

**MINIMAL VARIANCE HEDGING IN A
DISCRETE TIME MARKET DRIVEN BY
MARKOV PROCESS**

A THESIS SUBMITTED TO THE UNIVERSITY OF ZIMBABWE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN THE FACULTY OF SCIENCE

Author: Farai Julius Mhlanga
Supervisor: Prof. G. Di Nunno
Department: Mathematics
Date of Submission: 30 June 2005

Contents

Declaration	5
1 Preliminaries	13
1.1 Definitions and basic properties	13
1.2 Stopping Times	16
1.3 The Poisson Process	16
1.4 Martingales	22
1.4.4 Orthogonal Decomposition	26
1.5 Quadratic variation	28
1.6 Predictable Processes	29
2 Techniques in stochastic analysis	31
2.1 The Itô Calculus	31
2.1.5 Properties of the Itô Integral	35
2.2 The Itô formula	39

2.3	Stochastic Integral Representations	43
2.4	The Cauchy Criterion for Convergence	45
2.5	Stochastic Derivatives	45
3	Methods in quadratic hedging	57
3.1	Discrete-Time Economy	58
3.1.1	Trading strategies and the value processes	59
3.1.2	The cost process	60
3.1.3	Self-financing portfolio	60
3.1.6	Attainability and Completeness	61
3.2	Risk-minimising strategies	62
3.2.3	Mean Self-financing Strategies	63
3.3	Quadratic Hedging Approaches	66
3.3.1	Local risk Minimisation	66
3.3.2	Mean Variance Hedging	70
3.3.3	Minimal Variance Hedging	73
4	Applications	81
5	Markov processes	85
5.1	Markov Models	85

5.2	Markovian case: Local risk Minimisation	88
5.3	Markovian case: Mean Variance Hedging	89
5.4	Markovian case: Minimal variance hedging	91
6	Conclusion	95

Abstract

Techniques in stochastic analysis are presented in a continuous time framework. We then review methods in quadratic hedging approaches with focus on minimal variance hedging in a discrete time framework. We also consider specific exercises. We then relate the results obtained in quadratic hedging methods to the case of a discrete time market driven by a Markov process.

Declaration

No portion of the work in this thesis has been submitted for another degree or qualification of this or any other university or another institution of learning.

Acknowledgement

I would like to extend special thanks to my supervisor, Prof G Di Nunno, who took her time from her busy schedules to review now and then my manuscript. She provided invaluable suggestions and critical guidance throughout the project. She pored over the manuscript at every step of the process, provided truly innovative ideas and continually encouraged me to add value here, to update there and to clarify this discussion, what a supervisor she was. Once more, thanks.

My colleagues together with the Mathematics department staff of the University of Zimbabwe also deserve a recognition. Without their intellectual stimulation and support, my progression would suffer greatly.

Lastly, my thanks go to my relatives and friends, for they keep on encouraging me.

Dedication

To my Parents

Introduction

A problem that arises in many contexts is the problem of predicting the value of a given random variable that will be observed in the future. The newspapers often contain predictions of future values of Gross National Product (G.N.P), employment, consumer prices and other economic quantities. If the predicted output of a process is not satisfactory, a control or change is implemented to bring the process more in line with what is desired.

We consider a financial market model which consists of a bond of price (discounted) $X_0 = 1$ and d -risky assets of price process (normalized) $X_t^i = \frac{S^i}{S^0}$, for $i = 1, \dots, d$ where X_0 is the bond price and X_t^i the risky asset price process. The latter is driven by a d -dimensional Brownian motion.

In this project we are concerned with hedging of contingent claim $H(\omega)$ by means of dynamic trading strategies based on X . Hedging is whereby one insures or covers himself against random loss at time T . A contingent claim here is simply an \mathcal{F}_T -measurable random variable $H(\omega)$ describing the net payoff at time T of some financial instrument. Let the hedging strategy be of the form $(\theta, \eta) = (\theta_t, \eta_t)_{0 \leq t \leq T}$ where θ is a d -dimensional predictable process and η is adapted. In such a hedging strategy, θ_t^i represents the number of shares held at time t and η_t is the t -value of the bond. Thus the value of this strategy $\varphi_t = (\theta_t, \eta_t)$ is given by

$$V_t = \eta_t + \theta_t X_t$$

The question that naturally appears in the mind of a financial dealer is whether the claim $H(\omega)$ is achievable on the market. Now if the market is complete, then

$$V_T = H(\omega)$$

for every contingent claim $H(\omega)$. If the market is not complete, then there exist some claims such that $V_T < H(\omega)$. That is, perfect hedging is not always possible. Suppose that there exists a contingent claim $\hat{H}(\omega) < H(\omega)$ such that $V_T = \hat{H}(\omega)$ then we want to find θ_t such that $\hat{H}(\omega)$ is the “closest” to $H(\omega)$ in some sense, that is

$$\mathbb{E}[(H - \hat{H})^2]$$

is minimal. That is

$$\mathbb{E}[(H - V_T)^2]$$

is minimal.

Therefore we want to find θ which minimises

$$\mathbb{E}[(H - V_T)^2]$$

Here if we consider “closest” in terms of minimal variance, then we talk about “minimal variance hedging strategy”.

The chapters are structured as follows: Chapter one introduces some basic definitions and properties and recalls a few preliminaries to compliment the preceding discussion. The second chapter is devoted to the study of the techniques in stochastic analysis with respect to stochastic differentiation in the non-anticipating stochastic framework. This, however, is discussed in continuous time. The third chapter focusses on the methods in quadratic hedging approaches in discrete time with special emphasis on minimal variance hedging approaches. The fourth chapter deals with some specific exercises. The last chapter exposes the Markov process. More precisely, we relate the results obtained in the study of methods of quadratic approaches to the case of a discrete time market driven by a Markov process.

Chapter 1

Preliminaries

1.1 Definitions and basic properties

Let Ω be a set. A collection of \mathcal{F} of subsets of Ω is called a σ -field or σ -algebra if it contains the empty set and is closed under the operations of countable unions and compliments. The pair (Ω, \mathcal{F}) is called a measurable space. The elements of Ω are called the samples and those of \mathcal{F} are called the events. We will consider the works in [2], [6], [14], [17].

If P is taken to be σ -additive measure on (Ω, \mathcal{F}) , then it is called a probability if $P(\Omega) = 1$. The triple (Ω, \mathcal{F}, P) is called a probability space. The probability space is called complete if \mathcal{F} contains all subsets G of Ω with P -outer measure zero, that is with

$$P^*(G) = \inf\{P(F) : F \in \mathcal{F}, G \subseteq F\} = 0$$

Any probability space can be made complete by adding to \mathcal{F} all sets of outer measure zero and by extending P accordingly.

A real valued measurable function $X(\omega)$ defined on (Ω, \mathcal{F}) is called a random

variable. The random variable $X(\omega)$ may take values $\pm\infty$, but we assume that $X(\omega)$ takes finite values for almost all ω unless otherwise mentioned. We often suppress the sample ω and write it as X . If we can define the integral of X by the measure P , we denote it by $\mathbb{E}[X]$ and call it the expectation of X , a random variable, with respect to P , that is

$$\mathbb{E}[X] = \int_{\Omega} X(t) dP(\omega)$$

If the expectation $\mathbb{E}[X]$ exists then the process $m(t) = \mathbb{E}[X(t, \omega)]$ of a measurable process is also measurable. In this case the expectation is taken with respect to P on (Ω, \mathcal{F}) . If $A \in \mathcal{B}(\mathbb{R})$ and $\int_A \mathbb{E}[|X(t)|] dt < \infty$ then

$$\int_A |X(t)| dt < \infty$$

P a.s and

$$\int_A \mathbb{E}[X(t)] dt = \mathbb{E}\left[\int_A X(t) dt\right]$$

The subsets of Ω which belong to \mathcal{F} are called \mathcal{F} -measurable sets.

Definition 1.1.1 A “stochastic process, X ”, defined on a probability space (Ω, \mathcal{F}, P) , is a family $X = \{X_t = X(t, \omega), 0 \leq t < \infty, \omega \in \Omega\}$ of random variables.

In practice we note that the index $t \in [0, \infty)$ of the random variable X_t is usually regarded as time. For a fixed sample point $\omega \in \Omega$ the function $t \rightarrow X(t, \omega)$ is a sample path (trajectory) of the process X associated with ω . It provides the mathematical model for a random experiment whose outcome can be observed and recorded over a period of time. We also note that for a fixed time $t \in [0, \infty)$ the function $\omega \rightarrow X(t, \omega)$ is a random variable.

Definition 1.1.2 If we consider a stochastic process X_t , we say “ X_t is measurable” if for all $A \in \mathcal{B}(\mathbb{R})$, the set

$$\{(t, \omega); X(t, \omega) \in A\}$$

belongs to the σ -fields $\{\mathcal{B}([0, \infty)) \times \mathcal{F}\}$

Due to the fact that a stochastic process is defined on an interval of time, then we can talk of its ‘past’, ‘present’ and ‘future’ and we can always ask how much an observer of such a process knows about it at present as compared to how much he knew it at some point in the past or will know it at some point in the future. This information is modelled by a flow of σ -fields called filtration.

Definition 1.1.3 *Let $\{\mathcal{F}_t, t \geq 0\}$ be a family of sub σ -fields of \mathcal{F} . It is called a “filtration” of sub σ -fields of \mathcal{F} if $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $0 \leq s < t < \infty$.*

Intuitively, \mathcal{F}_t corresponds to the information known to an observer at time t . In particular, for a process X , \mathcal{F}_t^X is defined to be $\mathcal{F}_t^X = \sigma(X(s))$ for $s < t$. This means that \mathcal{F}_t^X is the information obtained by observing X up to time t . A filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is said to satisfy the ‘usual condition’ if it is right continuous and \mathcal{F}_0 contains all P -negligible events of \mathcal{F} . A filtration $\{\mathcal{F}_t : t \geq 0\}$ is right continuous if $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for $0 \leq t < \infty$.

A stronger concept of measurability of a stochastic process is ‘adaptedness’.

Definition 1.1.4 *A stochastic process X_t is said to be “adapted” to the filtration $\{\mathcal{F}_t\}$ if for every $t \geq 0$, X_t is an \mathcal{F}_t -measurable random variable.*

Since \mathcal{F}_t is increasing in t , X is $\{\mathcal{F}_t\}$ -adapted if and only if $\mathcal{F}_t^X \subset \mathcal{F}_t$ for each $t \geq 0$.

Definition 1.1.5 *The stochastic process X_t is said to be “progressively measurable” with respect to the filtration $\{\mathcal{F}_t\}$ if for each $t \geq 0$, the mapping*

$$(s, \omega) \longrightarrow X(s, \omega)$$

for $s \leq t$ is measurable.

Note that if the stochastic process X is right continuous and adapted to the filtration $\{\mathcal{F}_t\}$ then it is also progressively measurable with respect to $\{\mathcal{F}_t\}$.

1.2 Stopping Times

If we interpret t as time and the σ -algebra, \mathcal{F}_t , as accumulated information up to time t , then we may be interested in the occurrence of a certain phenomenon. We are thus forced to pay particular attention to instant $T(\omega)$ at which the phenomenon manifest itself for the first time. As a result, the event

$$\{\omega : T(\omega) \leq t\}$$

which occurs if and only if the phenomenon has occurred prior to time t , should be part of the information accumulated at that time. See [17].

Definition 1.2.1 *Let us consider a measurable sample (Ω, \mathcal{F}) equipped with filtration $\{\mathcal{F}_t\}_{t \geq 0}$. A random variable $T : \Omega \rightarrow [0, \infty]$ is a “stopping time” if the event $\{T \leq t\}$ belongs to the σ -field \mathcal{F}_t , for every $t \geq 0$.*

1.3 The Poisson Process

The Poisson process is one of the fundamental examples in the theory of continuous time stochastic processes. We assume given a probability space (Ω, \mathcal{F}, P) which satisfy the usual conditions. Let $\{T_n\}_{n \geq 0}$ be a strictly increasing sequence of positive random variables. We always take $T_0 = 0$ a.s. In fact $T_0 = 0$ is a stopping time (see definition 1.2.1). Define the indicator function $1_{t \geq T_n}$ as

$$1_{t \geq T_n} = \begin{cases} 1 & \text{if } t \geq T_n(\omega) \\ 0 & \text{if } t < T_n(\omega) \end{cases}$$

Definition 1.3.1 *The process $N = \{N_t\}_{t \geq 0}$ defined by*

$$N_t = \sum_{n \geq 1} 1_{\{t \geq T_n\}}$$

with values in $\mathbb{N} \cup \{\infty\}$ ($\mathbb{N} = 0, 1, 2, \dots$) is called the “counting process” associated with the sequence $\{T_n\}_{n \geq 1}$.

If we set $T = \sup_n T_n$, then

$$[T_n, \infty) = \{N \geq n\} = \{(t, \omega) : N_t(\omega) \geq n\}$$

as well as $[T_n, T_{n+1}) = \{N = n\}$, and $[T, \infty) = \{N = \infty\}$.

The random variable T is the “explosion time” of N . If $T = \infty$ a.s., then N is a counting process “without explosions”. For $T = \infty$, note that for $0 \leq s < t < \infty$ we have

$$N_t - N_s = \sum_{n \geq 1} 1_{\{s < T_n \leq t\}}$$

The increments $N_t - N_s$ counts the number of random times T_n that occur between fixed times s and t . Note that from the definition, the counting process, is not necessarily adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$. Thus we have the following:

Theorem 1.1 *A counting process N is adapted if and only if the associated random variables $\{T_n\}_{n \geq 1}$ are stopping times.*

Proof

If the $\{T_n\}_{n \geq 0}$ are stopping times with $T_0 = 0$ a.s., then the event

$$\{N_t = n\} = \{\omega : T_n(\omega) \leq t < T_{n+1}(\omega)\} \in \mathcal{F}_t$$

for each n . Thus $N_t \in \mathcal{F}_t$ and N is adapted.

If N is adapted, then

$$\{T_n \leq t\} = \{N_t \geq n\} \in \mathcal{F}_t$$

for each t and therefore T_n is a stopping time.

Note that a counting process without explosions has right continuous trajectories with left limits; hence a counting process without explosions is càdlàg.

Definition 1.3.2 *An adapted counting process N without explosions is a “Poisson Process” if*

1. For any $s, t, 0 \leq s < t < \infty$, $N_t - N_s$ is independent of \mathcal{F}_s
2. For any $s, t, u, v, 0 \leq s < t < \infty, 0 \leq u < v < \infty, t - s = v - u$ then the distribution of $N_t - N_s$ is the same as that of $N_v - N_u$

The conditions (1) and (2) are known respectively as increments independent of the past and stationary increments.

Theorem 1.2 *Let N be a Poisson process. Then*

$$P(N_t = n) = \frac{\exp^{-\lambda t} (\lambda t)^n}{n!}$$

$n = 0, 1, 2, \dots$ for some $\lambda \geq 0$. That is N_t has the Poisson distribution with parameter λt . Moreover, N is continuous in probability.

Proof

The proof will be given as in [17].

step 1

For all $t \geq 0$,

$$P(N_t = 0) = \exp^{-\lambda t}$$

for some constant $\lambda \geq 0$. Since

$$\{N_t = 0\} = \{N_s = 0\} \cap \{N_t - N_s = 0\}$$

for $0 \leq s < t < \infty$, by independence of increments,

$$\begin{aligned} P(N_t = 0) &= P(N_s = 0)P(N_t - N_s = 0) \\ &= P(N_s = 0)P(N_{t-s} = 0) \end{aligned}$$

by the stationarity of the increments.

Let $\alpha(t) = P(N_t = 0)$, we have

$$\alpha(t) = \alpha(s)\alpha(t - s)$$

for all $0 \leq s < t < \infty$. Since $\alpha(t)$ can be easily seen to be right continuous in t , we deduce that either

$$\alpha(t) = 0$$

for all $t \geq 0$, or

$$\alpha(t) = \exp^{-\lambda t}$$

for some $\lambda \geq 0$. If $\alpha(t) = 0$, it would follow that $N_t(\omega) = \infty$ a.s. for all t which would contradict that N is a counting process. Observe that

$$\begin{aligned} \lim_{u \rightarrow t} P(|N_u - N_t| > \epsilon) &= \lim_{u \rightarrow t} P(|N_{u-t}| > \epsilon) \\ &= \lim_{v \rightarrow 0} P(N_v > \epsilon) \\ &= \lim_{v \rightarrow 0} 1 - \exp^{-\lambda v} \\ &= 0 \end{aligned}$$

hence N is continuous in probability.

step 2

$P(N_t \geq 2)$ is $o(t)$. That is

$$\lim_{t \rightarrow 0} \frac{1}{t} P(N_t \geq 2) = 0$$

Let $\beta(t) = P(N_t \geq 2)$. Since the path of N_t is nondecreasing, β is also nondecreasing. We can quickly check that showing

$$\lim_{t \rightarrow 0} \frac{1}{t} \beta(t) = 0$$

is equivalent to showing that

$$\lim_{n \rightarrow \infty} n \beta\left(\frac{1}{n}\right) = 0$$

Divide $[0, 1]$ into n subintervals of equal length and let S_n denote the number of subintervals containing at least two arrivals. By the independence and stationarity of the increments S_n is the sum of n independently and identically distributed zero-one valued random variables and hence has a binomial distribution (n, p) where $p = \beta\left(\frac{1}{n}\right)$. Therefore

$$\mathbb{E}[S_n] = np = n \beta\left(\frac{1}{n}\right)$$

Since N is a counting process, we know that the arrival times are strictly increasing; that is $T_n < T_{n+1}$ a.s.

Therefore for fixed ω , for n sufficiently large no subinterval has more than one arrival, otherwise there would be an explosion. This implies

$$\lim_{n \rightarrow \infty} S_n(\omega) = 0$$

a.s. Since $S_n \leq N_1$, if $\mathbb{E}[N_1] < \infty$, we can use the Dominated Convergence Theorem to conclude

$$\lim_{n \rightarrow \infty} n\beta\left(\frac{1}{n}\right) = \lim_{n \rightarrow \infty} \mathbb{E}[S_n] = 0$$

step 3

$\lim_{t \rightarrow 0} \frac{1}{t}P(N_t = 1) = \lambda$ since

$$P(N_t = 1) = 1 - P(N_t = 0) - P(N_t \geq 2)$$

and

$$\lim_{t \rightarrow 0} \frac{1}{t}P(N_t = 1) = \lim_{t \rightarrow 0} \frac{1 - \exp^{-\lambda t} + o(t)}{t} = \lambda$$

step 4

We write $\varphi(t) = \mathbb{E}[\alpha^{N_t}]$ for $0 \leq \alpha \leq 1$. Then for $0 \leq s < t < \infty$, the independence and stationarity of the increments implies that $\varphi(t+s) = \varphi(t)\varphi(s)$ which in turn implies that

$$\varphi(t) = \exp^{t\psi(\alpha)}$$

But

$$\varphi(t) = \sum_{n=0}^{\infty} \alpha^n P(N_t = n) = P(N_t = 0) + \alpha P(N_t = 1) + \sum_{n=2}^{\infty} \alpha^n P(N_t = n) \quad (1.1)$$

and $\psi(\alpha) = \varphi'(0)$ that is the derivative of φ at 0. Therefore

$$\begin{aligned} \psi(\alpha) &= \lim_{t \rightarrow 0} \frac{\varphi(t) - 1}{t} \\ &= \lim_{t \rightarrow 0} \left\{ \frac{P(N_t = 0) - 1}{t} + \frac{\alpha P(N_t = 1)}{t} + \frac{1}{t}o(t) \right\} \\ &= -\lambda + \alpha\lambda \end{aligned}$$

Therefore

$$\varphi(t) = \exp^{-\lambda t + \lambda \alpha t}$$

hence

$$\varphi(t) = \sum_{n=0}^{\infty} \alpha^n P(N_t = n) = \exp^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n \alpha^n}{n!} \quad (1.2)$$

Equating coefficients of the two infinite series, (1.1) and (1.2) we have

$$P(N_t = n) = \frac{\exp^{-\lambda t} (\lambda t)^n}{n!}$$

for $n = 0, 1, 2, \dots$

Definition 1.3.3 *The parameter λ associated to a Poisson process in Theorem 1.2 is called the intensity, or the arrival rate, of the process.*

Corollary 1 *A Poisson process N with intensity λ satisfies*

$$\mathbb{E}[N_t] = \lambda t$$

$$\text{Variance}[N_t] = \text{Var}[N_t] = \lambda t$$

There are other, equivalent definitions of the Poisson process. For example, a counting process N without explosion can be seen to be a Poisson process if for all s, t $0 \leq s < t < \infty$, $\mathbb{E}[N_t] < \infty$ and

$$\mathbb{E}[N_t - N_s \mid \mathcal{F}_s] = \lambda(t - s)$$

Note that the Poisson process $\tilde{N} = \{\tilde{N}_t\}_{t \geq 0}$ where

$$\tilde{N}(t) = N(t) - \lambda t$$

is called the compensated Poisson process.

