STOCHASTIC PROCESSES FOR FINANCE RISK MANAGEMENT TOOLS

Notes for the Course by

F. Boshuizen, A.W. van der Vaart, H. van Zanten, K. Banachewicz and P. Zareba

CORRECTED 15 October 2006

CONTENTS

LITERATURE

[1] Baxter, M. and Rennie, A., (1996). *Financial calculus*. Cambridge University Press, Cambridge.

[2] Chung, K.L. and Williams, R.J., (1990). *Introduction to stochastic integration, second edition*. Birkhäuser, London.

[3] Etheridge, A., (2002). *A Course in Financial Calculus*. Cambridge University Press.

[4] Campbell, J.Y., Lo, A.W. and MacKinlay, A.C., (1997). *The Econometrics of Financial Markets*. Princeton University Press.

[5] Hunt, P.J. and Kennedy, J.E., (1998). *Financial Engineering*. Wiley.

[6] Jorion, P., (2001). *Value at Risk the New Benchmark for Managing Financial Risk*. McGraw-Hill, New York.

[7] Musiela, M. and Rutkowski, M., (1997). *Martingale Methods in Financial Modelling*. Springer-Verlag, Berlin.

[8] Smithson, C.W., Smith, C.W. and Wilford, D.S., (1995). *Managing Financial Risk*. Irwin, Burr Ridge, Illinois.

Baxter and Rennie is a book written for an audience of people in practice, but using the correct mathematical concepts and results. The result is a text with much intuition, and with mathematical theorems that are stated with some precision, but never proved or rigorously interpreted. Etheridge is a more mathematical version of Baxter and Rennie. Together these books are close to the content of the course. We recommend that you read these books, next to the following notes, which are really very brief.

The other books on the list are for further reading.

Musiela and Rutkowski is a step up in mathematical level. Hunt and Kennedy and Chung and Williams are mathematically completely rigorous. Chung and Williams has very little on finance. This is just a tiny selection of the books on mathematical finance that have appeared in the past ten years.

Campbell, Lo and MacKinLay gives a much wider view on finance, including some historical analysis and economic theories of price forming through utility, next to subjects close this course, all from a low-level mathematics point of view. The authors of “Managing Financial Risk” are bankers, not mathematicians. In their preface they write:

This stuff is not as hard as some people make it sound.

The financial markets have some complicated features, but good common sense goes a lot further than mathematical flash and dash.

Keep that mind when following this course. We do not entirely disagree, but do believe that some flash (and dash) will certainly help to clarify this complicated area.

1 Pricing derivatives

Financial instruments can be divided in two basic classes: underlying and derivative ones. The former can be stocks, bonds or various trade goods, while the latter are financial contracts that promise some payment to the owner, depending on the behavior of the "underlying" (e.g. a price at a given time T or an average price over a certain time period). Derivatives are extremely useful for risk management (apart from obvious investment properties) — we can reduce our financial vulnerability by fixing a price for a future transaction now.

In this chapter we introduce some basic concepts through examples; formal definitions and theory follow in later chapters.

1.1 Hedging a Forward

A *forward* is a contract that pays the owner an amount $S_T - K$ at a fixed time T in the future, the *expiry time*, where S_t is the price of an asset at time t and K is a fixed number, the *strike price*. Both T and K are written in the contract, but S_T will be known only at the expiry time. For which number K is the value of this contract equal to 0 at time 0?

Suppose that we have, besides buying the contract, two other options to invest our money:

(i) We can put our money in a savings account against a fixed, predetermined interest rate r. One unit of money placed in the account grows to e^{rt} units during a time interval $[0, t]$ and is freely available. A negative balance in our account is permitted, thus allowing us to borrow money at the same interest rate r. If we borrow one unit at time 0, than we owe e^{rt} units at time t, which is equivalent to having a capital of $-e^{rt}$ units.

(ii) We can invest in the asset. The asset price S_t at time t is a stochastic variable, dependent on t. It may be assumed that we know the probability distributions of these variables. For instance, a popular model is that the variable S_T/S_0 is log normally distributed.

We discuss two answers to the pricing question.

