}$ is a positive martingale with expected value 1 and that satisfies the SDE:*

$$\begin{aligned} dL_t &= -\lambda L_t dB_t, \\ L_0 &= 1. \end{aligned}$$

The random variable $L_T = \exp\left(-\lambda B_T - \frac{\lambda^2}{2}T\right)$ is a density in the probability space $(\Omega, \mathcal{F}_T, P)$, for which the new probability measure is defined:

$$Q(A) = E[\mathbf{1}_A L_T],$$

for every $A \in \mathcal{F}_T$.

- As $\{L_t, t \in [0,T]\}$ is a martingale, then the r.v. $L_t = \exp\left(-\lambda B_t - \frac{\lambda^2}{2}t\right)$ is a density in the probability space $(\Omega, \mathcal{F}_t, P)$ and in this space the probability measure Q has precisely the density L_t.

- In fact, if $A \in \mathcal{F}_t$, we have:

$$\begin{aligned} Q(A) &= E[\mathbf{1}_A L_T] = E[E[\mathbf{1}_A L_T|\mathcal{F}_t]] \\ &= E[\mathbf{1}_A E[L_T|\mathcal{F}_t]] = E[\mathbf{1}_A L_t], \end{aligned}$$

  where the conditional expectation properties and the martingale property of $\{L_t, t \in [0,T]\}$ were applied.

Theorem 8.2 *(Girsanov Theorem I): On the probability space $(\Omega, \mathcal{F}_T, Q)$, where Q is defined by $Q(A) = E[\mathbf{1}_A L_T]$, the stochastic process*

$$\widetilde{B}_t = B_t + \lambda t$$

is a Brownian motion

Before proving this theorem, we need the following lemma:

Lema 8.3 *Suppose X is a real r.v. and that $\mathcal{G}$ is a σ-algebra such that:*

$$E\left[e^{iuX}|\mathcal{G}\right] = e^{-\frac{u^2\sigma^2}{2}}.$$

Then the random variable X is independent from the σ-algebra $\mathcal{G}$ and has normal distribution $N(0, \sigma^2)$.

The proof of the above lemma may be found in [9], pgs. 63-64.

Proof. (Girsanov theorem) It suffices to show that in $(\Omega, \mathcal{F}_T, Q)$, the increment $\widetilde{B}_t - \widetilde{B}_s$, with $s < t \leq T$, is independent from $\mathcal{F}_s$ and has normal distribution $N(0, t-s)$. Taking into account the previous lemma, the result follows from the relation:

$$E_Q\left[\mathbf{1}_A e^{iu\left(\widetilde{B}_t - \widetilde{B}_s\right)}\right] = Q(A)\, e^{-\frac{u^2}{2}(t-s)}, \tag{8.2}$$

for all $s < t$, $A \in \mathcal{F}_s$ and $u \in \mathbb{R}$. In fact, if (8.2) is verified, then, from the definition of conditional expectation and the previous lemma, $\left(\widetilde{B}_t - \widetilde{B}_s\right)$ is independent from $\mathcal{F}_s$ and has normal distribution $N(0, t-s)$.

Proof of the equality (8.2):

$$\begin{aligned}
&E_Q\left[\mathbf{1}_A e^{iu\left(\widetilde{B}_t - \widetilde{B}_s\right)}\right] = E\left[\mathbf{1}_A e^{iu\left(\widetilde{B}_t - \widetilde{B}_s\right)} L_t\right] \\
&= E\left[\mathbf{1}_A e^{iu(B_t - B_s) + iu\lambda(t-s) - \lambda(B_t - B_s) - \frac{\lambda^2}{2}(t-s)} L_s\right] \\
&= E\left[\mathbf{1}_A L_s\right] E\left[e^{(iu-\lambda)(B_t - B_s)}\right] e^{iu\lambda(t-s) - \frac{\lambda^2}{2}(t-s)} \\
&= Q(A) e^{\frac{(iu-\lambda)^2}{2}(t-s) + iu\lambda(t-s) - \frac{\lambda^2}{2}(t-s)} \\
&= Q(A) e^{-\frac{u^2}{2}(t-s)},
\end{aligned}$$

Where the definition of E_Q and L_t, independence of $(B_t - B_s)$ from L_s and A, and the definition of Q were used. ■