We also note that $\mathbb{E}[\tilde{N}(t)] = 0$ and $\mathbb{E}[(\tilde{N}(t))^2] = \lambda t$ for each $t \geq 0$.

1.4 Martingales

In this section we give, mostly without proofs, only essential results from the theory of continuous time martingales. See [2], [14], [16], [17] and the references therein.

Assume as given a probability space (Ω, \mathcal{F}, P) where the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is assumed to be right continuous. Suppose that we are given also a stochastic process $X_t, t \geq 0$. Let \mathcal{F}_t be the smallest σ -field including $\bigcap_{\varepsilon > 0} \sigma(X_s : s \leq t + \varepsilon)$ and all null sets of \mathcal{F}_0 . Then \mathcal{F}_t is a filtration and X_t is an \mathcal{F}_t -adapted process. It is called the filtration generated by the process X_t . Let X_t be a real-valued \mathcal{F}_t -adapted process such that for each t , X_t is integrable. It is called a ‘martingale’ if it satisfies :

$$\mathbb{E}[X_t | \mathcal{F}_s] = X_s$$

a.s for any $t > s$. It is called a ‘submartingale’ if it satisfies

$$\mathbb{E}[X_t | \mathcal{F}_s] \geq X_s$$

a.s for any $t > s$. Further, if the converse inequality

$$\mathbb{E}[X_t | \mathcal{F}_s] \leq X_s$$

holds a.s it is called a ‘supermartingale’.

One of the most important examples of martingales is a Brownian motion.

Definition 1.4.1 *A ‘Brownian motion’ is a real-valued continuous stochastic process $\{\eta_t\}_{t \geq 0}$ with independent and stationary increments.*

That is

1. Continuity: the map $s \rightarrow \eta_s(\omega)$ is continuous P a.s
2. Independent increments: If $s \leq t$ then $\eta_t - \eta_s$ is independent of $\mathcal{F}_s = \sigma(\eta_u, u \leq s)$.

3. Stationary increment: If $s \leq t$ then $\eta_t - \eta_s$ and $\eta_{t-s} - \eta_0$ have the same probability law.

A Brownian motion is called ‘standard’ if $\eta_0 = 0$ P a.s, $\mathbb{E}[\eta_t] = 0$ and $\mathbb{E}[\eta_t^2] = t$.

Example 1.1 A Brownian motion η_t is a martingale with respect to the σ -algebra \mathcal{F}_t , generated by $\{\eta_s; s \leq t\}$.

Proof

$$\mathbb{E}[|\eta_t|^2] \leq \mathbb{E}[|\eta_t|^2] = \eta_0^2 + nt$$

If $s \geq t$ then

$$\begin{aligned} \mathbb{E}[\eta_s | \mathcal{F}_t] &= \mathbb{E}[\eta_s - \eta_t + \eta_t | \mathcal{F}_t] \\ &= \mathbb{E}[\eta_s - \eta_t | \mathcal{F}_t] + \mathbb{E}[\eta_t | \mathcal{F}_t] \\ &= 0 + \eta_t \\ &= \eta_t \end{aligned}$$

Note that since $\eta_s - \eta_t$ is independent of \mathcal{F}_t , $\mathbb{E}[\eta_s - \eta_t | \mathcal{F}_t] = 0$ and $\mathbb{E}[\eta_t | \mathcal{F}_t] = \eta_t$ since η_t is \mathcal{F}_t -measurable.

Theorem 1.3 Let $\eta = (\eta_t)_{t \geq 0}$ be a 1-dimensional standard Brownian motion with $\eta_0 = 0$. Then $M_t = \eta_t^2 - t$ is a martingale.

Proof

$\mathbb{E}[M_t] = \mathbb{E}[\eta_t^2 - t] = 0$. Also

$$\mathbb{E}[M_t - M_s | \mathcal{F}_s] = \mathbb{E}[\eta_t^2 - \eta_s^2 - (t - s) | \mathcal{F}_s]$$

and

$$\mathbb{E}[\eta_t \eta_s | \mathcal{F}_s] = \eta_s \mathbb{E}[\eta_t | \mathcal{F}_s] = \eta_s^2$$

since η is a martingale with $\eta_s, \eta_t \in L_2(\Omega, \mathcal{F}, P)$. Therefore

$$\begin{aligned}
\mathbb{E}[M_t - M_s \mid \mathcal{F}_s] &= \mathbb{E}[\eta_t^2 - 2\eta_t\eta_s + \eta_s^2 - (t - s) \mid \mathcal{F}_s] \\
&= \mathbb{E}[(\eta_t - \eta_s)^2 - (t - s) \mid \mathcal{F}_s] \\
&= \mathbb{E}[(\eta_t - \eta_s)^2] - (t - s) \\
&= 0
\end{aligned}$$

due to the independence of the increments from the past.

Proposition 1 *Given a poisson process N_t with intensity λ . The compensated Poisson process $\tilde{N}_t = N_t - \lambda t$ is a martingale.*

Proof

Let $\tilde{N}_t = N_t - \lambda t$ where N_t is a Poisson process with intensity λ . Now

$$\begin{aligned}
\mathbb{E}[\tilde{N}_t \mid \mathcal{F}_s] &= \mathbb{E}[(N_t - \lambda t) \mid \mathcal{F}_s] \\
&= \mathbb{E}[N_t \mid \mathcal{F}_s] - \lambda t \\
&= \mathbb{E}[(N_t - N_s + N_s) \mid \mathcal{F}_s] - \lambda t \\
&= \mathbb{E}[(N_t - N_s) \mid \mathcal{F}_s] + \mathbb{E}[N_s \mid \mathcal{F}_s] - \lambda t \\
&= \mathbb{E}[N_{t-s}] + N_s - \lambda t \\
&= \lambda(t - s) + N_s - \lambda t \\
&= N_s - \lambda s \\
&= \tilde{N}_s
\end{aligned}$$

The analogous statement holds for $(N_t - \lambda t)^2 - \lambda t$.

We now quote Doob's optional sampling Theorem and its consequences called Doob's inequalities.

Theorem 1.4 The Optional Sampling Theorem, see [17]: *If $\{M_t\}_{t \geq 0}$ is a continuous martingale with respect to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ and if τ_1 and τ_2 are*

stopping times such that $\tau_1 \leq \tau_2 \leq k$ where k is a finite real number, then M_{τ_2} is integrable and

$$\mathbb{E}[M_{\tau_2} | \mathcal{F}_{\tau_1}] = M_{\tau_1}$$

P a.s.

Therefore using the Optional Sampling Theorem, let $\tau_1 = 0$ and $\tau_2 = \tau$ then

$$\mathbb{E}[M_\tau | \mathcal{F}_0] = M_0$$

Theorem 1.5 Doob's martingale inequality see [17], [2]: *If M is a nonnegative submartingale with right continuous sample paths then*

$$\mathbb{E}[\sup_{0 \leq t \leq T} M_t]^p \leq \left(\frac{p}{p-1}\right)^p E[M_T]^p$$

for all $p > 1$.

In particular if $p = 2$ and M , a martingale then

$$\mathbb{E}[\sup_{0 \leq t \leq T} M_t]^2 \leq 4\mathbb{E}[|M_T|^2]$$

Theorem 1.6 The Doob-Meyer Decomposition, see [17]: *Let X be a positive supermartingale, and suppose $\mathcal{H} = \{X_T; T \text{ a stopping time}\}$ is uniformly integrable. Then X has a unique decomposition $X = M - A$ where M is a martingale and A is a right continuous, increasing, natural process with $A_0 = 0$.*

Definition 1.4.2 *A process A is called "natural" if for every bounded right continuous martingale $\{M_t\}_{t \geq 0}$ we have*

$$\mathbb{E}\left[\int_0^t M_s dA_s\right] = \mathbb{E}\left[\int_0^t M_{s-} dA_s\right]$$

for every $0 \leq t < \infty$.

Definition 1.4.3 For $X \in L_2(\Omega, \mathcal{F}, P)$ we define the “quadratic variation of X ” to be the process $\langle X \rangle_t = A_t$ where A is the natural increasing process in the Doob-Meyer decomposition of X .

In other words, $\langle X \rangle$ is that unique, adapted natural increasing process for which $\langle X \rangle_0 = 0$ a.s and $X^2 - \langle X \rangle$ is a martingale.

For any two martingales X, Y , we define their “cross variation” process $\langle X, Y \rangle$ by

$$\langle X, Y \rangle := \frac{1}{4}[\langle X + Y \rangle_t - \langle X - Y \rangle_t]$$

We then observe that $XY - \langle X, Y \rangle$ is a martingale. We also note that:

1. The uniqueness of the Doob-Meyer decomposition shows that $\langle X, Y \rangle$ is the only process of the form $A = A^{(1)} - A^{(2)}$ with A^j adapted and increasing natural process such that

$$XY - A$$

is a martingale.

2. In view of the identities

$$\mathbb{E}[(X_t - X_s)(Y_t - Y_s) \mid \mathcal{F}_s] = \mathbb{E}[(X_t Y_t - X_s Y_s) \mid \mathcal{F}_s] = \mathbb{E}[(\langle X, Y \rangle_t - \langle X, Y \rangle_s) \mid \mathcal{F}_s]$$

valid P a.s for every $0 \leq t < s < \infty$ the orthogonality of $X, Y \in L_2(\Omega, \mathcal{F}, P)$ is equivalent to the statements “ XY is a martingale” or the increments of X and Y over $[s, t]$ are conditionally uncorrelated given \mathcal{F}_s and we call $\langle M, M \rangle$, Meyer’s angle bracket.

1.4.4 Orthogonal Decomposition

Let M and N be martingales. If $\langle M, N \rangle = 0$ then M and N are said to be orthogonal. Since $M_t N_t - \langle M, N \rangle_t$ is a martingale then M and N are

orthogonal if and only if $M_t N_t$ is a martingale- see example 1.2.

Note that for M and N martingales, we have

$$\mathbb{E}[(M_t - M_s)(N_t - N_s) \mid \mathcal{F}_s] = 0$$

for all $t > s$. Hence the orthogonality means that $M_t - M_s$ and $N_t - N_s$ are orthogonal in L_2 -space.

Now, if $M_t - M_s$, $N_t - N_s$ and \mathcal{F}_s are independent and $\mathbb{E}[M_t - M_s] = 0$ or $\mathbb{E}[N_t - N_s] = 0$ holds, then

$$\mathbb{E}[(M_t - M_s)(N_t - N_s) \mid \mathcal{F}_s] = \mathbb{E}[(M_t - M_s)(N_t - N_s)] = 0$$

so that M_t and N_t are orthogonal.

Example 1.2 Check whether the process $X_t = \eta_1(t)\eta_2(t)$, where $(\eta_1(t), \eta_2(t))$ is 2-dimensional Brownian motion, is a martingale with respect to $\{\mathcal{F}_t\}$

Solution

$$\begin{aligned} \mathbb{E}[(\eta_1(t)\eta_2(t)) \mid \mathcal{F}_s] &= \mathbb{E}[(\eta_1(t) - \eta_1(s) + \eta_1(s))\eta_2(t) \mid \mathcal{F}_s] \\ &= \mathbb{E}[(\eta_1(t) - \eta_1(s))\eta_2(t) \mid \mathcal{F}_s] + \mathbb{E}[\eta_1(s)\eta_2(t) \mid \mathcal{F}_s] \\ &= \mathbb{E}[(\eta_1(t) - \eta_1(s))(\eta_2(t) - \eta_2(s)) \mid \mathcal{F}_s] + \eta_1(s)\eta_2(s) \\ &= \mathbb{E}[(\eta_1(t) - \eta_1(s))(\eta_2(t) - \eta_2(s)) \mid \mathcal{F}_s] \\ &\quad + \mathbb{E}[\eta_2(s)(\eta_1(t) - \eta_1(s)) \mid \mathcal{F}_s] + \eta_1(s)\eta_2(s) \\ &= \mathbb{E}[(\eta_1(t) - \eta_1(s))(\eta_2(t) - \eta_2(s))] + \eta_2(s)\mathbb{E}[\eta_1(t) - \eta_1(s)] + \eta_1(s)\eta_2(s) \end{aligned}$$

If η_1 and η_2 are independent then the above reduces to $\eta_1(s)\eta_2(s)$

Therefore

$$\mathbb{E}[(\eta_1(t)\eta_2(t)) \mid \mathcal{F}_s] = \eta_1(s)\eta_2(s)$$

Thus $\eta_1(t)\eta_2(t)$ is a martingale.

1.5 Quadratic variation

For the following results see [16], [14], [17].

Let $X = \{X_t\}_{t \geq 0}$ be a stochastic process. Fix $t \geq 0$ and let $\pi = t_0, t_1, \dots, t_m$ with $0 \leq t_0 \leq t_1 \leq t_2 \leq \dots \leq t_m = t$ be the partition of $[0, t]$. We define the “ p^{th} variation” of X , $p > 0$ over the partition π to be the limit of $V_t^{(p)}$ that is:

$$\lim_{\pi \rightarrow 0} V_t^{(p)}(\pi) = \sum_{k=1}^m |X_{t_k} - X_{t_{k-1}}|^p$$

Define the mesh of the partition as

$$\|\pi\| = \max_{1 \leq k \leq m} |t_k - t_{k-1}|$$

If $V_t^2(\pi)$ converges in some sense as $\|\pi\| \rightarrow 0$ then the limit is called a quadratic variation of X on $[0, t]$. We summarise this in the following theorem:

Theorem 1.7 *Let X be a continuous-square integrable martingale. For a partition π of $[0, t]$ we have*

$$\lim_{\|\pi\| \rightarrow 0} V_t^2(\pi) = \langle X \rangle_t$$

in probability, that is for all $\epsilon > 0$, $\eta > 0$ there exists $\delta > 0$ such that $\|\pi\| < \delta$ implies

$$P[|V_t^2(\pi) - \langle X \rangle_t| > \epsilon] < \eta$$

Proposition 2 *If $p = 1$, in the definition of the p^{th} variation of the process X , then we call*

$$\lim_{\pi \rightarrow 0} V_\pi(X) = \sum_{k=1}^m |X_{t_k} - X_{t_{k-1}}|$$

the variation of X and if

$$\sup_{\pi} V_\pi(X) < \infty$$

we say that X has finite variation on $[0, t]$.

Theorem 1.8 *With probability one, we have for every non constant martingale, M , with continuous sample paths*

1.

$$\sum_{k=1}^{2n} |M_{t_k} - M_{t_{k-1}}| \longrightarrow \infty$$

as $n \rightarrow \infty$

2.

$$\sum_{k=1}^{2n} (M_{t_k} - M_{t_{k-1}})^2 \longrightarrow \langle M \rangle_t$$

as $n \rightarrow \infty$

1.6 Predictable Processes

Definition 1.6.1 *The “predictable σ -field” is the least σ -field on the product space $[0, T] \times \Omega$ for which all continuous \mathcal{F}_t -adapted processes are measurable.*

Definition 1.6.2 *A “predictable process” is a process measurable with respect to the predictable σ -field.*

Note that a continuous \mathcal{F}_t -adapted process is predictable. See [14]. Let A_t be a continuous increasing process and let p be a positive number such that $p \geq 1$. We denote by $L_p(A)$ the set of all predictable process f_t such that

$$\int_0^T |f_s|^p dA_s < \infty$$

a.s.

Then the set of continuous \mathcal{F}_t -adapted process is dense in $L_p(A)$, that is for any $\mathcal{F}_t \in L_p(A)$, there exists a sequence of continuous \mathcal{F}_t -adapted process $\{f_s^n\}$ such that

$$\int_0^T |f_s^n - f_s|^p d \langle M \rangle_s \longrightarrow 0$$

a.s. where $\langle M \rangle_s$ is the quadratic variation.

For any $f \in L_2(\langle M \rangle)$, choose a sequence of continuous predictable processes $\{f^n\}$ converging to f in the above sense. Then the stochastic integrals $\int_0^t f_s^n dM_s$ forms a cauchy sequence.

Notation

1. $L_2(\Omega)$ - the set of \mathcal{F}_T -adapted square-integrable random variables with norm:

$$\| \xi \|_2 := (\mathbb{E} | \xi |^2)^{\frac{1}{2}}$$

Chapter 2

Techniques in stochastic analysis

2.1 The Itô Calculus

We will begin by a definition of a simple class of functions, f , and then extend to some form of approximation procedure. The approximation procedure will work out successfully provided that f has the property that each of the functions $\omega \rightarrow f(t_j, \omega)$ only depends on the behaviour of $X_s(\omega)$ up to time t_j . We will follow closely the work in [2]. Now we will describe the class of functions for which the Itô integral will be defined. See [2]

Definition 2.1.1 *Let $\mathcal{V} = \mathcal{V}(\mathcal{S}, \mathcal{T})$ be a class of functions*

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

such that

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

$$3. E[\int_S^T f^2(t, \omega) dt] < \infty$$

We will now make use of the functions $f \in \mathcal{V}$, to define the Itô integral:

$$I[f](\omega) = \int_S^T f(t, \omega) d\eta_t(\omega)$$

where η is a 1-dimensional Brownian motion. Firstly, we define $I[\varphi]$ for a simple class of functions φ . There after we show that for each $f \in \mathcal{V}$ we can approximate such φ 's. We will then use this idea to define $\int f d\eta$ as the limit of $\int \varphi d\eta$ as $\varphi \rightarrow f$. This is the most important concept we will make use of in trying to understand our topic. To consider some details of its construction, we begin by defining an elementary function.

Definition 2.1.2 *A function $\varphi \in \mathcal{V}$ of the form*

$$\varphi(t, \omega) = \sum_j e_j(\omega) 1_{[t_j, t_{j+1})}(t)$$

is called “an elementary function” where $1_{[t_j, t_{j+1})}$ is the indicator function.

In this case since $\varphi \in \mathcal{V}$, each function e_j must be \mathcal{F}_t -measurable. For elementary function $\varphi(t, \omega)$, we define the integral as

$$\int_S^T \varphi(t, \omega) d\eta_t(\omega) = \sum_{j \geq 0} e_j(\omega) [\eta_{t_{j+1}} - \eta_{t_j}](\omega)$$

If $\varphi(t, \omega)$ is not elementary then we have

$$\int_S^T \varphi(t, \omega) d\eta_t(\omega) = \lim_{n \rightarrow \infty} \int_S^T f_n(t, \omega) d\eta_t(\omega)$$

where f_n is a sequence of elementary functions such that

$$\mathbb{E}[\int_S^T (\varphi(t, \omega) - f_n(t, \omega))^2 dt] \rightarrow 0$$

Definition 2.1.3 The Itô Isometry: *If $\varphi(t, \omega)$ is bounded and elementary then*

$$\mathbb{E}[(\int_S^T \varphi(t, \omega) d\eta_t(\omega))^2] = \mathbb{E}[\int_S^T \varphi(t, \omega)^2 dt]$$

proof

Put $\Delta\eta_j = \eta_{t_{j+1}} - \eta_{t_j}$. Note that for $i < j$, $e_i e_j \Delta\eta_i$ and $\Delta\eta_j$ are independent

Then

$$\mathbb{E}[e_i e_j \Delta\eta_i \Delta\eta_j] = \begin{cases} 0 & \text{if } i \neq j \\ \mathbb{E}[e_j^2] \cdot (t_{j+1} - t_j) & \text{if } i = j \end{cases}$$

Therefore

$$\mathbb{E}[(\int_S^T \varphi d\eta)^2] = \sum_{i,j} \mathbb{E}[e_i e_j \Delta\eta_i \Delta\eta_j] = \sum_{i,j} \mathbb{E}[e_j^2] (t_{j+1} - t_j) = \mathbb{E}[\int_S^T \varphi^2 dt]$$

We now use the Itô Isometry to extend the definition from elementary functions to functions in \mathcal{V}

Step 1

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

$$\mathbb{E}[\int_S^T (g - \varphi_n)^2 dt] \longrightarrow 0$$

as $n \rightarrow \infty$

proof

Define $\varphi_n(t, \omega) = \sum_j g(t_j, \omega) 1_{[t_j, t_{j+1})}(t)$.