A naive (wrong) answer is to argue as follows. A payment of $S_T - K$ at time T is worth $e^{-rT}(S_T - K)$ at time 0. This is an unknown stochastic quantity from the perspective of time 0. The reasonable price is the expectation $\mathrm{E}\big(e^{-rT}(S_T - K)\big)$ of this variable. The strike price that gives the value 0 is therefore the solution of the equation

$$0 = \mathrm{E}\big(e^{-rT}(S_T - K)\big) = e^{-rT}(\mathrm{E}S_T - K).$$

In other words, $K = \mathrm{E}S_T$. For instance, using a log normal distribution with parameters μT and $\sigma^2 T$ for S_T/S_0, i.e. $\log(S_T/S_0)$ is normally distributed with these parameters, we find that

$$K = \mathrm{E}S_T = S_0 e^{\mu T + \frac{1}{2}\sigma^2 T}.$$

Somewhat surprisingly, this does not depend on the interest rate, meaning that somehow no trade-off has been made between investing in the asset or putting money in the savings account.

An accepted solution to the pricing question compares buying the contract to a "hedging strategy", as follows:

(i) at time 0 borrow an amount S_0 at interest rate r, and buy the asset at price S_0,

(ii) sit still until expiry time T.

At time T we own the asset, worth S_T, and we owe $S_0 e^{rT}$ to the bank. If $K = S_0 e^{rT}$, then together this is exactly the amount $S_T - K$ paid by the forward contract at time T. Thus in that case we ought to be indifferent towards buying the forward or carrying out the strategy (i)-(ii). The strategy costs us nothing at time 0 and hence $K = S_0 e^{rT}$ is the correct strike price for a forward with value 0 at time 0; this is also a maximal price, that the buyer of the contract can accept. Since the seller of the contract can perform a symmetric reasoning with $S_0 e^{rT}$ as a minimal accepted price, $K = S_0 e^{rT}$ is the unique *fair* price acceptable to both parties.

The correct solution $K = S_0 e^{rT}$ does not depend on the distribution of the asset price S_T, and in fact depends on the asset price process only through its (observable) value at time 0. This is remarkable if you are used to evaluate future gains through expectations on the random variables involved. Is there no role for probability theory in evaluating financial contracts?

There is. First we note that the expected gain of owning a contract is equal to $\mathrm{E}(S_T - K)$, which does depend on the distribution of the asset price. It evaluates to

$$S_0(e^{\mu T + \frac{1}{2}\sigma^2 T} - e^{rT})$$

if the asset price is log normally distributed and $K = S_0 e^{rT}$.

Second, it turns out that the correct solution can in fact be found from computing an expectation, but the expectation should be computed under a special probability measure, called a "martingale measure". To evaluate the price of a forward, this route would be overdone, as the preceding hedging strategy is explicit and simple. However, the prices of other contracts may not be so easy to evaluate, and in general probabilities and expectations turn out to be very useful. For instance, consider a *European call option*, which is a contract that pays an amount $(S_T - K)^+$ at time T, i.e. $S_T - K$ if $S_T > K$ and nothing otherwise. There is no simple universal hedging strategy to price this contract, but it turns out that given reasonable probabilistic models for the asset price process S_t, there are more complicated trading strategies that allow to parallel the preceding reasoning. These strategies require continuous trading during the term $[0, T]$ of the contract, and a big mathematical apparatus for their evaluation.

1.2 Note on Continuous Compounding

Our model for a savings account is that a capital of size R_0 placed in such an account at some time t increases to the amount $R_0 e^{r\Delta T}$ at time $t + \Delta t$. The capital in the account remains ours without restriction: we can withdraw it at no cost at any time. The constant r is the *continuously compounded interest rate* and is not quite an ordinary rate for a savings account, which is more often a yearly or monthly rate. If interest is added to the account at the end of a time period of one unit, then this would increase the capital from R_0 to $R_1 = (1 + r_1)R_0$, the interest being $r_1 R_0$, and r_1 being the rate per time unit. If instead we would obtain the interest in two installments, the first after half a time unit, and the second after one time unit, then the initial capital would first increase to $R_{1/2} = (1 + r_2)R_0$ and next to $R_1 = (1 + r_2)R_{1/2} = (1 + r_2)^2 R_0$. The second time we receive "interest on interest". The rate r_2 would be the rate per half time unit, and hence $2r_2$ should be compared to r_1. However, the comparison would not be exact, because $(1+r_1/2)^2 > 1+r_1$. It would be logical that r_1 and r_2 relate through the equation $(1 + r_2)^2 = 1 + r_1$, apart from possibly a correction for the benefit of early payment in the second scheme. The interest on interest is making the difference.

We could continue this thought experiment and break the time unit in n equal parts. The reasonable rate r_n per $(1/n)$th time unit would satisfy $(1 + r_n)^n = 1 + r_1$, or

$$\Big(1 + \frac{nr_n}{n}\Big)^n = 1 + r_1.$$

Here nr_n is the rate per time unit. Taking the limit as $n \to \infty$, and assuming that nr_n tends to a limit r we obtain the equation $e^r = 1 + r_1$. This r is the rate of interest per time unit earned in a savings account in which the interest is added "continuously in time". Using the continuous rate r is convenient, as exponentials multiply.