8.3 Girsanov Theorem - general version

Theorem 8.4 *(Girsanov Theorem II): Let $\{\theta_t, t \in [0,T]\}$ be an adapted stochastic process that satisfies the Novikov condition:*

$$E\left[\exp\left(\frac{1}{2}\int_0^T \theta_t^2 dt\right)\right] < \infty. \tag{8.3}$$

Then, the stochastic process

$$\widetilde{B}_t = B_t + \int_0^t \theta_s ds$$

is a Brownian motion with respect to the measure Q defined by $Q(A) = E[\mathbf{1}_A L_T]$, where

$$L_t = \exp\left(-\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right).$$

Note that L_t satisfies the linear SDE

$$L_t = 1 - \int_0^t \theta_s L_s dB_s.$$

It is necessary, for the process L_t to be a density, that $E[L_t] = 1$. However, condition (8.3) is sufficient to guarantee that it is in fact verified. The second version of the Girsanov theorem generalizes the first: note that, taking $\theta_t \equiv \lambda$, we obtain the previous version.

Chapter 9

Models for Financial Markets

9.1 The Black-Scholes model

The differential equations that define the Black-Scholes model are

$$dB\left(t\right) = rB\left(t\right)dt, \tag{9.1}$$

$$dS_t = \alpha S_t dt + \sigma S_t d\overline{W}_t, \tag{9.2}$$

where r, α and σ are constant. Represent by $B\left(t\right)$ the deterministic price of a riskless asset (a bond or a bank deposit), and by S_t the (stochastic) process of the price of a risky asset (a stock or an index). Consider $\overline{W}_t$ as a standard Brownian motion with respect to the original probability measure P, the risk-free interest rate r, the mean appreciation rate and the volatility of the risky asset α and σ, respectively.

We already know that the solution of (9.2) is the geometric Brownian motion:

$$S_t = S_0 \exp\left(\left(\alpha - \frac{1}{2}\sigma^2\right)t + \sigma\overline{W}_t\right).$$

Consider a contingent claim (e.g.: a financial derivative), with payoff given by

$$\chi = \Phi\left(S\left(T\right)\right). \tag{9.3}$$

Assume that this derivative may be traded in the market and that it's price process is given by

$$\Pi\left(t\right) = F\left(t, S_t\right), \quad t \in \left[0, T\right], \tag{9.4}$$

where F is a differentiable function of class $C^{1,2}$. Applying Itô's formula to (9.17) and considering (9.2), we get

$$dF(t,S_t) = \left(\frac{\partial F}{\partial t}(t,S_t) + \alpha S_t \frac{\partial F}{\partial x}(t,S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial x^2}(t,S_t)\right) dt + \left(\sigma S_t \frac{\partial F}{\partial x}(t,S_t)\right) d\overline{W}_t.$$

That is,

$$F(t,S_t) = F(0,S_0) + \int_0^t \left(\frac{\partial F}{\partial t}(r,S_r) + AF(r,S_r)\right) dr + \int_0^t \left(\sigma S_r \frac{\partial F}{\partial x}(r,S_r)\right) d\overline{W}_r,$$

where

$$Af(t,x) = \alpha x \frac{\partial f}{\partial x}(t,x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2}(t,x)$$

is the infinitesimal operator associated to the diffusion S_t that has the dynamics (9.2). We may also write

$$d\Pi(t) = \alpha_\Pi(t)\Pi_t dt + \sigma_\Pi(t)\Pi_t d\overline{W}_t, \quad (9.5)$$

where

$$\alpha_\Pi(t) = \frac{\left(\frac{\partial F}{\partial t}(t,S_t) + \alpha S_t \frac{\partial F}{\partial x}(t,S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial x^2}(t,S_t)\right)}{F(t,S_t)}, \quad (9.6)$$

$$\sigma_\Pi(t) = \frac{\sigma S_t \frac{\partial F}{\partial x}(t,S_t)}{F(t,S_t)}. \quad (9.7)$$

Consider a portfolio (a_t, b_t), where:

- a_t is the number of stocks (or units of the risky asset) in the portfolio at instant t.
- b_t is the number of bonds (or units of the riskless asset) in the portfolio at instant t.

Both a_t and b_t can be negative. If so, it means we are taking a short position on the respective asset in that instant.