Then φ_n is elementary since $g \in \mathcal{V}$ and

$$\int_S^T (g - \varphi_n)^2 dt \longrightarrow 0$$

as $n \rightarrow \infty$ for each ω

since $g(\cdot, \omega)$ is continuous for each ω , hence

$$\mathbb{E}[\int_S^T (g - \varphi_n)^2 dt] \longrightarrow 0$$

as $n \rightarrow \infty$ by bounded convergence.

Step 2

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

$$\mathbb{E}[\int_S^T (h - g_n)^2 dt] \longrightarrow 0$$

proof

Since h is bounded, it means there exists an $M \in \mathbb{R}$ such that $h(t, \omega) \leq M$ for all (t, ω) .

For each n , let ψ_n be a nonnegative continuous function on \mathbb{R} such that

1. $\psi_n(x) = 0$ for each $x \leq -\frac{i}{n}$ and $x \geq 0$
2. $\int_{-\infty}^{\infty} \psi_n(x) dx = 1$

Define

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

Then $g_n(t, \omega)$ is continuous for each ω and $|g_n(t, \omega)| \leq M$. Since $h \in \mathcal{V}$, we observe that $g_n(t, \omega)$ is \mathcal{F}_t -measurable for all t .

Furthermore,

$$\int_S^T (g_n(s, \omega) - h(s, \omega))^2 ds \longrightarrow 0$$

as $n \longrightarrow \infty$ for each ω , since $\{\psi_n\}_n$ constitute an approximate identity. Thus by Bounded Convergence Theorem,

$$\mathbb{E}\left[\int_S^T (h(t, \omega) - g_n(t, \omega))^2 dt\right] \longrightarrow 0$$

as $n \longrightarrow \infty$ as claimed.

Step 3

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

$$\mathbb{E}\left[\int_S^T (f - h_n)^2 dt\right] \longrightarrow 0$$

as $n \longrightarrow \infty$

proof

Put

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

Then the conclusion follows by the Dominated Convergence Theorem.

To complete the definition of the Itô integral

$$\int_S^T f(t, \omega) d\eta_t(\omega)$$

for $f \in \mathcal{V}$, we choose by steps 1 to 3 elementary functions $\varphi_n \in \mathcal{V}$ such that

$$\mathbb{E}\left[\int_S^T |f - \varphi_n|^2 dt\right] \longrightarrow 0$$

Then

$$I[f](\omega) := \int_S^T f(t, \omega) d\eta_t(\omega) = \lim_{n \rightarrow \infty} \int_S^T \varphi_n(t, \omega) d\eta_t(\omega)$$

The limit exists due to Cauchy criterion of convergence (see later, section 2.4) and Itô isometry.

We can now define the Itô Integral as follows:

Definition 2.1.4 *Let $f \in \mathcal{V}(S, T)$. Then the "Itô integral of f " is defined by*

$$\int_S^T f(t, \omega) d\eta_t(\omega) = \lim_{n \rightarrow \infty} \int_S^T \varphi_n(t, \omega) d\eta_t(\omega)$$

where φ_n is a sequence of elementary functions such that

$$\mathbb{E}\left[\int_S^T (f(t, \omega) - \varphi_n(t, \omega))^2 dt\right] \longrightarrow 0, \quad n \rightarrow \infty$$

Such sequences φ_n exists by steps 1 to 3 earlier on.

2.1.5 Properties of the Itô Integral

Theorem 2.1 *Let $f, g \in \mathcal{V}(0, T)$ and let $0 \leq S < U < T$ Then*

1. $\int_S^T f d\eta_t = \int_S^U f d\eta_t + \int_U^T f d\eta_t$ for a.a. ω
2. $\int_S^T (cf + g) d\eta_t = c \int_S^T f d\eta_t + \int_S^T g d\eta_t$ where c is a constant for a.a. ω

$$3. \mathbb{E}[\int_S^T f d\eta_t] = 0$$

4. $\int_S^T f d\eta_t$ is \mathcal{F}_T -measurable

proof of 1

Let $P = (t_0, \dots, t_n)$ be partitions of the interval $[S, T]$. let $Q = (t_0, \dots, t_{k-1}, U)$ and $R = (U, t_k, \dots, t_n)$ be partition of $[S, U]$ and $[U, T]$ respectively. Take $T = (x_1, \dots, x_n)$ to be any choice of points $x_i \in [t_{i-1}, t_i]$. Then $U = (x_1, \dots, x_{k-1}, U)$ and $V = (U, x_k, \dots, x_n)$ are the choices of points within the subintervals Q and R respectively. Now

$$S(F, Q, U) = \sum_{i=1}^{k-1} (t_i - t_{i-1})F(x_i) + (U - t_{i-1})F(U)$$

$$S(F, R, V) = \sum_{i=k+1}^n (t_i - t_{i-1})F(x_i) + (t_k - U)F(U)$$

$$S(F, P, T) = \sum_{i \neq k}^n (t_i - t_{i-1})F(x_i) + (t_k - t_{k-1})F(x_k)$$

Therefore

$$S(F, P, T) = S(F, Q, U) + S(F, R, V) + (t_k - t_{k-1})[F(x_k) - F(n)]$$

The idea is when $|P|$ is small, the first term on the right will be near $\int_S^U F$, the second will be near $\int_U^T F$ and the third will be small. This is due to the Cauchy criterion of convergence. To be precise, let $\varepsilon > 0$. Since f is integrable on $[S, U]$ there exists $\delta > 0$ such that

$$| \int_S^U f(t)d\eta_t - S(F, Q, U) | < \frac{\varepsilon}{3}$$

for all partitions Q of $[S, U]$ such that $\|Q\| \leq \delta$. Similarly there exists $\gamma > 0$ such that

$$| \int_U^T f(t)d\eta_t - S(F, R, V) | < \frac{\varepsilon}{3}$$

for all partitions R of $[U, T]$ such that $\|R\| \leq \gamma$ Now

$$| (t_k - t_{k-1})[F(x_k) - F(U)] | \leq 2 \|Q\| M$$

where $M = \sup_{S \leq t \leq T} |f(t)|$. Let P be the partition of $[S, T]$ such that

$$\|P\| \leq \min\left(\delta, \gamma, \frac{\varepsilon}{6 \|Q\| M}\right)$$

Then the result follows.

Proof of 2

Consider two simple functions φ_n and ψ_n . For a constant c ($c \neq 0$), if $c\varphi_n$ and ψ_n are monotone increasing sequences of simple functions converging to cf and g respectively then $c\varphi_n + \psi_n$ is a monotone increasing sequence converging to $cf + g$.

Then by Monotone Convergence Theorem,

$$\begin{aligned} \int_0^T (cf + g)d\eta_t &= \lim_{n \rightarrow \infty} \int_S^T (c\varphi_n + \psi_n)d\eta_t \\ &= \lim_{n \rightarrow \infty} \int_S^T c\varphi_n d\eta_t + \lim_{n \rightarrow \infty} \int_S^T \psi_n d\eta_t \\ &= c \lim_{n \rightarrow \infty} \int_S^T \varphi_n d\eta_t + \lim_{n \rightarrow \infty} \int_S^T \psi d\eta_t \\ &= c \int_S^T f d\eta_t + \int_S^T g d\eta_t \end{aligned}$$

An important property of the Itô integral is that it is a martingale. See chapter 1, section 1.4. For continuous martingales we use the important inequality due to Doob, see chapter 1, Theorem 1.5.

We now make use of the Doob's martingale inequality to prove that the Itô Integral

$$\int_0^t f(s, \omega) d\eta_s$$

can be chosen to depend continuously on t :

Theorem 2.2 *Let $f \in \mathcal{V}(0, T)$. Then there exists a t -continuous version of*

$$\int_0^t f(s, \omega) d\eta_s(\omega)$$

for $0 \leq t \leq T$. That is there exists a t -continuous stochastic process J_t on (Ω, \mathcal{F}, P) such that

$$P[J_t = \int_0^t f d\eta] = 1 \quad \text{for all } t \quad 0 \leq t \leq T$$

Proof

Let $\varphi_n = \varphi_n(t, \omega) = \sum_j e_j^{(n)}(\omega) 1_{[t_j^n, t_{j+1}^{(n)})}(t)$ be elementary functions such that

$$\mathbb{E}\left[\int_0^T (f - \varphi_n)^2 dt\right] \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

Put

$$I_n(t, \omega) = \int_0^t \varphi_n(s, \omega) d\eta_s(\omega)$$

and

$$I_t = I(t, \omega) = \int_0^t f(s, \omega) d\eta_s(\omega) \quad 0 \leq t \leq T$$

Then $I_n(\cdot, \omega)$ is continuous for all n . Moreover, $I_n(t, \omega)$ is a martingale with respect to \mathcal{F}_t , for all n when $t < s$ since

$$\begin{aligned} \mathbb{E}[I_n(s, \omega) \mid \mathcal{F}_t] &= \mathbb{E}\left[\left(\int_0^t \varphi_n d\eta + \int_t^s \varphi_n d\eta\right) \mid \mathcal{F}_t\right] \\ &= \int_0^t \varphi_n d\eta + \mathbb{E}\left[\sum_{t \leq t_j^{(n)} \leq t_{j+1}^{(n)} \leq s} e_j^{(n)} d\eta_j \mid \mathcal{F}_t\right] \\ &= \int_0^t \varphi_n d\eta + \sum_j \mathbb{E}[\mathbb{E}[e_j^{(n)} \Delta\eta_j \mid \mathcal{F}_{t_j}] \mid \mathcal{F}_t] \\ &= \int_0^t \varphi_n d\eta + \sum_j \mathbb{E}[e_j^{(n)} E[\Delta\eta_j \mid \mathcal{F}_{t_j}] \mid \mathcal{F}_t] \\ &= \int_0^t \varphi_n d\eta \\ &= I_n(t, \omega) \end{aligned}$$

Hence $I_n - I_m$ is also a martingale, so by the martingale inequality, it follows that

$$\begin{aligned} P\left[\sup_{0 \leq t \leq T} |I_n(t, \omega) - I_m(t, \omega)| > \epsilon\right] &\leq \frac{1}{\epsilon^2} \cdot \mathbb{E}[|I_n(T, \omega) - I_m(T, \omega)|^2] \\ &= \frac{1}{\epsilon^2} \mathbb{E}\left[\int_0^T (\varphi_n - \varphi_m)^2 ds\right] \\ &\rightarrow 0 \end{aligned}$$

as $n, m \rightarrow \infty$.

Thus we may choose a subsequence $n_k \rightarrow \infty$ such that

$$P\left[\sup_{0 \leq t \leq T} |I_{n_{k+1}}(t, \omega) - I_{n_k}(t, \omega)| > 2^{-k}\right] < 2^{-k}$$

By the Borel-Cantelli lemma

$$P\left[\sup_{0 \leq t \leq T} |I_{n_{k+1}}(t, \omega) - I_{n_k}(t, \omega)| > 2^{-k} \text{ for infinitely many } k\right] = 0$$

So for a.a. ω there exists $k_1(\omega)$ such that

$$\sup_{0 \leq t \leq T} |I_{n_{k+1}}(t, \omega) - I_{n_k}(t, \omega)| \leq 2^{-k} \text{ for } k \geq k_1(\omega)$$

Therefore $I_{n_k}(t, \omega)$ is uniformly convergent for $t \in [0, T]$, for a.a. ω and so the limit, denoted by $J_t(\omega)$, is t -continuous for $t \in [0, T]$, a.s. Since $I_{n_k}(t, \cdot) \rightarrow I(t, \cdot)$ in $L_2(P)$ for all t , we must have

$$I_t = J_t$$

a.s., for all $t \in [0, T]$.

That completes the proof.

From now on we shall always assume that $\int_0^t f(s, \omega) d\eta_s(\omega)$ means a t -continuous version of the integral.

Corollary 2 *Let $f(t, \omega) \in \mathcal{V}(0, T)$ for all T . Then*

$$M_t(\omega) = \int_0^t f(s, \omega) d\eta_s$$

is a martingale with respect to \mathcal{F}_t and

$$P\left[\sup_{0 \leq t \leq T} |M_t| \geq \lambda\right] \leq \frac{1}{\lambda^2} \cdot \mathbb{E}\left[\int_0^T f(s, \omega)^2 ds\right] \text{ for } \lambda, T > 0$$

2.2 The Itô formula

The Itô integral is not very useful when we try to evaluate a given integral. The important tool for evaluating Itô integral is the Itô formula.

We now state the main results without proofs (see [2] for proofs):

Theorem 2.3 The 1-dimensional Itô formula

Let X_t be an Itô process given by

$$dX_t = udt + v d\eta_t$$

Let $g(t, x) \in C^2([0, \infty) \times \mathbb{R})$. That is $g(t, x)$ is twice continuously differentiable on $[0, \infty) \times \mathbb{R}$. Then

$$Y_t = g(t, X_t)$$

is again an Itô process and

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

where $(dX_t)^2 = (dX_t) \cdot (dX_t)$ is computed according to the rules $dt \cdot dt = dt \cdot d\eta_t = d\eta_t \cdot dt = 0$ and $d\eta_t \cdot \eta_t = dt$

For higher dimensions, we use the general itô formula.

Theorem 2.4 The general Itô formula

Let

$$dX(t) = udt + v d\eta_t$$

be an n -dimensional Itô process. Let $g(t, x) = g_1(t, x), \dots, g_n(t, x)$ be a C^2 map from $[0, \infty) * \mathbb{R}^n$ into \mathbb{R}^p . Then the process is again an Itô process whose component number k , Y_k is given by

$$dY_k = \frac{\partial g_k}{\partial t}(t, X)dt + \sum_i \frac{\partial g_k}{\partial x_i}(t, X)dX_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t, X)dX_i dX_j$$

where $d\eta_i d\eta_j = \delta_{ij} dt$ and $d\eta_i dt = dt d\eta_i = 0$

Theorem 2.5 The Itô representation

For any \mathcal{F}_T -measurable random variable $\xi \in L_2(\Omega)$ there exists a unique \mathbb{F} -adapted (non-anticipating) stochastic function $\varphi = \varphi(t, \omega)$, $0 \leq t \leq T$ in $L_2(\Omega \times [0, T])$ such that

$$\xi(\omega) = \mathbb{E}[\xi] + \int_0^T \varphi(t, \omega) d\eta_t \tag{2.1}$$

Proof

Assume that $\xi = Y(t)$.

Define

$$Y_t(\omega) = \exp\left(\int_0^t h(s)d\eta_s(\omega) - \frac{1}{2} \int_0^t h^2(s)ds\right)$$

for $0 \leq t \leq T$.

Then by Itô's formula

$$dY_t = Y_t(h(t)d\eta_t - \frac{1}{2}h^2(t)dt) + \frac{1}{2}Y_t(h(t)d\eta_t)^2 = Y_t h(t)d\eta_t$$

so that

$$Y_t = 1 + \int_0^t Y_s h(s)d\eta_s$$

for $t \in [0, T]$

Therefore

$$\xi = Y_T = 1 + \int_0^T Y_s h(s)d\eta_s$$

and hence $\mathbb{E}[\xi] = 1$, so (2.1) holds in this case. By linearity (2.1) holds also for linear combinations of functions of the form

$$\exp\left(\int_0^T h(t)d\eta_s(\omega) - \frac{1}{2} \int_0^T h^2(t)dt\right) \quad (2.2)$$

$h \in L_2[0, T]$. So if $\xi \in L_2(\Omega, \mathcal{F}, P)$ is arbitrary, we approximate ξ in $L_2(\Omega, \mathcal{F}, P)$ by linear combinations ξ_n of functions of the form (2.2). Then for each n , we have

$$\xi_n(\omega) = \mathbb{E}[\xi_n] + \int_0^T f_n(s, \omega)d\eta_s(\omega)$$

$f_n \in \mathcal{V}(0, T)$

By the Itô isometry

$$\begin{aligned} \mathbb{E}[(\xi_n - \xi_m)^2] &= \mathbb{E}[(\mathbb{E}[\xi_n - \xi_m] + \int_0^T (f_n(t, \omega) - f_m(t, \omega))d\eta)^2] \\ &= (\mathbb{E}[\xi_n - \xi_m])^2 + \int_0^T \mathbb{E}[(f_n(t, \omega) - f_m(t, \omega))^2]dt \\ &\longrightarrow 0 \end{aligned}$$

as $n, m \longrightarrow \infty$

So $\{f_n\}$ is a cauchy sequence in $L_2([0, T] \times \Omega)$ and hence converges to some

$f \in L_2([0, T] \times \Omega)$. Since $f_n \in \mathcal{V}(o, T)$ we have $f \in \mathcal{V}(o, T)$.

Again using the Itô isometry we see that

$$\xi = \lim_{n \rightarrow \infty} \xi_n = \lim_{n \rightarrow \infty} (\mathbb{E}[\xi_n] + \int_0^T f_n(t, \omega) d\eta) = \mathbb{E}[\xi] + \int_0^T f(t, \omega) d\eta$$

the limit being taken in $L_2(\Omega)$. Hence the representation in (2.2) holds for all $\xi \in L_2(\Omega)$.

The uniqueness follows from the Itô isometry: Suppose

$$\xi_n = \mathbb{E}[\xi] + \int_0^T f_1(t, \omega) d\eta_t(\omega) = \mathbb{E}[\xi] + \int_0^T f_2(t, \omega) d\eta_t(\omega)$$

with $f_1(t, \omega), f_2(t, \omega) \in \mathcal{V}(0, T)$. Then

$$0 = \mathbb{E}[(\int_0^T (f_1(t, \omega) - f_2(t, \omega)) d\eta_t(\omega))^2] = \int_0^T \mathbb{E}[(f_1(t, \omega) - f_2(t, \omega))^2] dt$$

and therefore $f_1(t, \omega) = f_2(t, \omega)$ for a.a. $(t, \omega) \in [0, T] \times \Omega$

Theorem 2.6 The Martingale Representation

Let $\eta(t) = (\eta_1(t), \dots, \eta_n(t))$ be n -dimensional. Suppose M_t is an \mathcal{F}_t^n -martingale with respect to P and $M_t \in L_2(\Omega)$ for all $t \geq 0$. Then there exists a unique stochastic process $g(s, \omega)$ such that $g \in \mathcal{V}^n(0, t)$ for all $t \geq 0$ and

$$M_t(\omega) = \mathbb{E}[M_0] + \int_0^t g(s, \omega) d\eta(s)$$

for all $t \geq 0$.

Proof (n=1)

By Theorem 2.5 applied to $T = t$, $\xi = M_t$, we have that for all t there exists a unique $f^t(s, \omega) \in L_2(\Omega)$ such that

$$\begin{aligned} M_t(\omega) &= \mathbb{E}[M_t] + \int_0^t f^{(t)}(s, \omega) d\eta_s(\omega) \\ &= \mathbb{E}[M_0] + \int_0^t f^{(t)}(s, \omega) d\eta_s(\omega) \end{aligned}$$

Now assume $0 \leq t_1 \leq t_2$. Then

$$M_{t_1} = \mathbb{E}[M_{t_2} \mid \mathcal{F}_{t_1}] = \mathbb{E}[M_0] + \mathbb{E}\left[\int_0^{t_2} f^{(t_2)}(s, \omega) d\eta_s(\omega) \mid \mathcal{F}_{t_1}\right] = \mathbb{E}[M_0] + \int_0^{t_1} f^{(t_2)}(s, \omega) d\eta_s(\omega) \quad (2.3)$$

But we also have

$$M_{t_1} = \mathbb{E}[M_0] + \int_0^{t_1} f^{(t_1)}(s, \omega) d\eta_s(\omega) \quad (2.4)$$

Hence, comparing (2.3) and (2.4) we get that

$$0 = \mathbb{E}\left[\left(\int_0^{t_1} (f^{(t_2)}(s, \omega) - f^{(t_1)}(s, \omega)) d\eta\right)^2\right] = \int_0^{t_1} \mathbb{E}[(f^{(t_2)}(s, \omega) - f^{(t_1)}(s, \omega))^2] ds$$

and therefore

$$f^{(t_1)}(s, \omega) = f^{(t_2)}(s, \omega) \quad \text{for a.a. } (s, \omega) \in [0, t_1] * \Omega$$

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

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

and then we obtain

$$M_t = \mathbb{E}[M_0] + \int_0^t f^{(t)}(s, \omega) d\eta_s(\omega) = \mathbb{E}[M_0] + \int_0^t f(s, \omega) d\eta_s(\omega)$$

for all $t \geq 0$

2.3 Stochastic Integral Representations

We will now follow closely the work done in [9] and [8].