2
Binomial Tree Model

A *financial derivative* is a contract that is based on the price of an underlying asset, such as a stock price or bond price. An "option", of which there are many different types, is an important example. A main objective of "financial engineering" is to find a "fair price" of such a derivative, where by fair we mean a price acceptable both for the buyer and the seller. Following the work by Black and Scholes in the 1970s the prices of derivatives are found through the principle of "no arbitrage" introduced in the previous chapter.

In this chapter we discuss, as an introduction, the pricing of a European call option using a discrete time framework.

2.1 One Period Model

Suppose that at time 0 we can invest in an asset with price s_0, or put money in a savings account with a fixed interest rate. We model the asset price at time 1 as a random variable S_1 that can take only two values

$$\begin{aligned} \mathbb{P}(S_1 = us_0) &= p, \\ \mathbb{P}(S_1 = ds_0) &= 1 - p. \end{aligned} \tag{2.1}$$

Here u (for "up") and d (for "down") are two known constants, with $u > d$, and p is a number in $[0, 1]$ that may be unknown. We assume that an amount of one unit of money put in the savings account at time 0 grows to a guaranteed amount of e^r units at time 1.

We want to find the fair price at time 0 of a contract that pays the amount C at time 1, where $C = C(S_1)$ may depend on the (unknown) value of the asset at the payment date.

2.2 Example. A European call option corresponds to $C = (S_1 - K)^+$, for a given strike price K. The payment on a forward contract is equal to $C = S_1 - K$. □

Suppose that at time 0 we buy ϕ_0 assets and put an amount of ψ_0 money units in the savings account. Then we have a "portfolio" (ϕ_0, ψ_0) whose worth at time 0 is given by

$$V_0 = \phi_0 s_0 + \psi_0 \cdot 1. \tag{2.3}$$

If we do not trade between the times 0 and 1, then the value of the portfolio changes to its value V_1 at time 1, given by

$$V_1 = \phi_0 S_1 + \psi_0 e^r.$$

From the perspective of today (time 0) this is a random variable, that we cannot know with certainty. However, the asset can only take the values us_0 and ds_0 at time 1. In the first case the contract is worth $C(us_0)$ at time 1, whereas in the second case it is worth $C(ds_0)$ at time 1. The value of the portfolio is equal to $\phi_0 us_0 + \psi_0 e^r$ or $\phi_0 ds_0 + \psi_0 e^r$ in the two cases. Suppose that we fix the portfolio (ϕ_0, ψ_0) so that its value at time 1 agrees exactly with the contract, for each of the two possibilities, i.e.

$$\begin{cases} \phi_0 us_0 + \psi_0 e^r = C(us_0), \\ \phi_0 ds_0 + \psi_0 e^r = C(ds_0). \end{cases} \tag{2.4}$$

This portfolio will cost us V_0 at time 0, and is guaranteed to have the same value at time 1 as the contract with claim $C(S_1)$, whether the asset moves up or down. We should therefore have no preference for the portfolio or the contract, and hence a fair price for the contract at time 0 is the price of the portfolio, i.e. V_0 corresponding to the portfolio (ϕ_0, ψ_0) satisfying the equations (2.4).

The equations (2.4) form a system of two linear equations in the unknowns ϕ_0 and ψ_0 and can be solved to give

$$\begin{aligned} \phi_0 &= \frac{C(us_0) - C(ds_0)}{us_0 - ds_0}, \\ \psi_0 &= e^{-r}\Big(\frac{uC(ds_0) - dC(us_0)}{u - d}\Big). \end{aligned}$$

Inserting this in the equation (2.3), we see that this portfolio can be acquired at time zero for the amount

$$V_0 = e^{-r}\big(qC(us_0) + (1-q)C(ds_0)\big), \qquad q = \frac{e^r - d}{u - d}. \tag{2.5}$$

This is the fair price of the contract at time 0.

For $d \le e^r \le u$ the number q is contained in the interval $[0, 1]$ and can be considered an alternative probability for the upward move of the asset process. In general, this probability is different from the probability p, which turns out to be unimportant for the fair price. It can be seen that q is the unique probability such that

$$\mathrm{E}_q(e^{-r}S_1) = s_0. \tag{2.6}$$

Here the subscript q in E_q tells us to evaluate the expectation of the random variable S_1 using the distribution given in (2.1) but with q replacing p.