The value of the portfolio at instant t is given by

$$V(t) = a_t S_t + b_t B_t.$$

It is supposed that the portfolio is self-financed, that is, any variation on the value of the portfolio is only due to price changes of the assets, so cash infusion or withdrawal is not allowed. Mathematically, this can be written as

$$dV_t = a_t dS_t + b_t dB_t.$$

We can also consider a portfolio with two other assets: the risky asset and the derivative with the same underlying asset. Let $u_S(t)$ and $u_\Pi(t)$ be the relative quantities of each of these assets in the portfolio, so that $u_S(t) + u_\Pi(t) = 1$. The dynamics for the value of the portfolio (which is also assumed self-financed) are described by

$$dV_t = u_S(t) V_t \frac{dS_t}{S_t} + u_\Pi(t) V_t \frac{d\Pi_t}{\Pi_t}.$$

Substituting (9.2) and (9.5), we obtain

$$\begin{aligned} dV_t &= V_t \left[u_S(t)\alpha + u_\Pi(t)\alpha_\Pi(t)\right] dt \\ &\quad + V\left[u_S(t)\sigma + u_\Pi(t)\sigma_\Pi(t)\right] d\overline{W}_t. \end{aligned}$$

9.2 No-arbitrage principle and the Black-Scholes equation

We shall define the portfolio $(u_S(t), u_\Pi(t))$ in such a way so that the stochastic part of dV_t is zero. Let $u_S(t), u_\Pi(t)$ be solutions of the system of linear equations

$$\begin{cases} u_S(t) + u_\Pi(t) = 1, \\ u_S(t)\sigma + u_\Pi(t)\sigma_\Pi(t) = 0. \end{cases}$$

This system has as solution:

$$\begin{aligned} u_S(t) &= \frac{\sigma_\Pi(t)}{\sigma_\Pi(t) - \sigma}, \\ u_\Pi(t) &= \frac{-\sigma}{\sigma_\Pi(t) - \sigma}. \end{aligned}$$

Substituting (9.7) on the expressions above, we get:

$$u_S(t) = \frac{S_t \frac{\partial F}{\partial x}(t, S_t)}{S_t \frac{\partial F}{\partial x}(t, S_t) - F(t, S_t)}, \tag{9.8}$$

$$u_\Pi(t) = \frac{-F(t, S_t)}{S_t \frac{\partial F}{\partial x}(t, S_t) - F(t, S_t)}. \tag{9.9}$$

With this portfolio we have (value of the portfolio without a stochastic differential):

$$dV_t = V_t \left[u_S(t)\alpha + u_\Pi(t)\alpha_\Pi(t)\right] dt. \tag{9.10}$$

An arbitrage opportunity on a financial market is defined as a self-financed portfolio h such that:

$$V^h(0) = 0,$$
$$V^h(T) > 0 \quad q.c.$$

This means that an arbitrage opportunity is the possibility of obtaining a positive profit from no investment, with probability 1, i.e., with no risk involved.

The no-arbitrage principle simply states that, given a derivative with price $\Pi(t)$, we consider that $\Pi(t)$ is such that there are no arbitrage opportunities in the market.

Proposition 9.1 *If a self-financed portfolio h is such that the portfolio value has the dynamics*

$$dV^h(t) = k(t) V^h(t) dt,$$

where $k(t)$ is an adapted process, then we must have $k(t) = r$ for all t, or otherwise arbitrage opportunities exist.

More details on the no-arbitrage principle may be found in [1].

By the no-arbitrage principle we have, from (9.10), that

$$u_S(t)\alpha + u_\Pi(t)\alpha_\Pi(t) = r \tag{9.11}$$

Substituting (9.6), (9.8) and (9.9) in the no arbitrage condition (9.11), we get

$$\frac{\partial F}{\partial t}(t, S_t) + rS_t \frac{\partial F}{\partial x}(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial x^2}(t, S_t) - rF(t, S_t) = 0.$$

Furthermore, it is clear that in the maturity date of the derivative we have

$$\Pi(T) = F(T, S_T) = \Phi(S(T)) \tag{9.12}$$

So we may state the following theorem.