Let ξ_t , $0 \leq t \leq T$ and η_t , $0 \leq t \leq T$ be martingales (see section 1.4) in the standard L_2 -space with a filtration \mathcal{F}_t , $0 \leq t \leq T$. We recall the well known Kunita-Wanatabe representation, see [8],

$$\xi_t = \xi_t^0 + \int_0^t \varphi d\eta_s$$

$0 \leq t \leq T$ where φ_t is a stochastic function with finite L_2 -norm

$$\|\varphi_t\|_{L_2} = (\mathbb{E}[\int_0^T |\varphi_t|^2 d[\eta]_t])^{\frac{1}{2}}$$

The martingales ξ_t^0 , $0 \leq t \leq T$ and η_t , $0 \leq t \leq T$ are orthogonal (see chapter 1, section 1.4.4) in the sense that $\mathbb{E}[\Delta\xi^0 \Delta\eta | \mathcal{F}_t] = 0$ for all intervals $\Delta = (t, t + \Delta t]$ where $\Delta\xi_t^0 = \xi_{t+\Delta t}^0 - \xi_t^0$ and $\Delta\eta_t = \eta_{t+\Delta t} - \eta_t$.

We denote the indicator (see chapter 1, section 1.3) of $\Delta \subseteq (0, T]$ by

$$1_\Delta(s) = \begin{cases} 1 & \text{if } s \in \Delta \\ 0 & \text{if } s \notin \Delta \end{cases}$$

If we identify the martingales $\xi_t = \mathbb{E}[\xi | \mathcal{F}_t]$ and $\xi_t^0 = \mathbb{E}[\xi^0 | \mathcal{F}_t]$ by the following corresponding variables, $\xi = \xi_T$ and $\xi^0 = \xi_T^0$ then the Kunita-Wanatabe representation turns out to be equivalent to

$$\xi = \xi^0 + \int_0^T \varphi d\eta_s$$

Here the integral term

$$\hat{\xi} = \int_0^T \varphi d\eta_s$$

represents the best approximation in L_2 -space to ξ by means of stochastic integral with η_t , $0 \leq t \leq T$ as integrator. By saying that $\hat{\xi}$ is the best approximate to ξ , we mean that

$$\int_\Omega |\xi - \hat{\xi}|^2 dP = \mathbb{E}[|\xi - \hat{\xi}|^2] = \inf\{\mathbb{E}[|\xi - \psi|^2]; \psi \in \kappa\}$$

Here, (Ω, \mathcal{F}, P) is the probability space corresponding to the $(d+r)$ -dimensional Brownian motion starting at 0, \mathbb{E} denotes expectation with respect to P and

$$\kappa = \{\psi : \Omega \rightarrow \mathcal{R}^n; \psi \in L_2(P) \text{ and } \psi \text{ is } \mathcal{G}_T\text{-measurable}\}$$

where \mathcal{G}_T is the σ -algebra generated by $\{\xi_s(\cdot), s \leq t\}$. Thus, if ξ cannot be achieved we insist that close to it there are $\hat{\xi}'$ s which are achievable. So in this case we are looking for the $\hat{\xi}$ which is the closest to ξ which we can make use of to approximate ξ . Thus, this $\hat{\xi}$ becomes the the best approximation to ξ ,

$$\hat{\xi} = \int_0^T \varphi d\eta_s$$

by means of stochastic integral with respect to the integrator η_t , $0 \leq t \leq T$.

2.4 The Cauchy Criterion for Convergence

For arbitrary Y and Z , the parallelogram law states that

$$\| Y + Z \|^2 + \| Y - Z \|^2 = 2\| Y \|^2 + 2\| Z \|^2$$

Now let $Y = \varphi_n - \varphi$ and $Z = \varphi - \varphi_m$. If we set Y and Z in the parallelogram law, we obtain

$$\mathbb{E}[(\varphi_n - \varphi_m)^2] \leq 2\mathbb{E}[(\varphi_n - \varphi)^2] + 2\mathbb{E}[(\varphi_m - \varphi)^2]$$

hence if $\varphi_n \rightarrow \varphi$ then

$$\lim_{n,m \rightarrow \infty} \mathbb{E}[(\varphi_n - \varphi_m)^2] = 0$$

or equivalently

$$\lim_{n,m \rightarrow \infty} \| \varphi_n - \varphi_m \| = 0$$

Conversely, it can be shown that if φ_n is a sequence of random variables for which

$$\lim_{n,m \rightarrow \infty} \mathbb{E}[(\varphi_n - \varphi_m)^2] = 0$$

then there exists a random variable φ for which

$$\varphi_n \rightarrow \varphi$$

2.5 Stochastic Derivatives

For the standard L_2 -space $H = L_2(\Omega) = L_2(\Omega, \mathcal{F}, P)$ of real random variables ξ :

$$\|\xi\| = \left(\int_{\Omega} |\xi(\omega)|^2 P d(\omega) \right)^{\frac{1}{2}} = (\mathbb{E}[\xi^2])^{\frac{1}{2}}$$

on the probability space (Ω, \mathcal{F}, P) . We recall the Itô type non-anticipating integration scheme involving an arbitrary filtration \mathcal{F}_t , $0 \leq t \leq T$, and a general *right-continuous in $L_2(\Omega, \mathcal{F}, P)$* martingale η_t , $0 \leq t \leq T$, as integrator. By right

continuity in $L_2(\Omega, \mathcal{F}, P)$, we mean $\|\eta_{t+\Delta t} - \eta_t\| \rightarrow 0$ as $\Delta t \rightarrow 0$.

The integrands are considered as elements of a standard functional space $L_2(\Omega \times [0, t])$ of measurable stochastic functions

$$\varphi = \varphi(\omega, t), \quad (\omega, t) \in \Omega \times [0, T],$$

with a norm

$$\|\varphi\|_{L_2} = \left(\int \int_{\Omega \times (0, T]} |\varphi|^2 P(d\omega) \times d[\eta]_t(\omega) \right)^{\frac{1}{2}} = \left(\mathbb{E} \int |\varphi|^2 d[\eta]_t \right)^{\frac{1}{2}}$$

given by means of a product type measure $P(d\omega) \times d[\eta]_t(\omega)$. Here $[\eta]_t$ is associated with a stochastic function $[\eta]_t$, $0 \leq t \leq T$ ($[\eta]_0 = 0$), which have monotone right-continuous paths such that

$$\mathbb{E}[\Delta[\eta] \mid \mathcal{F}_t] = \mathbb{E}[|\Delta\eta|^2 \mid \mathcal{F}_t]$$

for the increments $\Delta[\eta]$ and $\Delta\eta$ to the intervals $\Delta = (t, t + \Delta t] \subseteq (0, T]$.

Observe that, here the integrands can be characterized as adapted process (see chapter 1, definition 1.1.4) or equivalently, thanks to the direct product measure $P(d\omega) \times d[\eta]_t$ applied, as predictable processes (see chapter 1, definition 1.6.2).

In particular, for the Brownian motion η_t , $0 \leq t \leq T$, as integrator ($\mathbb{E}[\eta_t] = 0$, $\mathbb{E}[\eta_t^2] = t$), the deterministic function $[\eta]_t = t$, $0 \leq t \leq T$, is applicable.

If we consider the integrand as a simple function φ^h of the form

$$\varphi_s^h = \sum_{\Delta} \varphi_{\Delta}^h 1_{\Delta}(s)$$

$0 \leq s \leq T$ with φ_{Δ}^h as an \mathcal{F}_t -measurable variable on the interval $\Delta = (t, t + \Delta t]$ of the finite h -partition: $\sum \Delta = (0, T]$, $\Delta t \leq h$, the stochastic integrals can be defined as

$$\int_0^T \varphi^h d\eta_s = \sum_{\Delta} \varphi_{\Delta}^h \cdot \Delta\eta$$

with the summation referring to the partition intervals. Here it is assumed that

$$\mathbb{E}[\varphi_{\Delta}^h \Delta\eta_t]^2 = \mathbb{E}(|\varphi_{\Delta}^h|^2 \cdot \mathbb{E}[|\Delta\eta_t|^2 \mid \mathcal{F}_t]) = \mathbb{E}[|\varphi_{\Delta}^h|^2 \cdot \mathbb{E}[\Delta[\eta]_t \mid \mathcal{F}_t]] = \mathbb{E} \left[\int_{\Delta} |\varphi_{\Delta}^h|^2 d[\eta]_s \right] < \infty$$

$\Delta\eta_t = \eta_{t+\Delta t} - \eta_t$ and $\Delta[\eta]_t = [\eta]_{t+\Delta t} - [\eta]_t$, which gives

$$\mathbb{E}\left[\int_0^T \varphi^h d\eta_s\right]^2 = \mathbb{E}\left[\int_0^T |\varphi^h|^2 d[\eta]_s\right]$$

Therefore the general integrands φ are identified as limits

$$\varphi = \lim_{h \rightarrow 0} \varphi^h \quad \text{that is} \quad \|\varphi - \varphi^h\|_{L_2} \longrightarrow 0 \quad h \longrightarrow 0$$

of the appropriate simple functions φ^h , in the functional space $L_2(\Omega * (0, T])$.

A natural question that follows is: How the integrands

$$\varphi = \lim_{h \rightarrow 0} \varphi^h$$

can be determined for a given integral

$$\hat{\xi}_t = \int_0^t \varphi_s ds$$

for $0 \leq t \leq T$.

To give a reasonable answer, we consider

$$\varphi = \lim_{h \rightarrow 0} \varphi^h \quad \text{for} \quad \varphi_t^h = \mathbb{E}\left[\frac{\Delta \hat{\xi}_t}{\Delta t} \mid \mathcal{F}_t\right]$$

on the h -partition interval $\Delta = (t, t + \Delta t]$ with $\Delta \hat{\xi}_t = \hat{\xi}_{t+\Delta t} - \hat{\xi}_t$. In particular the above simple formula implies that there is no martingale $\hat{\xi}_t$, $0 \leq t \leq T$ which admits integral representation

$$\hat{\xi}_t = \int_0^t \varphi_s ds$$

but $\hat{\xi}_t \equiv 0$ since

$$\mathbb{E}[\Delta \hat{\xi}_t \mid \mathcal{F}_t] \equiv 0$$

Another question that arise is how its stochastic differential equation

$$d\hat{\xi}_t = \varphi d\eta_t$$

can be determined for $\hat{\xi}_t$, $0 \leq t \leq T$, a martingale with respect to the filtration generated by independent Brownian motions η_t , $0 \leq t \leq T$.

To answer this question, we consider the known type L_2 -integral

$$\hat{\xi}_t = \int_0^t \varphi d\eta_s$$

for $0 \leq t \leq T$.

We are to recall that the very L_2 - integral definitions involves approximations

$$\varphi = \lim_{h \rightarrow 0} \varphi^h$$

in the functional L_2 -space.

Theorem 2.7 *The integrands of*

$$\xi_t = \sum_{j=1}^n \int_0^t \varphi_j d\eta_s^j$$

$0 \leq t \leq T$ can be identified as the limits

$$\varphi_j = \lim_{h \rightarrow 0} \varphi_j^h$$

($j = 1, 2, \dots, n$) for the adapted simple function φ_j^h :

$$\varphi_j^h = \sum_{\Delta} \mathbb{E}\left[\frac{\Delta \xi_t}{\Delta t} \Delta \eta^j | \mathcal{F}_t\right] 1_{\Delta}(s)$$

on the corresponding h -partition intervals $\Delta = (t, t + \Delta t]$, with

$$\Delta \xi_t = \xi_{t+\Delta t} - \xi_t \quad \Delta \eta_t = \eta_{t+\Delta t} - \eta_t$$

($j = 1, 2, \dots, n$)

Proof

Consider some adapted simple approximations $\tilde{\varphi}_j^h$ for the integrands φ^j such that

$$\lim_{h \rightarrow 0} \|\varphi_j - \tilde{\varphi}_j^h\|_{L_2} = 0$$

In the h-partition scheme applied; clearly

$$\begin{aligned}
\mathbb{E}\left[\frac{\Delta\hat{\xi}_t}{\Delta t} \Delta \eta^j \mid \mathcal{F}_t\right] &= \mathbb{E}\left[\frac{1}{\Delta t} \int_{\Delta} \tilde{\varphi}_k^h d\eta_s^k \Delta \eta^j \mid \mathcal{F}_t\right] \\
&= \frac{\tilde{\varphi}_k^h}{\Delta t} \mathbb{E}\left[\int_{\Delta} d\eta_s^k \Delta \eta^j \mid \mathcal{F}_t\right] \\
&= \frac{\tilde{\varphi}_k^h}{\Delta t} \mathbb{E}[\Delta \eta^k \Delta \eta^j \mid \mathcal{F}_t] \\
&= \tilde{\varphi}_k^h \frac{\mathbb{E}[\Delta \eta^k \Delta \eta^j \mid \mathcal{F}_t]}{\Delta t} \\
&= \begin{cases} \tilde{\varphi}_k^h \frac{\mathbb{E}[\Delta \eta^j]^2 \mid \mathcal{F}_t]}{\Delta t} & \text{if } k = j; \\ \tilde{\varphi}_k^h \frac{\mathbb{E}[\Delta \eta^k \Delta \eta^j \mid \mathcal{F}_t]}{\Delta t} & \text{if } k \neq j. \end{cases} \\
&= \begin{cases} \tilde{\varphi}_k^h * \frac{\Delta t}{\Delta t} & \text{if } k = j; \\ \tilde{\varphi}_k^h * 0 & \text{if } k \neq j. \end{cases} \\
&= \begin{cases} \tilde{\varphi}_k^h & \text{if } k = j; \\ 0 & \text{if } k \neq j. \end{cases}
\end{aligned}$$

Therefore on the partition interval $\Delta = (t, t + \Delta]$ we have

$$\begin{aligned}
\varphi_j^h - \tilde{\varphi}_j^h &= \frac{1}{\Delta t} \mathbb{E}\left[\left(\Delta \hat{\xi} - \sum_{k=1}^n \int_{\Delta} \tilde{\varphi}_k^h d\eta_s^k\right) \Delta \eta^j \mid \mathcal{F}_t\right] \\
&= \sum_{k=1}^n \frac{1}{\Delta t} \mathbb{E}\left[\int_{\Delta} (\varphi_k - \tilde{\varphi}_k^h) d\eta_s^k \Delta \eta^j \mid \mathcal{F}_t\right]
\end{aligned}$$

since $\Delta \xi := \sum_{k=1}^n \int_{\Delta} \varphi_k d\eta_s^k$ and by majorant condition

$$\|\varphi_j^h - \tilde{\varphi}_j^h\|^2 \leq C \sum_{k=1}^n \frac{1}{\Delta t} \|\mathbb{E}\left[\int_{\Delta} (\varphi_k - \tilde{\varphi}_k^h) d\eta_s^k \Delta \eta^j \mid \mathcal{F}_t\right]\|^2 \quad (2.5)$$

C is a positive constant and where

$$\begin{aligned}
\|\mathbb{E}\left[\int_{\Delta} (\varphi_k - \tilde{\varphi}_k^h) d\eta_s^k \Delta \eta^j \mid \mathcal{F}_t\right]\|^2 &\leq \mathbb{E}\left[\left|\int_{\Delta} (\varphi_k - \tilde{\varphi}_k^h) d\eta_s^k\right|^2 \mid \mathcal{F}_t\right] \mathbb{E}\left[|\Delta \eta_j|^2 \mid \mathcal{F}_t\right] \\
&= \mathbb{E}\left[\left|\int_{\Delta} (\varphi_k - \tilde{\varphi}_k^h) d\eta_s^k\right|^2 \mid \mathcal{F}_t\right] \Delta t
\end{aligned}$$

Thus

$$\begin{aligned}
\|\mathbb{E}[\int_{\Delta}(\varphi_k - \tilde{\varphi}_k^h)d\eta_s^k \Delta\eta^j | \mathcal{F}]\|^2 &\leq \mathbb{E}\mathbb{E}[\int_{\Delta}(\varphi_k - \tilde{\varphi}_k^h)d\eta_s^k]^2 | \mathcal{F}]\Delta t \\
&= \mathbb{E}[\int_{\Delta}(\varphi_k - \tilde{\varphi}_k^h)d\eta_s^k]^2 \Delta t \\
&= \int_{\Delta} \|(\varphi_k - \tilde{\varphi}_k^h)\|^2 ds \Delta t
\end{aligned}$$

Now we see that if we substitute the preceding inequality into (2.5) we have

$$\|\varphi_j^h - \tilde{\varphi}_j^h\|^2 \Delta t \leq C \sum_{j=1}^n \int_{\Delta} \|\varphi_k - \tilde{\varphi}_k^h\|^2 ds$$

which results

$$\begin{aligned}
\|\varphi_j^h - \tilde{\varphi}_j^h\|_{L_2}^2 &= \int_0^T \|\varphi_j^h - \tilde{\varphi}_j^h\|^2 ds \\
&= \sum_{\Delta} \|\varphi_j^h - \tilde{\varphi}_j^h\|^2 \Delta t \\
&\leq C \sum_{k=1}^n \sum_{\Delta} \int_{\Delta} \|\varphi_k - \tilde{\varphi}_k^h\|^2 ds \\
&= C \sum_{k=1}^n \int_0^T \|\varphi_k - \tilde{\varphi}_k^h\|^2 ds \\
&= C \sum_{k=1}^n T \|\varphi_k - \tilde{\varphi}_k^h\|_{L_2}^2 \longrightarrow 0
\end{aligned}$$

as $h \longrightarrow 0$ hence

$$\|\varphi_j - \varphi_j^h\|_{L_2} \leq \|\varphi_j^h - \tilde{\varphi}_j^h\|_{L_2} + \|\varphi_j - \tilde{\varphi}_j^h\|_{L_2} \longrightarrow 0$$

as $h \longrightarrow 0$ and this completes the proof.

As we have already indicated earlier on, we are asking the following questions: Does a random variable $\xi \in L_2(P)$ allow a representation by the stochastic integral? In any case, how can we find the best integral approximation $\hat{\xi}$ to ξ :

$$\hat{\xi} = \int_0^T \varphi d\eta_s? \tag{2.6}$$

Here $\hat{\xi}$ is the projection of ξ onto the subspace $L_2(\eta)$ of all stochastic integrals with the considered integrator η_t , $0 \leq t \leq T$. In regard to these questions, it is interesting to consider a more general case with the subspace $L_2(\eta)$ of random variable $\hat{\xi}$ in $L_2(P)$ allowing the stochastic integral representation of the form

$$\hat{\xi} = \int_0^T \varphi d\eta_s$$

with respect to the orthogonal martingale (see section 1.5.1) η_t , $0 \leq t \leq T$ as integrator.

Proposition 3 *Prove that $\xi = \xi_T = \mathbb{E}[E] + \int_0^T \varphi d\eta_s$ then*

$$\mathbb{E}[\Delta\xi_t \cdot \Delta\eta_t \mid \mathcal{F}_t] = \mathbb{E}[\xi \cdot \Delta\eta_t \mid \mathcal{F}_t]$$

$$\begin{aligned} \mathbb{E}[\xi \cdot \Delta\eta_t \mid \mathcal{F}_t] &= \mathbb{E}[\mathbb{E}(\xi - \xi_t \mid \mathcal{F}_t) + \mathbb{E}[\Delta\xi_t \cdot \Delta\eta_t \mid \mathcal{F}_t] + \xi_t \mathbb{E}[\Delta\eta_t \mid \mathcal{F}_t]] \\ &= \mathbb{E}[\xi - \xi_t] \mathbb{E}[\Delta\eta_t] + \mathbb{E}[\Delta\xi_t \cdot \Delta\eta_t \mid \mathcal{F}_t] + \xi_t \mathbb{E}[\Delta\eta_t] \\ &= \mathbb{E}[\Delta\xi_t \Delta\eta_t \mid \mathcal{F}_t] \end{aligned}$$

To address the above questions, we are going to answer by means of the non-anticipating stochastic derivative in the framework of Itô stochastic calculus.