Furthermore, the price of the contract can be written as

$$V_0 = \mathrm{E}_q\bigl(e^{-r}C(S_1)\bigr).$$

We can write equation (2.6) also in the form $\mathrm{E}_q(e^{-r}S_1|\,S_0) = S_0$ (with S_0 the random variable that is equal to the constant s_0 with probability one), which expresses that the expected value of the discounted asset price $e^{-r}S_1$ given S_0 is equal to $S_0 = e^{-0}S_0$, or that the process $S_0, e^{-r}S_1$ is a "martingale".

2.7 Example (Forward). The value at time 0 of a forward is $\mathrm{E}_q e^{-r}(S_1 - K) = \mathrm{E}_q e^{-r}S_1 - e^{-r}K = S_0 - e^{-r}K$. The strike price that makes this value equal to zero is $K = e^r S_0$, which is the value $e^{rT}S_0$ with $T = 1$ found in Chapter 1. □

2.2 Two Period Model

Suppose that at time 0 we have the same possibilities for investing as in the preceding section, but we now consider a full trading horizon of three times: 0, 1, 2. We wish to evaluate a claim on the asset process payable at time 2.

Let the price of the asset at the three times be modelled by S_0, S_1, S_2, where we assume that $S_0 = s_0$ is fixed, S_1 is equal to either dS_0 or uS_0, and S_2 is equal to either dS_1 or uS_1. Thus the asset prices follow a path in a binary tree. We assume that at each node of the tree the decision to move up or down is made with probabilities p and $1-p$, independently for the different nodes.

Besides investing in the asset we may put money in a savings account (also a negative amount, indicating that we borrow money) at a fixed interest rate r. One unit in the savings account grows to e^r units at time 1, and to e^{2r} units at time 2.

The contract pays the amount $C = C(S_2)$ at time 2, and we wish to find its fair price at time 0.

We can evaluate the claim recursively, backwards in time. At time 2 the claim is worth $C = C(S_2)$. At time 1 there are two possibilities: either the asset price is dS_0 or it is uS_0. If we put ourselves at the perspective of time 1, then we know which of the two possibilities is realized. If s_1 is the realized value of S_1, then we can calculate the value of the claim (at time 1) using the one-period model as (cf. (2.5))

$$e^{-r}\bigl(qC(us_1) + (1-q)C(ds_1)\bigr).$$

For the two possibilities for s_1 this gives

$$\begin{cases} e^{-r}\bigl(qC(uds_0) + (1-q)C(d^2 s_0)\bigr), & \text{if } S_1 = ds_0,\\ e^{-r}\bigl(qC(u^2 s_0) + (1-q)C(dus_0)\bigr), & \text{if } S_1 = us_0.\end{cases}$$

This is the value of the contract at time 1, as a function of the asset price S_1 at time one. We can think of this value as the pay-off on our contract at time 1, and next apply the one-period model a second time to see that the value of the contract at time 0 is given by (cf. (2.5))

$$e^{-r}\Bigl(q\bigl[e^{-r}\bigl(qC(u^2 s_0) + (1-q)C(dus_0)\bigr)\bigr] \\ + (1-q)\bigl[e^{-r}\bigl(qC(uds_0) + (1-q)C(d^2 s_0)\bigr)\bigr]\Bigr).$$

This equation can be rearranged as

$$e^{-2r}\bigl(q^2 C(u^2 s_0) + 2q(1-q)C(uds_0) + (1-q)^2 C(d^2 s_0)\bigr) = \mathrm{E}_q\bigl(e^{-2r}C(S_2)\bigr).$$

Hence once again the price is the expectation of the discounted claim, presently $e^{-2r}C(S_2)$, under the probability measure on the tree given by the branching probability q.

From the one-period model we know that $\mathrm{E}_q(e^{-r}S_2|\, S_1, S_0) = S_1$ and $\mathrm{E}_q(e^{-r}S_1|\, S_0) = S_0$. Together these equations show that, for $n = 0, 1$,

$$\mathrm{E}_q(e^{-r(n+1)}S_{n+1}|\, S_n, \ldots, S_0) = e^{-nr}S_n.$$

We shall later summarize this by saying that the process $S_0, e^{-r}S_1, e^{-2r}S_2$ is a martingale.

2.3 N Period Model

We can price a claim in a binomial tree model with N periods by extending the backwards induction argument. The fair price of a claim $C(S_N)$ is given by

$$\mathrm{E}_q\bigl(e^{-Nr}C(S_N)\bigr).$$

Here S_N is equal to $u^{X_N}d^{Y_N}S_0$ for X_N and Y_N the number of up-moves and down-moves, respectively. Of course $Y_N = N - X_N$ and the variable X_N possesses a binomial distribution with parameters N and success probability

q in the preceding display (and p in the real world). This allows to write the fair price as a sum.