Theorem 9.2 *(Black-Scholes eq.) Assume that the market is specified by eqs. (9.1)-(9.2) and we want to price a derivative with payoff given by (9.3). Then, the only price function of the form (9.17) that is consistent with the principle of no arbitrage is the solution F of the following boundary values problem, defined in the domain $[0,T]\times\mathbb{R}^+$:*

$$\frac{\partial F}{\partial t}(t,x)+rx\frac{\partial F}{\partial x}(t,x)+\frac{1}{2}\sigma^2x^2\frac{\partial^2 F}{\partial x^2}(t,x)-rF(t,x)=0, \qquad (9.13)$$
$$F(T,x)=\Phi(x).$$

Note that, in order to determine the Black-Scholes equation (9.13), we need to assume that the derivative price takes the form $\Pi(t)=F(t,S_t)$ and that there exists a market for the derivative to be traded. However it is not unusual for derivatives to be traded "over the counter" (OTC), so it is not always the case. To solve this problem, we shall se how may we obtain the same equation (9.13) without those hypothesis.

Consider the portfolio $(h^0(t),h^*(t))$ where $h^0(t)$ is the number of bonds (or riskless asset units) and $h^*(t)$ is the number of shares at instant t. The value of the portfolio at instant t is

$$V^h(t)=h^0(t)B_t+h^*(t)S_t.$$

It is supposed that the portfolio is self-financed, that is,

$$dV_t^h=h^0(t)dB_t+h^*(t)dS_t.$$

In integral form,

$$\begin{aligned}V_t^h&=V_0+\int_0^t h^*(s)dS_s+\int_0^t h^0(s)dB_s\\&=V_0+\int_0^t\left(\alpha h^*(s)S_s+rh^0(s)B_s\right)ds+\sigma\int_0^t h^*(s)S_sd\overline{W}_s. \qquad (9.14)\end{aligned}$$

Assume that the contingent claim (or financial derivative) has the payoff

$$\chi=\Phi(S(T)). \qquad (9.15)$$

and it is replicated by the portfolio $h=(h^0(t),h^*(t))$, that is, assume that $V_T^h=\chi=\Phi(S(T))$ a.s. Then, the unique price process that is compatible with the no-arbitrage principle is

$$\Pi(t)=V_t^h,\quad t\in[0,T]. \qquad (9.16)$$

Moreover, suppose that

$$\Pi(t) = V_t^h = F(t, S_t). \tag{9.17}$$

where F is a differentiable function of class $C^{1,2}$. By applying Itô's formula to (9.17) and considering (9.2), we obtain

$$\begin{aligned} dF(t,S_t) = {} & \left(\frac{\partial F}{\partial t}(t,S_t) + \alpha S_t \frac{\partial F}{\partial x}(t,S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial x^2}(t,S_t)\right) dt \\ & + \left(\sigma S_t \frac{\partial F}{\partial x}(t,S_t)\right) d\overline{W}_t. \end{aligned}$$

That is,

$$\begin{aligned} F(t,S_t) = {} & F(0,S_0) + \int_0^t \left(\frac{\partial F}{\partial t}(s,S_s) + AF(s,S_s)\right) ds \\ & + \int_0^t \left(\sigma S_s \frac{\partial F}{\partial x}(s,S_s)\right) d\overline{W}_s, \end{aligned} \tag{9.18}$$

where

$$Af(t,x) = \alpha x \frac{\partial f}{\partial x}(t,x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2}(t,x)$$

is the infinitesimal generator associated to the diffusion S_t that has the dynamics (9.2). Comparing (9.14) and (9.18), we have that

$$\begin{aligned} \sigma h^*(s) S_s &= \sigma S_s \frac{\partial F}{\partial x}(s,S_s), \\ \alpha h^*(s) S_s + r h^0(s) B_s &= \frac{\partial F}{\partial t}(s,S_s) + AF(s,S_s). \end{aligned}$$

Hence

$$\begin{aligned} \frac{\partial F}{\partial x}(s,S_s) &= h^*(s), \\ \frac{\partial F}{\partial t}(s,S_s) + rS_s \frac{\partial F}{\partial x}(s,S_s) + \frac{1}{2}\sigma^2 S_s^2 \frac{\partial^2 F}{\partial x^2}(s,S_s) - rF(s,S_s) &= 0. \end{aligned}$$

Therefore we have:

- A portfolio h with value $V_t^h = F(t,S_t)$, composed of risky assets with price S_t and riskless assets of price B_t.