Definition 2.5.1 *For the random variable ξ in $L_2(P)$, the “stochastic derivative”*

$$\mathcal{D}\xi = \mathcal{D}_\eta \xi = \varphi = \varphi_s \quad 0 \leq s \leq T$$

with respect to η_t , $0 \leq t \leq T$, in the functional space $L_2(\Omega \times [0, T])$ is defined as the limit

$$\varphi = \lim_{h \rightarrow 0} \varphi^h \quad \text{that is} \quad \|\varphi - \varphi^h\|_{L_2} \rightarrow 0 \quad h \rightarrow 0 \quad (2.7)$$

of simple functions φ^h of the form

$$\varphi_s^h = \sum_{\Delta} \varphi_{\Delta}^h 1_{\Delta}(s) \quad 0 \leq s \leq T$$

with $\varphi_{\Delta}^h \in L_2(P)$ defined as

$$\varphi_{\Delta}^h = \mathbb{E}\left[\xi \cdot \frac{\Delta\eta}{\|\Delta\eta\|_t^2} \mid \mathcal{F}_t\right] \quad (2.8)$$

for each interval $\Delta = (t, t + \Delta t]$ of the finite h -partition of $(0, T] : \sum \Delta = (0, T]$, $\Delta t \leq h \rightarrow 0$ where

$$\Delta\eta = \Delta\eta \cdot 1_{\|\Delta\eta\|_t > 0} \quad \text{in } L_2(P)$$

and

$$\|\Delta\eta\|_t = (\mathbb{E}[|\Delta\eta|^2 \mid \mathcal{F}_t])^{\frac{1}{2}}$$

The stochastic function $\mathcal{D}_\eta\xi$ is preferred to as the stochastic derivative with respect to the integrator η .

Let $L_0(P) \subseteq L_2(P)$ be a subspace of all variables ξ characterized by $\mathcal{D}_\eta\xi$, and $L_2(\eta) \subseteq L_2(P)$ the subspace of all stochastic integrals with integrator η . We can now consider the following theorem on integral representation for random variables in $L_2(P)$, see [8].

Theorem 2.8 *The stochastic derivative $\mathcal{D}\xi$ defined in definition 2.5.1 is well defined for any $\xi \in L_2(P)$ and ξ admits unique integral representation in terms of its derivative:*

$$\xi = \xi^0 + \int_0^T (\mathcal{D}\xi) d\eta_s \quad (2.9)$$

with the corresponding $\xi^0 \in L_0(P)$ where $L_0(P) \subseteq L_2(P)$ is the subspace of all variables ξ^0 characterised by $\Delta\xi^0 = 0$. Moreover, the following orthogonal decomposition holds true

$$L_2(P) = L_0(P) \oplus L_2(\eta)$$

where $L_2(\eta) \subseteq L_2(P)$ is the subspace of all stochastic integrals.

\oplus means the orthogonal sum of $L_0(P)$ and $L_2(\eta)$

Proof

Using the monotone h -partitions, we have

$$L_2(\eta) = \lim_{h \rightarrow 0} \sum \oplus L_2(\Delta\eta) \quad (2.10)$$

as the limit of the indicated orthogonal sums. Their components $L_2(\Delta\eta)$ are taken as subspaces of the corresponding variables in $L_2(P)$ of the form $\psi \cdot \Delta \eta$. The multipliers ψ are \mathcal{F}_t -measurable for the increments

$$\Delta\eta = \eta_{t+\Delta t} - \eta_t$$

to the h -partition interval $\Delta = (t, t + \Delta t]$.

The projection of ξ onto $L_2(\Delta\eta)$ is $\psi_\Delta^h \cdot \Delta \eta$. We denote the multipliers φ_Δ^h as

$$\varphi_\Delta^h = \mathbb{E}\left[\xi \cdot \frac{\Delta\eta}{\|\Delta\eta\|_t^2} \mid \mathcal{F}_t\right]$$

Thus

$$\mathbb{E}[|\varphi_\Delta^h \cdot \Delta \eta|^2] < \infty$$

since

$$|\varphi_\Delta^h|^2 \mathbb{E}[|\Delta \eta|^2 \mid \mathcal{F}_t] \leq \mathbb{E}[\xi^2 \mid \mathcal{F}_t]$$

Then

$$\begin{aligned} \mathbb{E}[(\xi - \varphi_\Delta^h \Delta \eta)(\psi \Delta \eta) \mid \mathcal{F}_t] &= \psi \mathbb{E}[(\xi \Delta \eta - \varphi_\Delta^h |\Delta \eta|^2) \mid \mathcal{F}_t] \\ &= \psi \mathbb{E}[\xi \Delta \eta \mid \mathcal{F}_t] - \psi \varphi_\Delta^h \mathbb{E}[|\Delta \eta|^2 \mid \mathcal{F}_t] \\ &= \psi \mathbb{E}[\xi \Delta \eta \mid \mathcal{F}_t] - \psi \mathbb{E}\left[\xi \cdot \frac{\Delta\eta}{\|\Delta\eta\|_t^2} \mid \mathcal{F}_t\right] \mathbb{E}[|\Delta \eta|^2 \mid \mathcal{F}_t] \\ &= \psi \mathbb{E}[\xi \Delta \eta \mid \mathcal{F}_t] - \psi \mathbb{E}[\xi \Delta \eta \mid \mathcal{F}_t] \\ &= 0 \end{aligned}$$

This implies that the orthogonality condition

$$\mathbb{E}[(\xi - \varphi_\Delta^h \Delta \eta)(\psi \Delta \eta)] = 0$$

holds.

Therefore the projection of ξ onto the pre-limit orthogonal sums in (2.10) are

$$\sum_{\Delta} \varphi_\Delta^h \Delta \eta = \int_0^T \varphi^h d\eta_s$$

where the integrands φ^h are the simple functions with the values $\psi = \varphi_{\Delta}^h$ for the interval $\Delta = (t, t + \Delta t]$. The projection $\hat{\xi}$ of ξ onto the subspace $L_2(\eta)$ of all integrals is represented by a particular integral since it is in fact a limit

$$\hat{\xi} = \int_0^T \varphi d\eta_s = \lim_{h \rightarrow 0} \int_0^T \varphi^h d\eta_s \quad \text{in } L_2(P)$$

Note that the integrand φ is the limit of the simple functions φ^h , according to

$$\left\| \int_0^T \varphi d\eta_s - \int_0^T \varphi^h d\eta_s \right\| = \|\varphi - \varphi^h\|_{L_2}$$

Thus in the representation

$$\xi = \xi^0 + \int_0^T (\mathcal{D}\xi) d\eta_s$$

with the integrand $\varphi = \mathcal{D}\xi$, the difference

$$\xi^0 = \xi - \int_0^T \varphi d\eta_s = \xi - \int_0^T (\mathcal{D}\xi) d\eta_s$$

Now

$$\mathcal{D}\xi^0 = \mathcal{D}\xi - \mathcal{D}\xi = 0$$

Thus the difference

$$\xi^0 = \xi - \int_0^T \varphi d\eta_s$$

is orthogonal to $L_2(\eta)$. This completes the proof.

We want to stress that, in representation (2.9),

$$\hat{\xi} = \int_0^T (\mathcal{D}\xi) d\eta_s$$

is the best integral approximation to ξ namely

$$\|\xi - \hat{\xi}\| = \min_{\psi} \left\| \xi - \int_0^T \psi d\eta_s \right\|$$

where the minimum is considered over all possible integrands ψ .

Corollary 3 For orthogonal martingale η_t , $0 \leq t \leq T$ and the subspace $L_2(\eta)$ of the variables in $L_2(P)$ defined in (2.6), the projection $\hat{\xi}$ of ξ onto $L_2(\eta)$ is

$$\hat{\xi} = \int_0^T (\mathcal{D}\xi) d\eta_s$$

with the integrands

$$\mathcal{D}\xi = \lim_{h \rightarrow 0} \sum_{\Delta} \mathbb{E}[\xi \cdot \frac{\Delta\eta}{\|\Delta\eta\|_t^2} \mid \mathcal{F}_t] 1_{\Delta}(s) \quad 0 \leq s \leq T$$

as the stochastic derivatives with respect to the corresponding integrators (2.7)-(2.8)

If we consider the deterministic case, we can define the Radon Nykodim derivative through the h -partitions of the limit

$$\mathcal{D}\xi = \lim_{h \rightarrow 0} \sum_{\Delta} \frac{\Delta\xi}{\Delta\eta} 1_{\Delta}(t) = \lim_{h \rightarrow 0} \sum_{\Delta} [\Delta\xi \frac{\Delta\eta}{|\Delta\eta|^2}] 1_{\Delta}(t)$$

In the stochastic case, the stochastic analogue of the above limit represents the stochastic Radon-Nykodim derivative $\mathcal{D}\xi = \varphi$ in the integrands space $L_2(\Omega * (0, T])$, namely

$$\mathcal{D}\xi = \lim_{h \rightarrow 0} \sum_{\Delta} \mathbb{E}[\Delta\xi \frac{\Delta\eta}{\|\Delta\eta\|_s^2} \mid \mathcal{F}_s] 1_{\Delta}(t)$$

where $\Delta = (t, t + \Delta t]$.

Note that the increments are on the interval of the partition $\Delta = (t, t + \Delta t]$ and the expectations are with respect to the element of the filtration \mathcal{F}_s (the initial time of the interval). This provides predictability (see chapter 1, section 1.6) and adaptedness (see chapter 1, definition 1.1.4).

The above limit can be redefined as

$$\mathcal{D}\xi = \lim_{h \rightarrow 0} \mathbb{E}[\xi \frac{\Delta\eta}{\|\Delta\eta\|_s^2} \mid \mathcal{F}_s] 1_{\Delta}(t)$$

to be more precise, $\mathcal{D}\xi = \varphi$ is identified as the limit of the simple function with the corresponding values

$$\varphi_{\Delta}^h = \mathbb{E}[\xi \frac{\Delta\eta}{\|\Delta\eta\|_s^2} \mid \mathcal{F}_s]$$

on the set Δ of the h -partition.

Chapter 3

Methods in quadratic hedging

Having considered some important aspects of stochastic analysis and stochastic calculus in continuous time framework, this chapter will take a new direction to consider the discrete time market, where the trading occurs at discrete time, as we move towards our main objective, which is to generalize minimal variance hedging to the case of a discrete time market driven by Markov process. We will follow closely the works in [3], [5], [10], [12], [15], [18],[19], [20].

A contingent claim is a random variable $H(\omega)$ representing a payoff at time $t = T$. Note that $H(\omega)$ is \mathcal{F}_T measurable. It can be thought of as part of a contract, that is a legal agreement between two parties, a buyer and a seller, to trade an underlying asset at a date in the future. So we consider an agent who knows, at time $t = 0$, that at some future time T he must make a payment $H(\omega)$ but the size of the payment depends on a number of factors which are still undetermined and not within his control.

This agent would like to set aside a fixed amount of money V_0 at time $t = 0$ and be assured that this will enable him to make payments at time T . A possible strategy will be to set aside an amount equal to the maximum possible payment size, $\sup_{\omega \in \Omega} H(\omega)$, if this maximum is finite.

A reasonable strategy entails setting aside less money but investing this money

in the market so that the terminal value of this investment is as close as possible to the value of the liability $H(\omega)$. This is the process of hedging the risk inherent in the random payment. Thus in this chapter we consider methods in quadratic hedging with focus on minimal variance hedging in complete and incomplete markets.

3.1 Discrete-Time Economy

To describe a financial market operating in discrete time, we let (Ω, \mathcal{F}, P) be a fixed probability space. The filtration model of information, $\mathbb{F} = \{\mathcal{F}_t; t = 0, 1, \dots\}$ will be a collection of σ -algebras with $\mathcal{F}_t \subseteq \mathcal{F}_{t+1} \subseteq \dots \subseteq \mathcal{F}$ for all t . We can view \mathcal{F}_t as the information available to all investors on the market at time t and is sometimes called the σ -algebra of events upto time t . The horizon T will often correspond to the maturity of the contract where $T \in \mathbb{Z}^+ = \{0, 1, 2, \dots\}$. From now, we will assume that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_\infty = \mathcal{F}$ and for all $\omega \in \Omega$, $P(\{\omega\}) > 0$. The market consists of $(d + 1)$ financial assets whose prices at time t are given by non-negative random variables $S_t^0, S_t^1, \dots, S_t^d$, measurable with respect to \mathcal{F}_t , see chapter 1, definition 1.1.2. That is the investors know the past and the present prices but obviously not the future ones.

We assume that one asset, say S^0 , has a strictly positive price. We then use S^0 as the numeraire and immediately pass to quantities discounted with S^0 . This means that asset 0 has (discounted) price 1 at all times and the other assets' (discounted) prices are $X^i = \frac{S^i}{S^0}$ for $i = 1, 2, \dots, d$. From now on, all subsequently appearing quantities will be expressed in discounted units.

3.1.1 Trading strategies and the value processes

Having specified the information describing the model, the next step is to define several quantities of interest.

We consider a trading strategy given by two stochastic processes $(\theta_t)_{t \geq 0}$ and $(\eta_t)_{t \geq 0}$ where θ is a d -dimensional predictable process (see chapter 1, definition 1.6.2) and η is adapted (see chapter 1, definition 1.1.4)to the filtration \mathcal{F}_t . Thus measurability holds for all t 's.

In such a strategy, θ_t^i describes the number of shares of asset i held at time t where $i = 1, 2, \dots, d$ and η_t is the amount invested in the bond at time t . A trading strategy describes an investor's portfolio as carried forward from $t = 0$ to time $t = T$. It should be a rule that specifies the investor's position in each security at each point in time and in each state of the world, see [18]. This rule should allow the investor to choose a position in the securities based on all the available information.

The value of the portfolio (θ_t, η_t) at any time t , is given by

$$V_t = \theta_t X_t + \eta_t \tag{3.1}$$

The value process V_t describes the total value of the portfolio at each point in time. Note that V_t is an adapted stochastic process.

Denote $\Delta X_{t+1} = X_{t+1} - X_t$ for the change in the value of the stochastic process X_t between times t and $t + 1$. Thus the cumulative gains, G_t , from trade up to time t are given by

$$G_t(\theta) = \sum_{j=0}^t \theta_j \Delta X_j = \sum_{j=0}^{t-1} \theta_j (X_{j+1} - X_j), \quad t \geq 0 \quad \text{and} \quad G_0 \tag{3.2}$$

The gain process G_t is a random variable that describes the total profit or loss generated by the portfolio between times $t = 0$ and time $t = T$ if T is our final time.

3.1.2 The cost process

Consider now the flow of capital from time t to time $t + 1$. First, the seller of the contract may adjust the number of stocks to θ_{t+1} by buying additional $(\theta_{t+1} - \theta_t)$ stocks, which leads to the costs $(\theta_{t+1} - \theta_t)X_t$. The new portfolio (θ_{t+1}, η_t) is then held until time $t + 1$. Thus, the investment in stocks leads to a gain $\theta_{t+1}(X_{t+1} - X_t)$. These actions are responsible for changes in the value process, which add up to

$$V_{t+1}(\varphi) - V_t(\varphi) = (\theta_{t+1} - \theta_t)X_t + \theta_{t+1}(X_{t+1} - X_t) + (\eta_{t+1} - \eta_t)$$

The first and the third terms on the right hand side represent costs to the hedger, whereas the second term is trading gains obtained from the strategy φ during $(t, t + 1]$.

The cumulative costs, C_t , that arises from trading because of the fluctuations of the price process X_s up to time t , incurred by using the portfolio (θ, η) is given by

$$C_t = V_t - \sum_{j=1}^t \theta_j \Delta X_j = V_t - G_t(\theta) \quad \text{by (3.2)} \quad (3.3)$$

Note that if its cumulative cost process, C_t , is constant over time, that is $C_0 = C_1 = \dots = C_T$ then the strategy is called self-financing. This is equivalent to $(\theta_{t+1} - \theta_t)X_{t+1} + \eta_{t+1} - \eta_t = 0$ for all $t \geq 0$.

3.1.3 Self-financing portfolio

Definition 3.1.4 *A strategy is called “self-financing” if its value process V_t is given by*

$$V_t = V_0(\varphi) + \sum_{j=1}^t \theta_j \Delta X_j = V_0(\varphi) + G_t(\theta) \quad (3.4)$$

where V_0 is the amount invested at time 0. Thus, for a self-financing strategy, the current value of the portfolio at time t is strictly the initial invested amount

plus trading gains. Intuitively, once the strategy has started, no money is added to or withdrawn from the portfolio, thus any change in the portfolio's value is due to a gain or loss in the investments. Therefore, any fluctuations in the stock price is neutralised by the rebalancing θ and η since there is no inflow or outflow of capital.

Definition 3.1.5 *A portfolio $\varphi(t)$ which is self-financing is called “admissible” if the corresponding value process $V^\varphi(t)$ is (t, ω) a.s. lower bounded. That is there exists $K = K(\varphi) < \infty$ such that*

$$V^\varphi(t, \omega) \geq -K$$

for a.a $(t, \omega) \in L_2(\Omega \times [0, T])$

Note that the restriction reflects a natural condition in real life finance: There must be a limit to how much debt the creditors can tolerate.

3.1.6 Attainability and Completeness

We say a contingent claim H is attainable if there exists a self-financing strategy with $V_T = H - P$ a.s.

This means that if we start with some initial fortune, we can find an admissible portfolio $\varphi(t)$ which generates the value $V^\varphi(T)$ at time T which is a.s equals H . Thus a claim is attainable if and only if it can be represented as a constant H_0 plus a stochastic integral with respect to the discounted stock price process. Thus for $H_0 = V^\varphi(0)$, equation (3.4) can be written as

$$H = H_0 + \sum_{j=0}^{t-1} \theta_j (X_{j+1} - X_j) = H_0 + G_t(\theta) - P \quad \text{a.s.} \quad (3.5)$$

A financial market is said to be complete if all claims are attainable. Now we consider the seller of the option with some contingent claim H payable at time T

whose objective is to reduce risk associated with the contingent claim. Suppose that it is possible to determine a self-financing strategy which replicates the claim perfectly, such that there exists a dynamic self-financing strategy φ , which sets out with some amount $V_0(\varphi)$ and has terminal value $V_T(\varphi) = H$, P -a.s, in this case, the initial value $V_0(\varphi)$ is the only reasonable price for the contingent claim H . The claim is then said to be attainable, and $V_0(\varphi)$ is exactly the no-arbitrage price of H .

3.2 Risk-minimising strategies

In more realistic situations, contingent claims cannot be hedged perfectly by use of a self-financing strategy. It is therefore relevant to consider the problem of defining an optimal criterion, see [19]. One possible objective would be to minimize the conditional expected value of the square of the costs occurring during the next period, subject to the condition that the terminal value at time T of the strategy is equal to the contingent claim, H , that is $V_T(\varphi) = H$. Thus in general, it is not possible to minimise via a self-financing strategy, and we therefore allow for the possibility of adding and withdrawing of capital after the initial time 0.

Definition 3.2.1 *The “risk process of φ ” is defined by*

$$R_t(\varphi) = \mathbb{E}[(C_T(\varphi) - C_t(\varphi))^2 \mid \mathcal{F}_t] \quad t \geq 0 \quad (3.6)$$

Since the contingent claim H is \mathcal{F}_T -measurable and η is adapted, see chapter 1, definition 1.1.4. Therefore, we can always find risk minimisation strategies with $V_T = H$ provided $H \in L_2(P)$.

The simplest is “wait, then pay” where $\theta \equiv 0$ and $\eta_t = HI_{t=T}$. Note that these strategies will not be self-financing; in fact (3.4) tells us that there is a self-financing risk-minimising strategy φ with $V_T(\varphi) = H$ if and only if H admits a

representation of the form (3.5). In that case, the cost process $C(\varphi)$ is constant and the risk process $R(\varphi)$ is identically 0.

Definition 3.2.2 *A risk-minimisation strategy φ is called "risk minimising" if for any risk minimising strategy $\bar{\varphi}$ such that $V_T(\bar{\varphi}) = V_T(\varphi)$ P - a.s we have*

$$R_t(\varphi) \leq R_t(\bar{\varphi}) \quad p - \text{ a.s } \text{ for every } t \geq 0 \quad (3.7)$$

Any self-financing strategy is clearly risk-minimizing since $R_t(\varphi) \equiv 0$.

3.2.3 Mean Self-financing Strategies

Although risk minimisation strategies with $V_T(\varphi) = H$ will in general not be self-financing, it turns out that good risk minimisation strategies are still "self-financing on average" (see [15]) in the following sense:

Definition 3.2.4 *A risk minimisation strategy φ is "mean-self-financing" if its cost process $C = (C_t)_{t \geq 0}$ is a martingale.*

Lemma

Any risk-minimising strategy φ is also mean self-financing.