The induction argument to prove this formula is straightforward, but tedious. We shall give a much prettier derivation after developing some "martingale theory", which will also give us the intuition needed to tackle the continuous time models later on.

Note that the derivation of the formula involves portfolios (ϕ, ψ) that are defined in terms of the constants u and d. Thus intuitively the reasoning seems to require that we know these constants. In practice one might observe the prices of some options (at least two for u and d and one more for q) on the market, and next calibrate the constants u and d so that the prices given by the formula agree with the market prices. We would do this only if we believe the binomial tree model. Most people would consider the continuous time models more believable. One unpleasant aspect of the binomial tree model is that it is essential that the splits in the tree are binary. If there were three or more possible moves, everything would fall down.

2.8 EXERCISE. Verify this for the one-period model, replacing (2.1) by the assumption that $\mathbb{P}(S_1/S_0 = x) > 0$ for $x \in \{d, i, u\}$ with probabilities p_d, p_i and p_u. Revisit (2.4), which becomes a system of three equations. Why is there no solution if $d < i < u$?

3
Discrete Time Stochastic Processes

3.1 Stochastic Processes

A *stochastic process* in discrete time is a (finite or infinite) sequence $X = (X_0, X_1, \ldots)$ of random variables or vectors, defined on a given probability space. Mathematically, random variables are maps $X_n: \Omega \to \mathbb{R}$ that map outcomes $\omega \in \Omega$ into numbers $X_n(\omega)$. The stochastic process $X_0, X_1, \ldots,$ maps every outcome ω into a sequence of numbers $X_0(\omega), X_1(\omega), \ldots$, called a *sample path.*

The best way to think of a stochastic process is to visualize the sample paths as "random functions". We generate an outcome ω according to some probability measure on the set of all outcomes and next have a function $n \mapsto X_n(\omega)$ on the domain $\mathbb{N} \cup \{0\}$. This domain is referred to as the set of "discrete times".

3.1 Example (Binomial tree model). The binomial tree model for the stock price is a stochastic process $S_0, S_1, \ldots, S_N$, where each possible sample path is given by a path in the binomial tree, and the probability of a sample path is the product of the probabilities on the branches along the path.

As indicated before, the best way to think about this stochastic process is as a random function, generated according to the probabilities attached to the different paths on the tree. The preceding description gives an intuitively clear description of the binomial tree process, but for later use it is instructive to define the stochastic process also formally as a map on a given outcome space. One possibility is to take Ω to be equal to the set of N-tuples $\omega = (\omega_1, \ldots, \omega_N)$, where each $\omega_i \in \{0, 1\}$. The appropriate probability measure is

$$\mathbb{P}(\{(\omega_1,\ldots,\omega_N)\}) = p^{\#1\le i\le N:\omega_i=1}(1-p)^{\#1\le i\le N:\omega_i=0},$$

and the stochastic process can be formally defined by setting $S_0 = s_0$ and, for $n = 1,2,\ldots,N$,

$$S_n(\omega_1,\ldots,\omega_N) = S_0 u^{\#1\le i\le n:\omega_i=1} d^{\#1\le i\le n:\omega_i=0}.$$

Thus $\omega_i = 1$ indicates that the sample path goes up in the tree at time i, whereas $\omega_i = 0$ indicates a down move. The value S_n at time n is determined by the total number of moves up and down in the tree up till that time. □

3.2 Conditional Expectation

For a discrete random variable X and a discrete random vector Y, the *conditional expectation* of X given the event $Y = y$ is given by

$$\mathrm{E}(X|\,Y=y) = \sum_x x\mathbb{P}(X=x|\,Y=y).$$

If we write this function of y as $f(y) = \mathrm{E}(X|\,Y=y)$, then we write $\mathrm{E}(X|\,Y)$ for $f(Y)$. This is a random variable, called the conditional expectation of X given Y. Some important rules are given in the following lemma.

3.2 Lemma.
(i) $\mathrm{E}\mathrm{E}(X|\,Y) = \mathrm{E}X$.
(ii) $\mathrm{E}(\mathrm{E}(X|\,Y,Z)|\,Z) = \mathrm{E}(X|\,Z)$ *(tower property).*
(iii) $\mathrm{E}(X|\,Y) = X$ *if* $X = f(Y)$ *for some function* f.
(iv) $\mathrm{E}(X|\,Y) = \mathrm{E}X$ *if* X *and* Y *are independent.*

These rules can be proved from the definition, but are intuitively clear. The first rule says that the expectation of a variable X can be computed in two steps, first using the information on another variable Y, and next taking the expectation of the result. Assertion (ii) gives exactly the same property, with the difference that every of the expectations are computed conditionally on a variable Z. Rule (iii) says that we can predict a variable X exactly if X is a function of a known variable Y, which is obvious.