- Portfolio h replicates the contingent claim χ in each instant t, and

$$\Pi(t) = V_t^h = F(t,S_t).$$

- In particular:

$$F(T, S_T) = \Phi\left(S\left(T\right)\right) = \text{Payoff}.$$

The portfolio should be continuously updated by acquiring (or selling) $h^*(t)$ shares of the risky asset and $h^0(t)$ units of the riskless asset, where

$$\begin{aligned} h^*(t) &= \frac{\partial F}{\partial x}(t, S_t), \\ h^0(t) &= \frac{V_t^h - h^*(t) S_t}{B_t} = \frac{F(t, S_t) - h^*(t) S_t}{B_t}. \end{aligned}$$

The derivative price function satisfies the partial differential equation (Black-Scholes eq.)

$$\frac{\partial F}{\partial t}(t, S_t) + r S_t \frac{\partial F}{\partial x}(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial x^2}(t, S_t) - rF(t, S_t) = 0.$$

Theorem 9.3 *(Black-Scholes eq.) Suppose that the market is specified by eqs. (9.1)-(9.2) and we want to price a derivative with payoff (9.3). Then, the only pricing function that is consistent with the no-arbitrage principle is the solution F of the following boundary value problem, defined in the domain $[0, T] \times \mathbb{R}^+$:*

$$\begin{aligned} \frac{\partial F}{\partial t}(t, x) + r x \frac{\partial F}{\partial x}(t, x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 F}{\partial x^2}(t, x) - rF(t, x) &= 0, \qquad (9.19) \\ F(T, x) &= \Phi(x). \end{aligned}$$

The Black-Scholes equation may be solved analytically or with probabilistic methods. By applying Feynman-Kac formula, we have the following result.

Proposition 9.4 *(Feynman-Kac formula) Let F be a solution of the boundary values problem*

$$\begin{aligned} \frac{\partial F}{\partial t}(t, x) + \mu(t, x) \frac{\partial F}{\partial x}(t, x) + \frac{1}{2}\sigma^2(t, x) \frac{\partial^2 F}{\partial x^2}(t, x) - rF(t, x) &= 0, \qquad (9.20) \\ F(T, x) &= \Phi(x). \end{aligned}$$

Assume that $\sigma(s, X_s) \frac{\partial F}{\partial x}(s, X_s)$ is a process in L^2 (i.e. $E \int_0^t \left(\frac{\partial F}{\partial x}(s, X_s)\, \sigma(s, X_s)\right)^2 ds < \infty$). Then,

$$F(t, x) = e^{-r(T-t)} E_{t,x}\left[\Phi(X_T)\right],$$

where X satisfies

$$\begin{aligned} dX_s &= \mu(s, X_s)\, ds + \sigma(s, X_s)\, dB_s, \\ X_t &= x. \end{aligned}$$

Applying the Feynman-Kac formula from the previous proposition to the eq. (9.19), we obtain:

$$F(t,x) = e^{-r(T-t)} E_{t,x}\left[\Phi(X_T)\right], \tag{9.21}$$

where X is a stochastic process with dynamics:

$$\begin{aligned} dX_s &= rX_s ds + \sigma X_s d\overline{W}_s, \\ X_t &= x. \end{aligned} \tag{9.22}$$

Note that the process X is not the same as the process S, as the drift of X is rX and not αX. This means that S has an mean appreciation rate α, while X has the appreciation rate the risk-free interest rate r. To pass from process X to the process S, we'll apply the Girsanov theorem.

9.3 The martingale measure and risk-neutral valuation

Denote by P the original probability measure ("objective" or "real" probability measure). The P-dynamics of the process S is given in (9.2). Note that (9.2) is equivalent to

$$\begin{aligned} dS_t &= rS_t dt + \sigma S_t \left(\frac{\alpha - r}{\sigma} dt + d\overline{W}_t \right) \\ &= rS_t dt + \sigma S_t d \underbrace{\left(\frac{\alpha - r}{\sigma} t + \overline{W}_t \right)}_{W_t}. \end{aligned}$$

By the Girsanov Theorem, there exists a probability measure Q such that, in the probability space $(\Omega, \mathcal{F}_T, Q)$, the process

$$W_t := \frac{\alpha - r}{\sigma} t + \overline{W}_t$$

is a Brownian motion, and S has the Q-dynamics:

$$dS_t = rS_t dt + \sigma S_t dW_t. \tag{9.23}$$

Now consider the following notation: E denotes the expected value with respect to the original measure P, while E^Q denotes the expected value with respect to the new probability measure Q (that comes from the application of the Girsanov theorem). Also, let $\overline{W}_t$ denote the original Brownian motion (under the measure P) and W_t denote the Brownian motion under the measure Q.