Proof

Fix $t_0 \in [0, T]$ and define $\tilde{\varphi}$ by setting $\tilde{\theta} := \theta$ and

$$\tilde{\theta} X_t + \tilde{\eta}_t = V_t(\tilde{\varphi}) := V_t(\varphi) I_{[0, t_0)}(t) + \mathbb{E}[V_T(\varphi) - \sum_{j=0}^T \theta_j \Delta X_j \mid \mathcal{F}_t] I_{(t_0, T]}(t)$$

Then $\tilde{\theta}$ is a risk-minimisation strategy with $V_T(\tilde{\varphi}) = V_T(\varphi)$, $C_T(\varphi) = C_T(\tilde{\varphi})$ and $C_{t_0}(\tilde{\varphi}) = \mathbb{E}[C_T(\tilde{\varphi}) \mid \mathcal{F}_{t_0}]$, we have

$$C_T(\varphi) - C_{t_0}(\varphi) = C_T(\tilde{\varphi}) - C_{t_0}(\tilde{\varphi}) + \mathbb{E}[C_T(\tilde{\varphi}) \mid \mathcal{F}_{t_0}] - C_{t_0}(\varphi)$$

This implies that

$$R_{t_0} = R_{t_0}(\tilde{\varphi}) + (C_{t_0}(\tilde{\varphi}) - \mathbb{E}[C_T(\varphi) | \mathcal{F}_{t_0}])^2$$

Because φ is risk-minimising we conclude that

$$C_{t_0}(\varphi) = \mathbb{E}[C_T(\varphi) | \mathcal{F}_{t_0}]$$

P a.s and since t_0 is arbitrary, the result follows.

This lemma implies that the value process of a risk-minimising strategy has to be a martingale. Thus, $V_t(\varphi) = E[H | \mathcal{F}_t]$. This motivates the use of the well-known Galtchouk-Kunita-Wanatabe decomposition to find the risk-minimising strategies, see [13] and the references therein. For $H \in L_2(P)$, the decomposition is uniquely written as

$$H = \mathbb{E}[H | \mathcal{F}_0] + \sum_{j=0}^{T-1} \theta_j^H \Delta X_j + L_T^H$$

P a.s for some $\theta^H \in L^2(P)$ and L^H -martingale.

We characterise the risk-minimising strategy by two properties:

1. Its cost process C must be a martingale.
2. This martingale C , must be orthogonal to the martingale part M of the prices of X .

The first characterisation shows that the strategy is no longer self-financing but still remains mean self-financing. M in the second characterisation is the martingale part of X in the Doob Meyer decomposition, (see chapter 1, Theorem 1.6). If we relate this to the conditions on the contingent claim, it shows that there exists a risk-minimizing strategy for H if and only if H admits a decomposition of the form

$$H = H_0 + \sum_{j=0}^{T-1} \theta_j^H (X_{j+1} - X_j) + L_T^H \tag{3.8}$$

P a.s where L^H is a martingale orthogonal to M .

This decomposition form has a financial importance in the sense that it directly provides the risk-minimising strategy for H . That is the stock component θ is given by θ_j^H and η is determined by the requirement that the cost process C_t should coincide with $H_0 + L^H$.

Note that in the case of finite discrete time, θ_j^H and L^H can be computed recursively backward in time.

Proposition 4 *The existence of an optimal strategy is equivalent to a decomposition*

$$H = H_0 + \sum_{j=0}^T \theta_j^H \Delta X_j + L_T^H \quad (3.9)$$

with H_0 constant and $L^H = (L_t^H)_{t \geq 0}$ a square integrable martingale orthogonal to M .

Proof

For a decomposition

$$H = H_0 + \sum_{j=0}^T \theta_j \Delta X_j + L_T^H$$

with L^H -square integrable martingale orthogonal to M , the cost process associated with the strategy (θ, η) defined by

$$\theta = \theta^h, \quad \eta = V - \theta \cdot X \quad \text{and} \quad V_t = H_0 + \sum_{j=0}^T \theta_j \Delta X_j + L_t^H$$

for $t \geq 0$ is given by

$$C_t = V_t - G_t(\theta) = (H_0 + \sum_{j=0}^T \theta_j \Delta X_j + L_t^H) - (\sum_{j=0}^T \theta_j \Delta X_j) = H_0 + L_t^H$$

Thus we see that

$$H = H_0 + \sum_{j=0}^T \theta_j \Delta X_j + L_T^H = H_0 + L_T^H + \sum_{j=0}^T \theta_j \Delta X_j = C_T + \sum_{j=0}^T \theta_j \Delta X_j$$

and so (θ, η) is optimal.

Conversely, an optimal strategy leads to the decomposition

$$H = C_T + \sum_{j=0}^T \theta_j \Delta X_j = C_T + \sum_{j=0}^T \theta_j \Delta X_j + C_0 - C_0 = C_0 + \sum_{j=0}^T \theta_j \Delta X_j + (C_T - C_0)$$

Thus we have

$$H = H_0 + \sum_{j=0}^T \theta_j \Delta X_j + L_T^H$$

with

$$\theta^H = \theta, \quad L_t^H = C_t - C_0 \quad \text{for } t \geq 0 \quad \text{and} \quad H_0 = C_0$$

This completes the proof.

3.3 Quadratic Hedging Approaches

We now focus on three types of quadratic approaches to hedge a given square integrable contingent claim

1. Local risk minimisation (LRM) and
2. Mean-variance hedging (MVH)
3. Minimal variance hedging

3.3.1 Local risk Minimisation

The local risk-minimisation strategy keeps the replicating condition and relaxes the self-financing condition (see [13]). Taking H as our contingent claim, we look for risk-minimisation strategy φ with $V_T(\varphi) = H$. Note that this strategy cannot be always self-financing, so we should choose the optimal trading strategy

to minimise the incremental costs incurred from adjusting the portfolio at each hedging time. In discrete time framework, we consider a situation where trading is only done at dates $t = 0, 1, \dots$. See [15]. At time t we choose the numbers θ_{t+1} of shares to be held at time period $(t, t + 1]$ and the number η_t the amount invested in the bond at time t . Observe that predictability of θ forces us to determine the date $t + 1$ holdings, θ_{t+1} , already at date t . The time t portfolio is $\varphi_t = (\theta_t, \eta_t)$ and its value is $V_t(\varphi) = \theta_t^{tr} X_t + \eta_t$ where θ_t^{tr} means the transpose of θ_t . Since we want to minimise risk locally, we now consider the incremental cost incurred by adjusting the portfolio from θ_t to θ_{t+1} . Because θ_{t+1} is already chosen at time t with price given by X_t , this cost increment is

$$\begin{aligned}
C_{t+1}(\varphi) - C_t(\varphi) &= \theta_{t+1}^{tr} X_t + \eta_{t+1} - \theta_t^{tr} X_t - \eta_t \\
&= (\theta_{t+1} - \theta_t) X_t + \eta_{t+1} - \eta_t \\
&= \theta_{t+1}^{tr} X_t - \theta_t^{tr} X_t + \eta_{t+1} - \eta_t \\
&= \theta_{t+1}^{tr} X_t - \theta_t^{tr} X_t + (V_{t+1}(\varphi) - \theta_{t+1} X_{t+1}) - (V_t(\varphi) - \theta_t^{tr} X_t) \\
&= V_{t+1}(\varphi) - V_t(\varphi) - \theta_{t+1}^{tr} (X_{t+1} - X_t) \\
&= \Delta V_{t+1}(\varphi) - \theta_{t+1}^{tr} \Delta X_{t+1}
\end{aligned}$$

with the difference operator $\Delta \cup_{t+1} = \cup_{t+1} - \cup_t$ for any discrete-time stochastic process \cup .

The quadratic criterion for this local risk minimising strategy is given by minimising

$$\mathbb{E}[(C_{t+1}(\varphi) - C_t(\varphi))^2 \mid \mathcal{F}_t] \tag{3.10}$$

with respect to the time t , control variable θ_{t+1} and η_t .

By using the expression for $C_{t+1}(\varphi) - C_t(\varphi)$ and the fact that $V_t(\varphi)$ does not influence the conditional variance given \mathcal{F}_t , we can write,

$$\begin{aligned}
\mathbb{E}[(C_{t+1}(\Delta) - C_t(\varphi))^2 \mid \mathcal{F}_t] &= \text{Var}[V_{t+1}(\varphi) - \theta_{t+1}^{tr} \Delta X_{t+1} \mid \mathcal{F}_t] \\
&\quad + (\mathbb{E}[V_{t+1}(\varphi) - \theta_{t+1}^{tr} \Delta X_{t+1} \mid \mathcal{F}_t] - V_t(\varphi))^2
\end{aligned}$$

Since the first term on the right hand side does not depend on η_t , it is clearly optimal to choose η_t in such a way that

$$V_t(\varphi) = \mathbb{E}[V_{t+1}(\varphi) - \theta_{t+1}^{tr} \Delta X_{t+1} \mid \mathcal{F}_t] \quad (3.11)$$

This is equivalent to

$$0 = \mathbb{E}[\Delta V_{t+1}(\varphi) - \theta_{t+1}^{tr} \Delta X_{t+1} \mid \mathcal{F}_t] = \mathbb{E}[C_{t+1}(\varphi) - C_t(\varphi) \mid \mathcal{F}_t] \quad (3.12)$$

so that an optimal strategy should again be mean self-financing. In order to determine the holdings in the hedging portfolio at a certain time, we need to solve backward in time from maturity of the option.

Since $V_T = H$ is fixed, (3.11) implies by a backward induction argument that for the purpose of minimising

$$\mathbb{E}[(C_{t+1}(\varphi) - C_t(\varphi))^2 \mid \mathcal{F}_t]$$

at time t , the value $V_{t+1}(\varphi)$ may be considered as given.

Therefore it only remains to minimise

$$Var[V_{t+1}(\varphi) - \theta_{t+1}^{tr} \Delta X_{t+1} \mid \mathcal{F}_t]$$

with respect to \mathcal{F}_t -measurable quantity θ_{t+1} . This can be achieved if and only if

$$Cov[V_{t+1}(\varphi) - \theta_{t+1}^{tr} \Delta X_{t+1}, \Delta X_{t+1} \mid \mathcal{F}_t] = 0 \quad (3.13)$$

Using the Doob decomposition of X (see chapter 1, Theorem 1.6) into a martingale M and a predictable process A (see chapter 1, definition 1.6.2) given by $M_0 = 0 = A_0$,

$$\Delta A_{t+1} = \mathbb{E}[\Delta X_{t+1} \mid \mathcal{F}_t] \quad \text{and} \quad \Delta M_{t+1} = \Delta X_{t+1} - \Delta A_{t+1}$$

We can therefore rewrite (3.13) as

$$\begin{aligned} 0 &= Cov[C_{t+1}(\varphi) - C_t(\varphi), \Delta M_{t+1} \mid \mathcal{F}_t] \\ &= \mathbb{E}[(C_{t+1}(\varphi) - C_t(\varphi)) \Delta M_{t+1} \mid \mathcal{F}_t] \end{aligned}$$

which says that the product of the two martingales $C(\varphi)$ and M must be a martingale (see chapter 1, Example 1.2) or that $C(\varphi)$ and M must be strongly orthogonal under P . Thus in discrete case

a suitably integrable strategy φ is locally risk-minimising if and only if its cost process $C(\varphi)$ is a martingale and strongly orthogonal to the martingale part of X .

Suppose for simplicity that $d = 1$, since θ_{t+1} is \mathcal{F}_t -measurable, we can solve for θ_{t+1} from (3.13) to obtain

$$\begin{aligned}\theta_{t+1} &= \frac{\text{Cov}[V_{t+1}(\varphi), \Delta X_{t+1} \mid \mathcal{F}_t]}{\text{Var}[\Delta X_{t+1} \mid \mathcal{F}_t]} \\ &= \frac{\mathbb{E}[V_{t+1}(\varphi) \Delta M_{t+1} \mid \mathcal{F}_t]}{\mathbb{E}[(\Delta M_{t+1})^2 \mid \mathcal{F}_t]}\end{aligned}$$

If we make use of

$$\mathbb{E}[\theta_{t+1} \Delta X_{t+1} \mid \mathcal{F}_t] = \theta_{t+1} \Delta A_{t+1}$$

and substitute it in (3.11), we obtain

$$\begin{aligned}V_t(\varphi) &= \mathbb{E}[V_{t+1}(\varphi) - \theta_{t+1} \Delta A_{t+1} \mid \mathcal{F}_t] \\ &= \mathbb{E}[V_{t+1}(\varphi) (1 - \frac{\Delta A_{t+1}}{\mathbb{E}[(\Delta M_{t+1})^2 \mid \mathcal{F}_t]} \Delta M_{t+1}) \mid \mathcal{F}_t] \\ &= \mathbb{E}[V_{t+1}(\varphi) \frac{Z_{t+1}}{Z_t} \mid \mathcal{F}_t]\end{aligned}$$

so that for locally risk minimising strategy φ , the product $ZV(\varphi)$ is a martingale, if the process Z is defined by the difference equation

$$Z_{t+1} - Z_t = Z_t \left(\frac{Z_{t+1}}{Z_t} - 1 \right) = -Z_t \lambda_{t+1} \Delta M_{t+1}$$

and $Z_0 = 0$ with predictable process

$$\lambda_{t+1} = \frac{\Delta A_{t+1}}{\mathbb{E}[(\Delta M_{t+1})^2 \mid \mathcal{F}_t]} = \frac{\mathbb{E}[\Delta X_{t+1} \mid \mathcal{F}_t]}{\text{Var}[\Delta X_{t+1} \mid \mathcal{F}_t]}$$

Thus if we start from $V_T = \eta_T = H$, for $t = T-1, \dots, 0$, we choose θ_t, η_t recursively (see [5]) to minimise

$$\mathbb{E}[(V_{t+1} - V_t - \theta_t(X_{t+1} - X_t))^2 \mid \mathcal{F}_t] = \mathbb{E}[(X_{t+1}(\theta_{t+1} - \theta_t) + (\eta_{t+1} - \eta_t)^2) \mid \mathcal{F}_t]$$

The hedging strategy constructed in this way is given explicitly by

$$\begin{aligned}\theta_T &= 0 \\ \eta_T &= H \\ \theta_t &= \frac{Cov[V_{t+1}(\varphi), \Delta X_{t+1} \mid \mathcal{F}_t]}{Var[\Delta X_{t+1} \mid \mathcal{F}_t]} \\ \eta_t &= \mathbb{E}[(\theta_{t+1} - \theta_t)X_{t+1} + \eta_{t+1} \mid \mathcal{F}_t]\end{aligned}$$

3.3.2 Mean Variance Hedging

The local risk minimisation strategies are typically not self-financing. A related criterion that still leads to self-financing strategies is the mean variance hedging. The difference between local risk minimisation and mean variance hedging is that we no longer impose on our trading strategies the replication requirement $V_T(\varphi) = H$ P -a.s. but insist instead on the self-financing constraint (3.4).

With mean variance hedging approaches the main idea is essentially to “approximate” the claim H as closely as possible by the terminal value of a self-financing strategy using a quadratic criterion (see [19]). More precisely, this amounts to finding a self-financing strategy φ which minimises

$$\mathbb{E}[(H - V_T)^2] = \| (H - V_T(\varphi)) \|_{L_2(P)}^2 \quad (3.14)$$

over all self-financing strategies φ , that is a strategy which approximates H in the L_2 -sense. The use of L_2 -norm is mainly for convenience because it allows fairly explicit results and at the same time lead to interesting mathematical questions. By (3.4), this strategy is completely determined by the pair $(V_0(\varphi), \theta)$, so that the solution to the problem of minimising (3.14) is obtained in principle by projecting the random variable H in $L_2(P)$ on the subspace $\mathcal{A} = \mathbb{R} + G_T(\Theta)$ of the attainable claims. Thus for a chosen linear subspace Θ which is self-financing, the linear subspace

$$\mathcal{G} = G_T(\Theta) = \left\{ \int_0^T \theta_s dX_s \mid \theta \in \Theta \right\}$$

of $L_2(P)$ describes all outcomes of self-financing Θ -strategies with initial wealth $V_0 = 0$ and

$$\mathcal{A} := \mathbb{R} + \mathcal{G} = \left\{ V_0 + \int_0^T \theta dX_s \mid (V_0, \theta) \in \mathbb{R} \times \Theta \right\}$$

is the subspace of contingent claim by self-financing Θ -strategies.

The optimal initial capital $V_0(\varphi)$ is often called the approximation price of H , and the optimal strategy is the mean variance hedging strategy. For the case where $X = (X_t)_{t \geq 0}$ is a real-valued square integrable process in discrete time with bounded mean variance tradeoff, explicit recursive formulae for $\tilde{\theta}$ have been given in Schweizer (1995b), see [15] and the references therein.

We now provide a brief outline of the results developed by Schweizer (1995), which also include those of Schal (1994) as a particular case. See [10] and the references therein.

It is assumed that $\text{Var}[\Delta X_t \mid \mathcal{F}_{t-1}] > 0$ for each $t = 1, 2, \dots$ and the discounted stock price has a bounded mean-variance tradeoff, that is

$$\frac{(\mathbb{E}[\Delta X_t \mid \mathcal{F}_{t-1}])^2}{\text{Var}[\Delta X_t \mid \mathcal{F}_{t-1}]}$$

is P a.s uniformly bounded.

We consider the following problem

$$\text{minimize } \mathbb{E}\left\{ \left(H - V_0 - \sum_{t=1}^T \theta_t \Delta X_t \right)^2 \right\} \quad (3.15)$$

for all $\theta \in \Theta$ where H is the fixed contingent claim and V_0 is the initial capital to start the strategy.

To describe the solution of the problem, we have to define the predictable process $\beta = \{\beta_t : t = 1, 2, \dots\}$ by

$$\beta_t = \frac{\mathbb{E}[\Delta X_t \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j) \mid \mathcal{F}_{t-1}]}{\mathbb{E}[\Delta X_t^2 \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j)^2 \mid \mathcal{F}_{t-1}]} \quad t = 1, 2, \dots$$

and the predictable process $\rho = \{\rho_t : t = 1, 2, \dots\}$ by

$$\rho_t = \frac{\mathbb{E}[H \Delta X_t \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j) \mid \mathcal{F}_{t-1}]}{\mathbb{E}[\Delta X_t^2 \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j)^2 \mid \mathcal{F}_{t-1}]} \quad t = 1, 2, \dots$$

Consider the random variable \tilde{Z}_0 and the process \hat{Z} respectively defined by

$$\tilde{Z}_0 := \prod_{t=1}^T (1 - \beta_t \Delta X_t)$$

$$\hat{Z}_t := \prod_{j=1}^T \frac{1 - \alpha_j \Delta X_j}{1 - \alpha_j \mathbb{E}[\Delta X_j | \mathcal{F}_{t-1}]} \quad t = 0, 1, \dots$$

where

$$\alpha_t = \frac{\mathbb{E}[\Delta X_t | \mathcal{F}_{t-1}]}{\mathbb{E}[\Delta X_t^2 | \mathcal{F}_{t-1}]} \quad t = 1, 2, \dots$$

We now define the measures \hat{P} and \tilde{P} on (Ω, \mathcal{F}) as follows

$$\frac{d\hat{P}}{dP} = \hat{Z}_T = \prod_{j=1}^T \frac{1 - \alpha_j \Delta X_j}{1 - \alpha_j \mathbb{E}[\Delta X_j | \mathcal{F}_{t-1}]}$$

$$\frac{d\tilde{P}}{dP} = \frac{\tilde{Z}_0}{\mathbb{E}[\tilde{Z}_0]}$$

where \hat{P} is called variance-optimal martingale measure and \tilde{P} is the discrete-time version of the minimal martingale measure. Note that the random variable may take negative values since there may exist a time $j \in \{1, 2, \dots\}$ such that $P(\alpha_j \Delta X_j > 1) > 0$ or $P(\beta_j \Delta X_j > 1) > 0$. Therefore the measures \hat{P} and \tilde{P} are signed martingale measures and not probability measures in general. A signed measure Q is said to be a signed martingale measure for X if $Q(\Omega) = 1$, Q is absolutely continuous with respect to P with

$$\frac{dQ}{dP} \in L_2(P)$$

and

$$\mathbb{E}\left[\frac{dQ}{dP} \Delta X_t | \mathcal{F}_{t-1}\right] = 0 \quad P - a.s \quad t = 1, 2, \dots$$

Proposition 3.1 *For every fixed contingent claim H , the solution of the problem (3.15) is given by the pair (V_0, ξ^{V_0}) where $V_0 = \mathbb{E}[H]$*

and

Proposition 5 For every fixed contingent claim H , and an initial outlay V_0 , the solution of problem (3.15) is given by

$$\theta_t = \theta_t^{V_0} = \theta_t^L + \beta_t(\hat{\mathbb{E}}[H | \mathcal{F}_{t-1}] - V_0 - \sum_{j=1}^{t-1} \theta_j^{V_0} \Delta X_j) + \gamma_t \quad t = 1, 2, \dots$$

where

$$\gamma_t = \frac{\mathbb{E}[(L_T^H - L_{t-1}^H) \Delta X_t \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j) | \mathcal{F}_{t-1}]}{\mathbb{E}[\Delta X_t^2 \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j)^2 | \mathcal{F}_{t-1}]}, \quad t = 1, 2, \dots$$

and the conditional expectation $\hat{\mathbb{E}}$, is defined by

$$\hat{\mathbb{E}}[H | \mathcal{F}_t] = \frac{\mathbb{E}[\hat{Z}_T H | \mathcal{F}_t]}{\mathbb{E}[\hat{Z}_T | \mathcal{F}_t]}$$

which is consistent with the properties we have if \hat{P} is a probability measure.