3.3 EXERCISE. Suppose you generate N points in the interval $[0,1]$ as follows. First you choose N from the Poisson distribution with mean 100. Next given $N=n$ you generate a random sample of n random variables from a given distribution F on $[0,1]$. What is the expected number of points in an interval $B\subset[0,1]$?

We shall use the notation $\mathrm{E}(X|\,Y)$ also if X or Y are continuous random variables or vectors. Then the preceding definition does not make sense, because the probabilities $\mathbb{P}(X=x|\,Y=y)$ are not defined if $\mathbb{P}(Y=y)=0$, which is the case for continuous random variable X. However, the conditional expectation $\mathrm{E}(X|\,Y)$ can still be defined as a function of Y, namely as the function such that, for every function g,

$$\mathrm{E}\bigl(\mathrm{E}(X|\,Y)g(Y)\bigr)=\mathrm{E}\bigl(Xg(Y)\bigr).$$

The validity of this equality in the case of discrete random variables can be checked in the same manner as the validity of the three rules in the lemma. For general random variables X and Y we take this as a definition of conditional expectation, where it is also understood that $\mathrm{E}(X|\,Y)$ must be a function of Y. The three rules of the lemma continue to hold for this extended definition of conditional expectation.

In most cases this abstract definition agrees perfectly with your intuition of the expected value of X given Y. However, in some cases where there are many sets $\{Y=y\}$, all with probability zero, your intuition could deceive you. The problem is then usually that there are several equally "good", but incompatible intuitions.

3.3 Filtration

A *σ-field* is a collection of events.[†] A *filtration* in discrete time is an increasing sequence $\mathcal{F}_0\subset\mathcal{F}_1\subset\cdots$ of σ-fields, one per time instant. The σ-field $\mathcal{F}_n$ may be thought of as the events of which the occurrence is determined at or before time n, the "known events" at time n. Filtrations are important to us, because they allow to model the flow of information. Of course, the information increases as time goes by.

The filtrations that are of interest to us are generated by stochastic processes. The *natural filtration* of a stochastic process $X_0,X_1,\ldots$ is defined by

$$\mathcal{F}_n=\bigl\{(X_0,X_1,\ldots,X_n)\in B\colon B\subset\mathbb{R}^{n+1}\bigr\}.$$

[†] A rigorous mathematical definition includes the requirements (i) $\emptyset\in\mathcal{F}$; (ii) if $F\in\mathcal{F}$, then $F^c\in\mathcal{F}$; (iii) if $F_1,F_2,\ldots,\in\mathcal{F}$, then $\cup_iF_i\in\mathcal{F}$. The *Borel σ-field* is the smallest σ-field of subsets of $\mathbb{R}^n$ that satisfies the properties and contains all intervals. In the following the sets B should be required to be Borel sets.

Thus $\mathcal{F}_n$ contains all events that depend on the first $(n+1)$ elements of the stochastic process. It gives the "history" of the process up till time n.

A convenient notation to describe a σ-field corresponding to observing a random vector X is $\sigma(X)$. Thus $\sigma(X)$, called the *σ-field generated by* X, consists of all events that can be expressed in X: events of the type $\{X \in B\}$. In this notation, the natural filtration of a stochastic process $X_0, X_1, \ldots$ can be written as $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$.

We say that a process $X_0, X_1, \ldots$ is *adapted* to the filtration $(\mathcal{F}_n)_{n\geq 0}$ if $\sigma(X_n) \subset \mathcal{F}_n$ for every n. Thus the events connected to an adapted process up to time n are known at time n. The natural filtration corresponding to a process is the smallest filtration to which it is adapted. If the process $Y_0, Y_1, \ldots$ is adapted to the natural filtration of a stochastic process $X_0, X_1, \ldots$, then for each n the variable Y_n is a function $\phi_n(X_0, X_1, \ldots, X_n)$ of the sample path of the process X up till time n.

We say that a process $Y_0, Y_1, \ldots$ is *predictable* relative to the filtration $(\mathcal{F}_n)_{n\geq 0}$ if $\sigma(Y_n) \subset \mathcal{F}_{n-1}$ for each n. Thus the events connected to a predictable process are known one time instant before they happen. If $\mathcal{F}_n$ is generated by (X_n), then this equivalent to Y_n being a function $Y_n = \phi_n(X_0, \ldots, X_{n-1})$ of the history of the process $X_0, X_1, \ldots$ up till time $n-1$, for some ϕ_n.