Getting back to (9.21) and (9.22), and taking into account that under the measure Q the equations (9.22) and (9.23) are the same, we may represent the solution of the Black-Scholes equation by

$$F(t,s) = e^{-r(T-t)} E^Q_{t,s}\left[\Phi(S_T)\right],$$

where the dynamics of S under the measure Q is

$$dS_t = rS_t dt + \sigma S_t dW_t.$$

We may finally state the theorem that provides us a pricing formula for the contingent claim in terms of the new measure Q.

Theorem 9.5 *The price (absent of arbitrage) of the contingent claim $\Phi(S_T)$ is given by the formula*

$$F(t,S_t) = e^{-r(T-t)} E^Q_{t,s}\left[\Phi(S_T)\right], \tag{9.24}$$

where the dynamics of S under the measure Q is

$$dS_t = rS_t dt + \sigma S_t dW_t.$$

In the Black-Scholes, the diffusion coefficient σ may depend on t and S - be a function $\sigma(t, S_t)$ - and in this case, the calculations needed would be analogous to the ones we've done.

The measure Q is called equivalent martingale measure. The reason for this nomenclature has to do with the fact that the discounted process

$$\widetilde{S}_t := \frac{S_t}{B_t}$$

is a Q-martingale (martingale under the measure Q). In fact,

$$\begin{aligned}\widetilde{S}_t &= \frac{S_t}{B_t} = e^{-rt}S_t = e^{-rt}S_0 \exp\left(\left(\alpha - \frac{1}{2}\sigma^2\right)t + \sigma\overline{W}_t\right)\\ &= S_0 \exp\left(-\frac{1}{2}\sigma^2 t + \sigma W_t\right)\end{aligned}$$

is a martingale.

9.4 The Black-Scholes formula

Computing explicitly the price of the derivative, we have

$$\begin{aligned} e^{r(T-t)}F(t,s) &= E^Q_{t,s}\left[\Phi(S_T)\right] \\ &= E^Q_{t,s}\left[\Phi\left(s\exp\left(\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma\left(W_T-W_t\right)\right)\right)\right] \\ &= E^Q\left[\Phi\left(se^Z\right)\right], \end{aligned}$$

where $Z=\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma\left(W_T-W_t\right)\sim N\left(\left(r-\frac{1}{2}\sigma^2\right)(T-t),\sigma^2(T-t)\right).$ Therefore,

$$F(t,s)=e^{-r(T-t)}\int_{-\infty}^{+\infty}\Phi\left(se^y\right)f(y)\,dy, \tag{9.25}$$

where f is the density of the gaussian random variable Z. The integral formula (9.25), for a given function Φ, should be, in the general case, computed using numerical methods. However, there are some particular cases where (9.25) may be obtained analytically. For example, for a European "call" option with payoff

$$\Phi(x)=(x-K)^+=\max(x-K,0),$$

we have

$$\begin{aligned} F(t,s) &= e^{-r(T-t)}\int_{-\infty}^{+\infty}\max\left(se^y-K,0\right)f(y)\,dy \\ &= e^{-r(T-t)}\int_{\ln(K/s)}^{+\infty}\left(se^y-K\right)f(y)\,dy \\ &= e^{-r(T-t)}\left(s\int_{\ln(K/s)}^{+\infty}e^yf(y)\,dy-K\int_{\ln(K/s)}^{+\infty}f(y)\,dy\right) \end{aligned} \tag{9.26}$$

and the first integral may be computed in the following way:

$$\int_{\ln(K/s)}^{+\infty} e^y f(y)\,dy =$$

$$= \int_{\ln(K/s)}^{+\infty} \frac{\exp\left(y - \frac{\left(y-\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}\right)}{\sigma\sqrt{2\pi(T-t)}} dy$$

$$= \int_{\ln(K/s)}^{+\infty} \frac{\exp\left(\frac{2\sigma^2(T-t)y-\left(y-\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}\right)}{\sigma\sqrt{2\pi(T-t)}} dy$$

$$= e^{r(T-t)} \int_{\ln(K/s)}^{+\infty} \frac{\exp\left(-\frac{\left(y-\left(r+\frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}\right)}{\sigma\sqrt{2\pi(T-t)}} dy$$