If the mean-variance tradeoff is deterministic, it can be shown that $\beta_t = \alpha_t$ for each $t = 1, 2, \dots$ that $\gamma_t = 0$ for each $t = 1, 2, \dots$ and that the measures \hat{P} and \tilde{P} coincide.

3.3.3 Minimal Variance Hedging

For this section, we consider the work in [1], [3].

Suppose we have a financial market consisting of $d + 1$ trading assets where the the bond and stocks prices X_0, X_1, \dots, X_d are modelled as follows

Bond price

$$dX_0(t) = r(t, \omega)dt \quad X_0(0) = x_0$$

Stock prices

$$dX_i(t) = b_i(t, \omega)dt + \sum_{j=1}^m \sigma_{ij}(t, \omega)d\eta_j(t) \quad X_i(0) = x_i$$

for $i = 1, \dots, d$ and $t \geq 0$.

where η_j are Brownian motions on a probability space (Ω, \mathcal{F}, P) . A random

variable $H \in L_2(\Omega, \mathcal{F}_t, P)$ is called hedgeable in the market if there exists an (predictable) adapted (see chapter 1, definition 1.1.4) process $\varphi(t) = (\varphi_1(t), \varphi_2(t), \dots, \varphi_d(t))$, $0 \leq t \leq T$ (a portfolio) such that φ is admissible, that is

$$\sum_{j=1}^d \mathbb{E} \left[\int_0^T \varphi_j^2(s) ds \right] < \infty$$

The restriction reflects a natural condition in real life finance: There must be a limit to how debt the creditors can tolerate.

Suppose we trade in some, say k , $k < d$, of the securities X_0, \dots, X_d . Let \mathcal{K} be the set of $i \in \{1, \dots, d\}$ such that we trade in X_i . Then, according to the model, the value hedged by an initial value $V_0 \in \mathbb{R}$ and an admissible portfolio $\varphi(t)$ up to time t is

$$V(t) = V^\varphi(t) = V_0 + \sum_{i \in \mathcal{K}} \sum_{\Delta} \varphi_i(s) \Delta X_i(s)$$

for $t \geq 0$ and $\Delta \subseteq (0, T]$.

Now, the portfolio φ replicates (hedges) H in the sense that

$$H = \mathbb{E}[H] + \sum_{j=1}^d \sum_{\Delta} \varphi_j(s) \Delta \eta_j(s) \quad P \quad a.e$$

where \mathbb{E} is the expectation with respect to P . Thus in this case, $\varphi(t)$,

$0 \leq t \leq T$ is called the hedging portfolio. Note that the market is called complete if every $H \in L_2(\Omega, \mathcal{F}, P)$ is replicable (hedgeable). Denote the set of all admissible portfolio by \mathcal{A} .

It is natural to ask: given a market as described earlier on and given

$H \in L_2(\Omega, \mathcal{F}, P)$, how close can we get to H by hedging with an admissible portfolio φ ? If we interpret closeness in terms of variance, then the problem is to find V_0 and $\varphi(t) \in \mathcal{A}$ such that

$$\mathbb{E}[(H - V^\varphi(T))^2] = \inf_{V_0, \varphi} \mathbb{E}[(H - V^\varphi(T))^2] \quad (3.16)$$

This question is often referred to the problem of “minimal variance hedging” for incomplete markets in mathematical finance.

This is a difficult problem even in the classical Brownian motion setting. The

problem has been studied in the context of martingales, see [8].

We are going to answer problem (3.16) by means of the non-anticipating stochastic derivative in the framework of Itô stochastic calculus

Definition 3.3.4 *For any random variable $H \in L_2(\Omega)$ we can define the “non-anticipating stochastic derivative”*

$$\varphi = \mathcal{D}H = \mathcal{D}_s H \quad 0 \leq s \leq T$$

of H with respect to the integrator η_t , $t \geq 0$ as the element in $L_2(\Omega \times [0, T])$ given by the limit

$$\mathcal{D}H = \lim_{n \rightarrow 0} \varphi^n \quad \text{in } L_2(\Omega \times [0, T])$$

with

$$\varphi^n(s) = \sum_{\Delta} \mathbb{E}[H \cdot \frac{\Delta \eta_t}{\Delta t} \mid \mathcal{F}_t] 1_{\Delta}(s) \quad 0 \leq s \leq T$$

where $\Delta \eta_t = \eta_{t+\Delta t} - \eta_t$ $\Delta = (t, t + \Delta t]$

Theorem 3.1 *For any \mathcal{F}_T -measurable random variable $\xi \in L_2(\Omega)$ the integrand $\varphi = \varphi_s$, $0 \leq s \leq T$, appearing in the Itô representation (see chapter 2, theorem 2.5) can be determined by the stochastic non-anticipating derivative*

$$\varphi_s = \mathcal{D}_s H \quad 0 \leq s \leq T$$

Proof

We will give the proof in steps, see [7].

step 1

Consider the process

$$H_t = \mathbb{E}[H \mid \mathcal{F}_t] = \mathbb{E}[H] + \sum_{\Delta} \varphi_s \Delta \eta_s \quad 0 \leq t \leq T$$

which is a martingale with respect to \mathcal{F}_t . Then we can see that

$$\mathbb{E}[H \Delta \eta_t \mid \mathcal{F}_t] = \mathbb{E}[\Delta H_t \Delta \eta_t \mid \mathcal{F}_t]$$

for $\Delta H_t = H_{t+\Delta t} - H_t$. In fact we have

$$\begin{aligned}\mathbb{E}[H\Delta\eta_t \mid \mathcal{F}_t] &= \mathbb{E}[\mathbb{E}(H - H_t \mid \mathcal{F}_t)\Delta\eta_t \mid \mathcal{F}_t] + \mathbb{E}[\Delta H_t\Delta\eta_t \mid \mathcal{F}_t] + H_t\mathbb{E}[\Delta\eta_t \mid \mathcal{F}_t] \\ &= \mathbb{E}[H - H_t]\mathbb{E}[\Delta\eta_t] + \mathbb{E}[\Delta H_t\Delta\eta_t \mid \mathcal{F}_t] + H_t\mathbb{E}[\Delta\eta_t] \\ &= \mathbb{E}[\Delta H_t\Delta\eta_t \mid \mathcal{F}_t]\end{aligned}$$

Hence the simple function can be rewritten in the equivalent form

$$\varphi^n(s) = \sum_{\Delta} \mathbb{E}[\Delta H_t \cdot \frac{\Delta\eta_t}{\Delta t} \mid \mathcal{F}_t] 1_{\Delta}(s) \quad 0 \leq s \leq T$$

step 2

Being the integral $\int_0^T \varphi_s d\eta_s$ an Itô integral, then by construction there exists a sequence of simple integrands $\psi^n, n = 1, 2, \dots$ with

$$\psi_s^n = e_0^n 1_{\{0\}}(s) + e_t^n 1_{\Delta}(s), \quad 0 \leq s \leq T$$

such that

$$\varphi = \lim_{n \rightarrow \infty} \psi^n \quad \text{in } L_2(\Omega \times [0, T]) \quad \text{that is} \quad \|\varphi - \psi^n\|_{L_2} \rightarrow 0, \quad n \rightarrow \infty$$

and thus

$$\sum_{\Delta} \varphi(s)\Delta\eta(s) = \lim_{h \rightarrow \infty} \sum_{\Delta} \psi^n(s)\Delta\eta(s) \quad \text{in } L_2(\Omega)$$

step 3

We note that in the construction of the Itô integral there is no statement of uniqueness of the sequence of simple integrands. We are going to exploit this fact and prove that the sequence $\varphi^n, n = 1, 2, \dots$ of simple integrands also characterises the representation, that is

$$\sum_{\Delta} \varphi(s)\Delta\eta(s) = \lim_{n \rightarrow \infty} \sum_{\Delta} \varphi_s^n \Delta\eta(s) \quad \text{in } L_2(\Omega)$$

To this aim it is enough to prove that

$$\varphi = \lim_{n \rightarrow \infty} \varphi^n \quad \text{in } L_2(\Omega \times [0, T]) \quad \text{that is} \quad \|\varphi - \varphi^n\|_{L_2} \rightarrow 0, \quad n \rightarrow \infty$$

step 4

Furthermore, since

$$\| \varphi - \varphi^n \|_{L_2} \leq \| \varphi - \psi^n \|_{L_2} + \| \psi^n - \varphi^n \|_{L_2} \quad (3.17)$$

we only need to show that

$$\| \psi^n - \varphi^n \|_{L_2} \longrightarrow 0, \quad n \longrightarrow \infty$$

step 5

$$\begin{aligned} \| \psi^n - \varphi^n \|_{L_2} &= \mathbb{E} \left[\sum_{\Delta} \sum_{\Delta} \frac{1}{\Delta t} (e^n(t) \Delta t - \mathbb{E}[\Delta H(t) \Delta \eta(t) \mid \mathcal{F}_t]^2 1_{\Delta}(s) \Delta s) \right] \\ &= \sum_{\Delta} \frac{1}{\Delta t} \mathbb{E} \left[(\mathbb{E}[e^n(t) \Delta t - \sum_{\Delta} \varphi(s) 1_{\Delta}(s) \Delta s \mid \mathcal{F}_t]^2) \right] \\ &\leq \sum_{\Delta} \frac{1}{\Delta t} \mathbb{E} \left[\left(\sum_{\Delta} (e^n(t) \Delta t - \varphi(s)) 1_{\Delta}(s) \Delta s \right)^2 \right] \\ &\leq \sum_{\Delta} \frac{1}{\Delta t} \mathbb{E} \left[\sum_{\Delta} (\psi^n(s) - \varphi(s))^2 1_{\Delta}(s) \Delta s \cdot \sum_{\Delta} 1_{\Delta}(s) \Delta s \right] \\ &= \| \psi^n(s) - \varphi(s) \|_{L_2} \\ &\longrightarrow 0, \quad n \longrightarrow \infty \end{aligned}$$

step 7

From (3.17), we have

$$\| \varphi - \varphi^n \|_{L_2}^2 \leq 2 \| \varphi - \psi^n \|_{L_2}^2 \longrightarrow 0 \quad n \longrightarrow \infty$$

This completes the proof.

Corollary 4 *The non-anticipating stochastic derivative $\mathcal{D}\xi$ is continuous with respect to ξ in $L_2(\Omega)$. Namely*

$$\xi = \lim_{n \rightarrow \infty} \xi_n \quad \text{that is} \quad \| \xi - \xi_n \| \longrightarrow 0, \quad n \longrightarrow \infty$$

implies

$$\mathcal{D}\xi = \lim_{n \rightarrow \infty} \mathcal{D}\xi_n \quad \text{that is} \quad \| \mathcal{D}\xi - \mathcal{D}\xi_n \|_{L_2} \longrightarrow 0 \quad \text{as } n \longrightarrow \infty$$

Note that the non-anticipating stochastic derivative $\mathcal{D}_\eta H = \varphi$ represents the minimal variance portfolio and the simple function in definition 3.3.4 gives a method of approximation.

We now consider the martingale η_t in a special nature

$$\eta_j(t) = \sigma_j W_j(t) + \sum_{j=1}^d \sum_{\Delta} x \tilde{N}_j(\Delta s, \Delta x)$$

where $\Delta s = (s, s + \Delta s]$, $\Delta x = (x, x + \Delta x]$. Note that $\eta_j(t)$ is also a levy process. $W_j(t)$ is the standard Wiener process (with diffusion coefficient 1), $\sigma_j > 0$ is a constant and $\tilde{N}_j(\Delta s, \Delta x)$ is the compensated Poisson random measure, that is

$$\tilde{N}_j(\Delta t, \Delta x) = N_j(\Delta t, \Delta x) - \nu_j(\Delta x) \Delta t$$

with $\nu_j(\Delta x)$ as the jump measure of the levy process $\eta_j(t)$ itself and $N_j(\Delta t, \Delta x)$ is a Poisson random measure such that

$$\mathbb{E}[N_j(\Delta t, \Delta x)] = \nu_j(\Delta x) \Delta t$$

Now, if $\eta = (\eta_1, \dots, \eta_d) = (W_1, \dots, W_d)$ is the Wiener process with independent components, the representation in Theorem 3.1 for an appropriate H reduces to

$$H = H_0 + \sum_{j=1}^d \sum_{\Delta} \mathbb{E}[D_{s,j} H \mid \mathcal{F}_s] \Delta W_j(s)$$

Thus in this case we have

$$\mathcal{D}_{\eta_j} H(s) = \mathbb{E}[D_{s,j} H \mid \mathcal{F}_s] \quad 0 \leq s \leq T$$

where $D_{s,j} H$, $0 \leq s \leq T$ is the stochastic Malliavin derivative in the direction of the Wiener process η_j .

If $k=d$, this gives the well known Clark-Haussmann-Ocone formula.

We now give an explicit formula for minimal variance portfolio (see [1]).

Note that

$$\begin{aligned} \mathbb{E}[(H - V_T(\varphi))^2] &= \mathbb{E}[(H - V_0 - \sum_{\Delta} \varphi(t) \Delta \eta_t)^2] \\ &= \mathbb{E}[(H - \mathbb{E}[H] - \sum_{\Delta} \varphi(t) \Delta \eta(t) + (\mathbb{E}[H] - V_0))^2] \\ &= \mathbb{E}[(H - \mathbb{E}[H] - \sum_{\Delta} \varphi(t) \Delta \eta(t))^2 + (\mathbb{E}[H] - V_0)^2] \end{aligned}$$

We see that it is optimal to choose $V_0 = \mathbb{E}[H]$

Theorem 3.2 *For any $H \in D_{1,2}$, the minimal variance portfolio $\varphi = (\varphi_1, \dots, \varphi_d)$ in (3.16):*

$$H = \mathbb{E}[H] + \sum_{j=1}^d \sum_{\Delta} \varphi_j(s) \Delta \eta_j(s)$$

admits the following representation

$$\varphi_j(s) = \frac{\sigma \mathbb{E}[D_{s,j}H \mid \mathcal{F}_s] + \sum_{\Delta} x \mathbb{E}[D_{s,x,j}H \mid \mathcal{F}_s] \nu_j(\Delta x)}{\sigma_j^2 + \sum_{\Delta} x^2 \nu_j(\Delta x)}$$

where $D_{s,j}H$ is the derivative of H with respect to the Wiener process W_j , $D_{s,x,j}H$ is the derivative with respect to the j^{th} compensated Poisson random measure with $s \geq 0$ and $x \in \mathbb{R}$ here and in the sequel.

Note that the summation is referred to all intervals $\Delta = (x, x + \Delta x]$. Here we are partitioning the real line (representing the size of the jumps of the Poisson random measure).

Chapter 4

Applications

In this chapter, we will consider a few computations.

Example 4.1 Consider the price process $X_t = (1, \eta_t)$ where η_t is a 1-dimensional Brownian motion starting at zero.

Is the claim $H(\omega) = \eta_T(\omega)$ attainable?

Solution

We look for a constant V_0 and a self-financing portfolio $\theta_t = (\theta_t^{(1)}, \theta_t^{(2)})$ such that

$$\eta_T(\omega) = V_0 + \sum_{s=0}^T \theta(s) \Delta X_s$$

Note that

$$X_t = (1, \eta_t)$$

$$dX_t = (0, d\eta_t)$$

Therefore

$$\eta_T(\omega) = V_0 + \sum_{s=0}^T \theta_t^{(1)} \cdot 0 + \sum_{s=0}^T \theta_t^{(2)} \Delta \eta_s \quad (4.1)$$

Taking expectation of both sides

$$0 = \mathbb{E}[\eta_T(\omega)] = V_0 + \mathbb{E}\left[\sum_{s=0}^T \theta_t^{(2)} (\eta_t - \eta_{t-1})\right] = V_0$$

hence

$$V_0 = 0 \tag{4.2}$$

we see immediately from (4.1) that

$$\theta_t^{(2)} = 1 \tag{4.3}$$

does the job.

We can now determine $\theta_t^{(1)}$ from (4.3) and the identity

$$\theta_t^{(1)} \cdot 1 + \theta_t^{(2)} \eta_t = V_t = V_0 + \sum_{s=0}^t \theta_s^{(2)} \eta_s = \eta_t$$

This implies that $\theta_t^{(1)} = 0$. Therefore $V_0 = 0$, $\theta_t^{(1)} = 0$ and $\theta_t^{(2)} = 1$

The claim $H(\omega) = \eta_T(\omega)$ can be hedged using $V_0 = 0$ and $\theta_t = (0, 1)$

Example 4.2 Suppose now $X_t = (1, Y_t)$ where

$$dY_t = d\eta^{(1)}(t) - 5d\eta^{(2)}(t)$$

with $(\eta_t^{(1)}, \eta_t^{(2)})$ being a 2-dimensional Brownian motion .

Find a T -claim, $H(\omega)$, $E[H^2(\omega)] < \infty$ which is not attainable in the market.

Solution

We look for $H(\omega)$ with $E[H^2(\omega)] < \infty$ such that is impossible to get either V_0 or $\theta_t = (\theta_t^{(1)}, \theta_t^{(2)})$ such that

$$H(\omega) = V_0 + \sum_{s=0}^T \theta_s^{(1)} \Delta X^{(1)}(s) + \sum_{s=0}^T \theta_s^{(2)} \Delta X^{(2)}(s) \tag{4.4}$$

by Itô representation (see chapter 2, Theorem 2.5) there exists a unique

$V_0 = \mathbb{E}[H(\omega)]$ and a unique $\vartheta_t = (\vartheta_t^{(1)}, \vartheta_t^{(2)})$ such that

$$H(\omega) = V_0 + \sum_{s=0}^T \vartheta_s^{(1)} \Delta X^{(1)}(s) + \sum_{s=0}^T \vartheta_s^{(2)} \Delta X^{(2)}(s) \tag{4.5}$$

From (4.4), we have

$$H(\omega) = V_0 + \sum_{s=0}^T \theta_s^{(2)} \Delta Y_s$$

Thus

$$\begin{aligned} H(\omega) &= V_0 + \sum_{s=0}^T \theta_s^{(2)} [\Delta\eta_s^{(1)} - 5\Delta\eta_s^{(2)}] \\ &= V_0 + \sum_{s=0}^T \theta_s^{(2)} \Delta\eta_s^{(1)} - \sum_{s=0}^T 5\theta_s^{(2)} \Delta\eta_s^{(2)} \end{aligned}$$

Suppose we choose

$$H(\omega) = \eta_T^{(1)} \quad , \quad V_0 = E[H] = 0$$

Therefore from (4.5) we get

$$H(\omega) = \eta_T^{(1)} = \sum_{s=0}^T \vartheta_s^{(1)} + \sum_{s=0}^T \vartheta_s^{(2)} \Delta\eta_s^{(2)}$$

One solution for this equation is $\vartheta = (1, 0)$ (see Example 4.1). That is

$$V_0 = 0 \quad , \quad \vartheta_t^{(1)} = 1 \quad , \quad \vartheta_t^{(2)} = 0$$

which is a contradiction

Therefore, for $H(\omega) = \eta_T^{(1)}$ we cannot find $\theta_t = (\theta_t^{(1)}, \theta_t^{(2)})$ such that

$$H(\omega) = V_0 + \sum_{s=0}^T \theta_s \Delta X_s$$

Example Suppose the market is modelled by S_0 and d independent Brownian motions $S = \eta = \eta_1, \dots, \eta_d$ and we are allowed to trade only in S_0 and η_1, \dots, η_k where $k < d$ securities. Note that the market is incomplete because of the given constraint.