If $\mathcal{F}$ is the σ-field generated by Y, then we also write $\mathrm{E}(X|\,\mathcal{F})$ for the random variable $\mathrm{E}(X|\,Y)$. Thus $\mathrm{E}(X|\,\mathcal{F})$ is the expected value of X given the information $\mathcal{F}$. The *trivial σ-field* $\{\emptyset, \Omega\}$ is the σ-field containing no information.

3.4 Lemma.

(i) $\mathrm{E}(X|\,\{\emptyset,\Omega\}) = \mathrm{E}X$.

(ii) for two σ-fields $\mathcal{F} \subseteq \mathcal{G}$ there holds $\mathrm{E}(\mathrm{E}(X|\,\mathcal{G})|\,\mathcal{F}) = \mathrm{E}(X|\,\mathcal{F})$ *(tower property).*

(iii) $\mathrm{E}(X|\,Y) = X$ *if* $\sigma(X) \subset \sigma(Y)$.

3.4 Martingales

A stochastic process $X_0, X_1, \ldots$ is a *martingale* relative to a given filtration $(\mathcal{F}_n)_{n\geq 0}$ if it is adapted to this filtration and $\mathrm{E}(X_n|\,\mathcal{F}_m) = X_m$ for every $m < n$.

The martingale property is equivalent to $\mathrm{E}(X_n - X_m|\,\mathcal{F}_m) = 0$ for every $m < n$, expressing that the *increment* $X_n - X_m$ given the "past" $\mathcal{F}_m$ has expected value 0. A martingale is a stochastic process that, on the average, given the past, does not grow or decrease.

3.5 Example (Random walk). Let $X_1, X_2, \ldots$ be independent random variables with mean zero. Define $S_0 = 0$ and $S_n = X_1 + \cdots + X_n$ for $n \in \mathbb{N}$. Then S is a martingale relative to its natural filtration. Indeed, S is adapted by construction and $\mathrm{E}(S_{n+1}|\,\mathcal{F}_n) = \mathrm{E}(S_n + X_{n+1}|\,\mathcal{F}_n) = S_n + \mathrm{E}X_{n+1} = S_n$, since $S_n \in \mathcal{F}_n$ and X_{n+1} is independent of $\mathcal{F}_n$. □

3.6 Example (Doob martingale). If Y is a random variable with $\mathrm{E}|Y| < \infty$ and $\mathcal{F}_n$ an arbitrary filtration, then $X_n = \mathrm{E}(Y|\,\mathcal{F}_n)$ defines a martingale. This can be proved from the tower property of conditional expectations: $\mathrm{E}\big(\mathrm{E}(Y|\,\mathcal{F}_n)|\,\mathcal{F}_m\big) = \mathrm{E}(Y|\,\mathcal{F}_m)$ for any $m < n$. □

The martingale property can also be equivalently described through one step ahead expectations. A process is a martingale if $\mathrm{E}(X_{n+1}|\,\mathcal{F}_n) = X_n$ for every n.

3.7 EXERCISE. Use the tower property to prove this.

3.8 EXERCISE. Let $Z_1, Z_2, \ldots$ be independent $N(0,1)$-variables. Show that the sequence $S_n = \exp(\sum_{i=1}^n Z_i - \frac{1}{2}n)$ forms a martingale.

3.9 EXERCISE. In a branching process we start with $N_0 = 1$ individuals at time 0, and at each time n each individual has a random number of offspring independent of the other individuals, where the number is chosen from a fixed distribution. The new generation consists of the offspring only. Thus given that there are $N_n = k$ individuals at time n the number of individuals N_{n+1} in the $(n+1)$th generation is distributed as $\sum_{i=1}^k X_i$ for i.i.d. random variables $X_1, \ldots, X_k$. Show that $N_0, N_1, \ldots,$ is a martingale if and only if $\mathrm{E}X_i = 1$.

3.5 Change of Measure

If there are two possible probability measures $\mathbb{P}$ and $\mathbb{Q}$ on the set of outcomes, and X is a martingale relative to $\mathbb{Q}$, then typically it is not a martingale relative to $\mathbb{P}$. This is because the martingale property involves the expected values, and hence the probabilities of the various outcomes. An important tool in finance is to change a given measure $\mathbb{P}$ into a measure $\mathbb{Q}$ making a discounted asset process into a martingale.