But $\frac{1}{\sigma\sqrt{2\pi(T-t)}} \exp\left(-\frac{\left(y-\left(r+\frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}\right)$ is the density function of a random variable Z^* with distribution $N\left(\left(r+\frac{1}{2}\sigma^2\right)(T-t), \sigma^2(T-t)\right)$ and so

$$\int_{\ln(K/s)}^{+\infty} e^y f(y)\,dy = e^{r(T-t)} Q\left(Z^* \geq \ln(K/s)\right)$$

$$= e^{r(T-t)} Q\left(\overline{Z} \geq \frac{\ln(K/s) - \left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}\right)$$

$$= e^{r(T-t)} Q\left(\overline{Z} \leq \frac{\ln(s/K) + \left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}\right)$$

$$= e^{r(T-t)} N\left[d_1(t,s)\right],$$

where $N[x]$ is the cumulative distribution function of the distribution $N(0,1)$ and

$$d_1(t,s) = \frac{\ln(s/K) + \left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}.$$

The second integral may be calculated as follows.

$$\int_{\ln(K/s)}^{+\infty} f(y)\,dy = Q\left(Z \geq \ln(K/s)\right)$$

$$= Q\left(\overline{Z} \geq \frac{\ln(K/s) - \left(r-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}\right)$$

$$= Q\left(\overline{Z} \leq \frac{\ln(s/K) + \left(r-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}\right) = N\left[d_2(t,s)\right],$$

where $N[x]$ is the cumulative distribution function of the distribution $N(0,1)$ and

$$d_2(t,s) = \frac{\ln(s/K) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{(T-t)}}$$

Returning to eq. (9.26), we obtain

$$\begin{aligned} F(t,s) &= e^{-r(T-t)}\left(se^{r(T-t)}N[d_1(t,s)] - KN[d_2(t,s)]\right) \\ &= sN[d_1(t,s)] - e^{-r(T-t)}KN[d_2(t,s)] \end{aligned}$$

and this is the well known Black-Scholes Formula:

$$F(t,s) = sN[d_1(t,s)] - e^{-r(T-t)}KN[d_2(t,s)].$$

Exercise 9.6 *Deduce the Black-Scholes Formula form (9.24) for an European put option (with payoff $\Phi(S_T) = \max(K - S_T, 0)$).*

Bibliography

[1] Tomas Björk, Arbitrage Theory in Continuous Time, 2nd Edition. Oxford University Press, 2004.

[2] Zdzislaw Brzezniak and Tomasz Zastawniak, Basic Stochastic Processes. Springer, 1999.

[3] Marek Capinski and Ekkehard Kopp, Measure, Integral and Probability. Springer, 1998.

[4] Robert Jarrow and Philip Protter, A short history of stochastic integration and mathematical finance: The early years, 1880–1970. IMS Lecture Notes Monograph, vol. 45 (2004), pages 1–17.

[5] Ioannis Karatzas and Steven Shreve, Brownian Motion and Stochastic Calculus, 2nd edition. Springer, 1991.

[6] Peter E. Kloeden and Platen, Numerical Solution of Stochastic Differential Equations. Springer, 1992.

[7] Thomas Mikosch, Elementary Stochastic Calculus with Finance in view. World Scientific, 1998.

[8] Daniel Müller, Probabilidade e processos estocásticos: uma abordagem rigorosa com vista aos modelos em finanças. Coimbra, Almedina, 2011.

[9] David Nualart, Stochastic Calculus (Lecture notes, Kansas University), http://www.math.ku.edu/~nualart/StochasticCalculus.pdf, 2008.

[10] Bernt Oksendal, Stochastic Differential Equations, 6th Edition. Springer, 2003.

[11] Daniel Revuz and Marc Yor, Continuous martingales and Brownian motion. Third Edition. Springer, 1999.

[12] Steven Shreve, Stochastic Calculus for Finance II: Continuous-Time Models. Springer, 2004.