For any $H \in D_{1,2}$ the minimal variance hedge \hat{H} in chapter 3, Theorem 3.2 is

$$\varphi_j(s) = \mathcal{D}_{\eta_j} H(s) = \mathbb{E}[D_{s,j} H \mid \mathcal{F}_s] \quad j = 1, \dots, d$$

Since

$$H = \mathbb{E}[H] + \sum_{j=1}^d \sum_{\Delta} \mathbb{E}[D_{s,j} H \mid \mathcal{F}_s] \Delta\eta_j(s)$$

and also

$$H = H^0 + \hat{H} = H^0 + \sum_{j=1}^d \sum_{\Delta} \varphi_j(s) \Delta\eta_j(s)$$

then the minimal risk can be given explicitly by

$$H^0 = \sum_{j=1}^d \sum_{\Delta} \varphi_j(s) \Delta \eta_j(s)$$

Theorem 4.1 *Let $g : \mathbb{R}^d \rightarrow \mathbb{R}$ be a bounded continuous function and let $0 \leq s_1 \leq \dots \leq s_d \leq T$. Then*

$$D_{t,x}[g(\eta(s_1), \dots, \eta(s_d))] = g(\eta(s_1) + x1_{[0,s_1]}(t), \dots, \eta(s_d) + x1_{[0,s_d]}(t)) - g(\eta(s_1), \dots, \eta(s_d))$$

where η_t is a martingale in a special nature (see chapter 3, section 3.3.3).

Examples

1. $D_{t,x}\eta(T) = \eta(T) + x1_{[0,T]}(t) - \eta(T) = x$
2. $D_{t,x}\eta^2(T) = (\eta(T) + x1_{[0,T]}(t))^2 - \eta^2(T) = 2x\eta(T) + x^2$
3. $D_{t,x}[(\eta(T) - c)^+] = (\eta(T) + x - c)^+ - (\eta(T) - c)^+$

Example 4.3 *Find the closest hedge of the claim $H(\omega) = \eta^2(T)$.*

Solution Take $d = 1$

If $H(\omega) = \eta^2(T)$ then thanks to the Theorem 4.1, we have

$$\begin{aligned} D_{t,x}H &= (\eta(T) + x)^2 - (\eta(T))^2 \\ &= x(2\eta(T) + x), \quad 0 \leq s \leq T \end{aligned}$$

here, since $\eta(t)$ is a martingale, the minimal variance portfolio is given by

$$\begin{aligned} \varphi(t) &= \frac{\sum_{\Delta} x \mathbb{E}[x(2\eta(T) + x) \mid \mathcal{F}_t] \nu(\Delta x)}{\sum_{\Delta} x^2 \nu(\Delta x)} \\ &= \frac{\sum_{\Delta} x^2 2\eta(t) \nu(\Delta x) + \sum_{\Delta} x^3 \nu(\Delta x)}{\sum_{\Delta} x^2 \nu(\Delta x)} \\ &= 2\eta(t) + \frac{\sum_{\Delta} x^3 \nu(\Delta x)}{\sum_{\Delta} x^2 \nu(\Delta x)} \end{aligned}$$

Chapter 5

Markov processes

5.1 Markov Models

In this chapter, we will focus on a class of stochastic processes that share what is called the “Markov property”: the future is independent of the past, given the present values of the process. See [4], [17], [18], [21].

The Markov processes are important models of security prices because they are often realistic representations of the true prices and also they lead to simplified computations. In the previous chapters we have considered different quadratic hedging methods. We would like to make use of the Markov property to generalize such results in the case of a discrete time market. We therefore begin by considering the filtration $\mathbb{F} = \{\mathcal{F}_t; t = 0, 1, \dots\}$ generated by a stochastic process $X = \{X_t; t = 0, 1, \dots\}$. This process takes values in some finite set E , called the state space. We will also consider a sample space Ω and a probability measure P on it and the information \mathcal{F} should be thought of as the history of the present and past values of the process X .

Definition 5.1.1 *The stochastic process X is called a "Markov chain" if*

$$P[X_{t+1} = j \mid \mathcal{F}_t] = P[X_{t+1} = j \mid X_t]$$

for all $j \in E$ and all t . By elementary probability calculations, it follows that

$$P[X_{t+s} = j \mid \mathcal{F}_t] = P[X_{t+s} = j \mid X_t]$$

for $s \geq 1$.

One thinks of this as “the best prediction of the future given the past and the present is the same as the best prediction of the future given the present”

Thus a Markov chain is a stochastic process where the only information useful for predicting the future values is the current state. Thus, in other words, given the history of the process, the past values can be ignored as long as you know the present state.

The Markov chain X is said to be “stationary or time-homogeneous” if the conditional probabilities

$$P(X_{t+1} = j \mid \mathcal{F}_t)$$

do not depend on time t . We can define the transition probabilities

$$P(i, j) \equiv P(X_{t+1} = j \mid X_t = i)$$

for $i, j \in E$. and to organise them in a transition matrix

$$P' \equiv [P(i, j)]$$

Observe that this is a square matrix with the number of rows equal to the number of elements in the state space. We also observe that the sum of elements in each row of P' is equal to one.

A strong requirement that is often satisfied is that the Markov property holds for stopping times. A useful property of Markov chain is

If $Y = f(X_t, X_{t+1}, \dots)$ for some function f , then

$$\mathbb{E}[Y \mid \mathcal{F}] = \mathbb{E}[Y \mid X_t]$$

Interpretation

$\mathbb{E}[Y \mid \mathcal{F}_t]$ is the conditional expectation from all information of X_t .

$\mathbb{E}[Y \mid X_t]$ only depends on the final value X_t .

Remark

This means that the future and the past are conditionally independent, given the present values.

Now suppose we have a security market model where the discounted price process X (see chapter 3, section 3.2) is a Markov chain. That is

$$P[X_{t+1} = j \mid \mathcal{F}_t] = P[X_{t+1} = j \mid X_t]$$

for all $j \in E$ and t . It is natural to ask whether the discounted security prices are Markov chain under the martingale measure if there is one. This is not a trivial question because a stochastic process may lose its Markov property when you change from one probability measure to an equivalent one.

We can answer this in the affirmative:

“If there are no arbitrage opportunities, if the discounted price process X is a Markov chain under P , and if the filtration \mathcal{F} is the one generated by X , then there exist a martingale measure Q under which X is a Markov chain.”

Now, from the previous chapters, X was considered to be a stochastic process. Bouleau and Lamberton (1989), see [15] and references therein, imposed the additional condition that X is a function of some Markov process, that is, the future is independent of the past given the present value. Thus we can now be in a position to get explicit results in the case of a discrete time market. See [10], and [15].

Suppose that instead of insisting on the replicability condition $V_T(\varphi) = H$ P - a.s for the claim H via the strategy φ , we focus on self-financing risk minimization strategies (see chapter 3). We describe such a strategy by a pair (V_0, θ) in $L_2(\Omega)$

and its shortfall at the terminal date T is

$$H - V_T(V_0, \theta) = H - V_0 - \sum_{\Delta} \theta_u \Delta X_u$$

If the contingent claim H is attainable by such a strategy in the sense that $H = V_T(V_0, \theta)$ for some pair (V_0, θ) , the shortfall can therefore be reduced to zero. But in general, one has a residual risk of

$$J_0(V_0, \theta) := \mathbb{E}[(H - V_T(V_0, \theta))^2]$$

if one uses a quadratic loss function, and the idea of Bouleau and Lamberton (1989) is to minimize this residual risk by choice of (V_0, θ) , see [15]. This clearly amounts to projecting the random variable H in $L_2(P)$ on the linear space spanned by $L_2(\mathcal{F}_0, P)$ and the stochastic integrals $\int_0^T \theta_s dX_s$ with $\theta \in L^2(X)$.

Recalling (see chapter 3, section 3.2.3) the decomposition

$$H = \mathbb{E}[H | \mathcal{F}_0] + \sum_{\Delta} \theta_u^H \Delta X_u + L_T^H$$

the solutions are given by

$$\begin{aligned} \tilde{V}_0 &= [H | X_0] \\ \tilde{\theta} &= \theta^H \end{aligned}$$

with minimal residual risk of

$$\begin{aligned} J_o(\tilde{V}_0, \tilde{\theta}) &= \mathbb{E}[(L_T^H)^2] \\ &= \text{Var}[L_T^H] \end{aligned}$$

5.2 Markovian case: Local risk Minimisation

We now analyse in detail the implementation of the local risk-minimising problem when using Markov property to describe the stock price. Consider the filtration

$\{\mathcal{F}\}_{t \geq 0}$ given by $\mathcal{F}_t = \sigma(X_i \mid i \leq t)$, the σ -field generated by the variables X_0, \dots, X_t . Suppose the stock prices are modelled using a Markov process with N periods, but hedging can only take place on $T < N$ time dates

$0 = i_0 < i_1 < \dots < i_{T-1} < N := i_T$. Thus, for all $0 \leq t \leq T$, at time i_t , there are $d = i_t + 1$ possible states for the stock price. Recall (from chapter 3, section 3.1) that the discounted stock price is given by

$$X_t^i = \frac{S_t^i}{S_t^0}, \quad 0 \leq t \leq T, \quad 0 \leq i \leq d$$

Now suppose that at time N , the discounted payoff of the option in state i is given by H_i .

Then, in the framework of the Markov model for the stock price, the Local risk-minimisation problem (see chapter 3, section 3.3.1) becomes: starting from $V_T = H$, for all the states i at time t , $t = T - 1, \dots, 0$, minimise

$$\mathbb{E}[(C_{t+1}(\varphi) - C_t(\varphi))^2 \mid X_t = X_t^i]$$

The explicit hedging strategy solving this problem is given by the following:

For each $1 \leq i \leq d$ define

$$\xi_T^i = 0, \quad \eta_T^i = H_i$$

and for each $t = T - 1, \dots, 0$ and for each $1 \leq i \leq d$, define

$$\theta_t = \frac{\text{Cov}[H - \sum_{j=t+2}^T \theta_j \Delta X_j, \Delta X_{t+1} \mid X_t = X_t^i]}{\text{Var}[\Delta X_{t+1} \mid X_t = X_t^i]}$$

and

$$\eta_t = \mathbb{E}[(\theta_{t+1} - \theta_t)X_{t+1} + \eta_{t+1} \mid X_t = X_t^i]$$

5.3 Markovian case: Mean Variance Hedging

In this section we assume that the parameters of the price evolution only depends on the current values of X_t and some state variables Y_t and the joint

process (X, Y) is Markovian (see [11]).

More precisely

$$\frac{dX_t^i}{X_t^i} = b^i(t, X_t, Y_t)dt + \sum_{j=1}^n \sigma^{ij}(t, X_t, Y_t)d\eta_t$$

and

$$dY_t = \alpha(t, X_t, Y_t)dt + \gamma(t, X_t, Y_t)d\eta_t$$

where $i = 1, \dots, d$ and Y is a $n - d$ dimensional process.

We introduce the additional process Y in order to be able to simply describe the information available to the traders. The matrix

$$\begin{bmatrix} \sigma \\ \gamma \end{bmatrix}$$

is invertible for all $t, t \geq 0$ P -a.s. and \mathcal{F}_t is simply the information generated by current and past values of X, Y . Note that in the case of complete market $d = n$ and the introduction of Y is unnecessary.

We now relate the mean variance hedging discussed in chapter 3, section 3.3.2 to the case of Markov process. Suppose there are $d+1$ assets being traded. Consider the filtration $\mathcal{F}_t = \sigma(X_i \mid i \leq t)$ the σ -field generated by the variables X_0, \dots, X_t . Now in the case of Markov process, the problem reduces to finding a mean variance portfolio V_0 and $\varphi(t)$, such that

$$\mathbb{E}[(H - V_T)^2] = \| (H - V_T(\varphi)) \|_{L_2(P)}^2$$

The Hedging Price

Note that the cash-flow to be hedged only depends on the terminal values of X and Y :

$$H_T = H(X_T, Y_T)$$

From Chapter 3, section 3.3.2 the hedging price is given by

$$V_t(H_T) = E[H(X_T, Y_T)]$$

This follows from a Markov property of (X, Y) .

The Hedging Portfolio

The allocation in the risky asset of the hedging portfolio (see chapter 3, section 3.3.2) driven by Markov process is given by

$$\theta_t = \theta_t^{V_0} = \theta_t^L + \beta_t(\hat{\mathbb{E}}[H \mid X_{t-1} = X_{t-1}^i] - V_0 - \sum_{j=1}^{t-1} \theta_j^{V_0} \Delta X_j) + \gamma_t \quad t = 1, 2, \dots$$

where

$$\beta_t = \frac{\mathbb{E}[\Delta X_t \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j) \mid X_{t-1} = X_{t-1}^i]}{\mathbb{E}[\Delta X_t^2 \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j)^2 \mid X_{t-1} = X_{t-1}^i]} \quad t = 1, 2, \dots$$

and

$$\gamma_t = \frac{\mathbb{E}[(L_T^H - L_t^H) \Delta X_t \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j) \mid X_{t-1} = X_{t-1}^i]}{\mathbb{E}[\Delta X_t^2 \Pi_{j=t+1}^T (1 - \beta_j \Delta X_j)^2 \mid X_{t-1} = X_{t-1}^i]}$$

Observe that again the allocation in the risky assets depends only on the Markov process. Thus the hedging portfolio (θ_t, V_t) shares a Markov property where by the future value of the portfolio depends only on the current value.

5.4 Markovian case: Minimal variance hedging

We now relate the results obtained in chapter 3, section 3.3.3 for minimal variance hedging to the case of Markov process, where the future value does not depend on the past history given the present value.

Suppose there are $d+1$ assets, one of which is the bond and the rest of them are the stock holdings. We consider the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ given by $\mathcal{F}_t = \sigma(X_j \mid j \leq t)$, the σ -field generated by the variables X_0, \dots, X_t . Suppose that at time $t < T$, the discounted payoff of the option in state j is given by H_j . If the stock prices are modelled by the Markov process, then the problem (3.16) reduces to finding V_0 and $\varphi(t) \in \mathcal{A}$ such that

$$\mathbb{E}[(H - V^\varphi(T))^2] = \inf_{V_0, \varphi} \mathbb{E}[(H - V^\varphi(T))^2] \quad (5.1)$$

The explicit formula for minimal variance portfolio solving problem (5.1) is therefore given by

$$\varphi_j(t) = \frac{\sigma \mathbb{E}[D_{t,j}H \mid X_t = X_t^j] + \sum_{\Delta} x \mathbb{E}[D_{t,x,j}H \mid X_t = X_t^j] \nu_j(\Delta x)}{\sigma_j^2 + \sum_{\Delta} x^2 \nu_j(\Delta x)} \quad (5.2)$$

Example 5.1 Find the closest hedge φ of the European call option

$$H(\Omega) = (X_1(T) - c)^+$$

where $c > 0$ is a constant, $X_1(T) = \eta_1(T)$

Solution Choose $d = 1$. The closest hedge (see chapter 3, Theorem 3.2) is given by

$$\varphi(t) = \frac{\sum_{\Delta} x \mathbb{E}[D_{t,x}H \mid X_s] \nu(\Delta x)}{\sum_{\Delta} x^2 \nu(\Delta x)}$$

By theorem 4.1, we have

$$D_{t,x}[(\eta(T) - c)^+] = (\eta(T) + x - c)^+ - (\eta(T) - c)^+ \quad \text{for } t \leq T$$

and hence by the Markov property

$$\varphi(t) = \frac{\sum_{\Delta} x \mathbb{E}^{X_s}[(\eta(T-t) + x - c)^+ - (\eta(T-t) - c)^+]_{X_s=\eta(s)} \nu(\Delta x)}{\sum_{\Delta} x^2 \nu(\Delta x)}$$

Now the intuitive meaning of Markov property is that the future behaviour of the process $(X_t)_{t \geq 0}$ after t depends on the value X_t and is not influenced by the history of the process before t . This is an important property of the Markov property model and it will have great consequences in the pricing of options. For instance, it will allow us to show how the price of an option on an underlying asset whose price is Markovian only depends on the price of this underlying asset at time t .

Mathematically, an \mathcal{F}_t -adapted process $(X_t)_{t \geq 0}$ satisfies the Markov property if for any bounded Borel function f and for any s and t such that $s \leq t$, we have

$$\mathbb{E}[f(X_t) \mid \mathcal{F}_s] = \mathbb{E}[f(X_t) \mid X_s]$$

Example 5.2 Prove that $X_t = X_t^x = x \exp(bt + \sigma\eta(t))$ is a Markov process

Proof

We want to show that for f bounded, Borel measurable, then

$$\mathbb{E}[f(X_{t+h}) \mid \mathcal{F}_t] = \mathbb{E}^{X_t}[f(X_h)] \quad X(0) = x$$

Now

$$\begin{aligned} \mathbb{E}^x[f(X_{t+h}) \mid \mathcal{F}_t] &= \mathbb{E}[f(X_{t+h}^x) \mid \mathcal{F}_t] \\ &= \mathbb{E}[f(x \exp(b(t+h) + \sigma\eta(t+h))) \mid \mathcal{F}_t] \\ &= \mathbb{E}[f(x \exp(bt + bh + \sigma\eta(t+h) - \sigma\eta(t) + \sigma\eta(t))) \mid \mathcal{F}_t] \\ &= \mathbb{E}[f(x \exp(bt + \sigma\eta(t)) \cdot \exp(bh + \sigma(\eta(t+h) - \eta(t)))) \mid \mathcal{F}_t] \\ &= \mathbb{E}[f(x \exp(bt + \sigma\eta(t)) \cdot \exp(bh + \sigma\eta(h)))] \\ &= \mathbb{E}[f(y \exp(bh + \sigma\eta(h))) \mid_{y=x \exp(bt + \sigma\eta(t))=X_t}] \\ &= \mathbb{E}^{X_t}[f(x \exp(bh + \sigma\eta(h)))] \\ &= \mathbb{E}^{X_t}[f(X_h)] \end{aligned}$$

Therefore, the process $X_t = x \exp(bt + \sigma\eta(t))$ is a Markov process.

Chapter 6

Conclusion

In the study of the techniques in stochastic analysis, we have introduced the Itô integral with respect to a martingale which by the way is a martingale. We then apply integration directly for continuous martingales. This had guaranteed us of the smooth construction of Itô integrals. We have also shown how a stochastic differential (the non-anticipative derivative) is determined for a given stochastic function.

In the review of the methods in quadratic hedging, our analysis of local risk-minimisation shows that it is a hedging approach designed to control riskness of a strategy as measured by its local cost fluctuations. For non-attainable claims, the mean variance hedging seeks a best approximation of the contingent claim H by the terminal value $V_0 + G_T(\Theta)$. Thus, to find a mean variance optimal strategy one has to project H in $L_2(P)$ on the space $\mathbb{R} + G_T(\Theta)$ of attainable claims. We have also considered the minimal variance hedging method whereby we seek to look for “closest” achievable claim \hat{H} to the claim H . The hedging portfolios have explicitly been given.

We were also able to relate these results to the case of Markov processes. Thanks to the Markov property: The future values are independent of the past given the present value.

References

- [1] Benth F, E and e.t al 2002: Explicit Representation of the Minimal Variance Portfolio in Market driven by Levy process.
- [2] Bernt Ø : *Stochastic Differential Equations* ;fifth edition.
- [3] Biagini F, and Bernt O 2003: “Minimal Variance Hedging for Fractional Brownian motion.”
- [4] Cerny A : “Dynamic Programming and Mean-Variance Hedging in Discrete Time.”
- [5] Coleman T.F e.t al 2002: Discrete Hedging under Piecewise Linear Risk Minimization.
- [6] Da Prato G. and Zabczyk J: *Stochastic Equations in Infinite Dimensions*.
- [7] Di Nunno, G, MAT 4700/4710: Short note on the Stochastic Non-anticipating Derivative.
- [8] Di Nunno ,G, 2001 : Stochastic Integral Representations, Stochastic Derivatives and Minimal Variance Hedging.
- [9] Di Nunno, G, and Yu.A. Rozanov 1999 : “On Stochastic integration and Differentiation”.
- [10] Fabio M : Claim Pricing and Hedging under Market Imperfection.

- [11] Gouriéroux C et al (1995): Quadratic hedging and numeraire.
- [12] Hans Föllmer and Martin Schweizer : Hedging of Contingent Claims under incomplete information.
- [13] Luciano Campi : Some Results on Quadratic Hedging with Insider Trading.
- [14] Kunita .H : “*Stochastic flows and Stochastic Differential Equations*.”
- [15] Martin Schweizer : In “Option Pricing, Interest Rates and Risk Management” Cambridge University Press 2001.
- [16] Mataramvura S: Lecture Notes on Stochastic Differential equations, 2003.
- [17] Protter P :*Stochastic Integration and Differential Equations; A New Approach*.
- [18] Stanley R Pliska :*Introduction to Mathematical Finance: Discrete Time Models*, Blackwell, Oxford, 1997.
- [19] Thomas Møller (2001): On valuation and risk management at the interface of insurance and finance.
- [20] Thomas Møller (2003): Risk Minimisation.
- [21] Yin G. and Zhou X.Y : Mean-Variance Portfolio Selection under Markov Regime; Discrete-time Models and Continuous-time limits.