3.10 Example. In the binomial tree model with $\mathcal{F}_n$ the natural filtration of $S_0, S_1, \ldots$ and $\mathbb{P}(S_{n+1} = uS_n|\,\mathcal{F}_n) = 1 - \mathbb{P}(S_{n+1} = dS_n|\,\mathcal{F}_n) = p$, we have that

$$\mathrm{E}(S_{n+1}|\,\mathcal{F}_n) = uS_n p + dS_n(1-p) = S_n\big[up + d(1-p)\big].$$

This is equal to S_n if u, d and p satisfy the equation $up + d(1-p) = 1$. For instance, if $u = 2$ and $d = 1/2$, then this is true if and only if $p = 1/3$. □

3.11 Example (Discounted stock). In the binomial tree model as in the preceding example consider the discounted process $S_0, e^{-r}S_1, e^{-2r}S_2, \ldots$. The one step ahead expectations are given by

$$\mathrm{E}(e^{-(n+1)r}S_{n+1}|\,\mathcal{F}_n) = ue^{-(n+1)r}S_n p + de^{-(n+1)r}S_n(1-p).$$

The discounted process is a martingale only if the right side is equal to $e^{-nr}S_n$. This is the case only if

$$p = \frac{e^r - d}{u - d}.$$

This value of p is contained in the unit interval and defines a probability only if $d \leq e^r \leq u$. In that case the discounted process is a martingale. □

3.6 Martingale Representation

In Example 3.11 we have seen that the process $\tilde{S}$ defined by $\tilde{S}_n = e^{-rn}S_n$ in the binomial tree model is a martingale if the tree is equipped with the probability

$$q = \frac{e^r - d}{u - d}.$$

In this section we shall show that all other martingales in this setting can be derived from $\tilde{S}$ in the sense that the increments $\Delta M_n = M_n - M_{n-1}$ of any martingale $M_0, M_1, \ldots$ must be multiples $\phi_n \Delta \tilde{S}_n$ for a predictable process $\phi = (\phi_0, \phi_1, \ldots)$. In other words, the change ΔM_n of an arbitrary martingale at time $n-1$ is proportional to the change in $\tilde{S}$, with the proportionality constant ϕ_n being a function of the preceding values $\tilde{S}_0, \ldots, \tilde{S}_{n-1}$ of the process $\tilde{S}$. At time $n-1$ the only randomness to extend $M_0, \ldots, M_{n-1}$ into M_n is in the increment $\Delta \tilde{S}_n$.

3.12 Theorem. *If M is a martingle on the binomial tree model of Example 3.1 with $q = (e^r - d)/(u - d)$ with filtration $\mathcal{F}_n = \sigma(S_0, \ldots, S_n)$, then there exists a predictable process $\phi_0, \phi_1, \ldots$ such that, for every $n \in \mathbb{N}$,*

$$\Delta M_n = \phi_n \Delta \tilde{S}_n.$$

Proof. Because M is adapted to the filtration generated by $S_0, S_1, \ldots$, for each n the variable M_n is a function of $S_0, \ldots, S_n$. Given $\mathcal{F}_{n-1}$ the values of $S_0, \ldots, S_{n-1}$ are fixed and hence M_n can assume only two possible values, corresponding to a downward or upward move in the tree. By a similar argument we see that the variable M_{n-1} is fixed given $\mathcal{F}_{n-1}$, and hence ΔM_n has two possible values given $\mathcal{F}_{n-1}$. If we fix $S_0, \ldots, S_{n-1}$, then we can write $\Delta M_n = g_n(S_n)$ for some function g_n (which depends on the fixed values of $S_0, \ldots, S_{n-1}$). Similarly we can write $\Delta \tilde{S}_n = f_n(S_n)$. The martingale properties of the processes M (by assumption) and $\tilde{S}$ (by Example 3.11) give that $\mathrm{E}_q(\Delta M_n | \mathcal{F}_{n-1}) = 0 = \mathrm{E}_q(\Delta \tilde{S}_n | \mathcal{F}_{n-1})$, or

$$qg_n(uS_{n-1}) + (1-q)g_n(dS_{n-1}) = 0,$$
$$qf_n(uS_{n-1}) + (1-q)f_n(dS_{n-1}) = 0.$$

It follows from this that $g_n(uS_{n-1})/f_n(uS_{n-1}) = g_n(dS_{n-1})/f_n(dS_{n-1})$. We can define ϕ_n as this common ratio. ∎

4
Binomial Tree Model Revisited

Suppose that the price S_t at time t is a stochastic process described by the binomial tree model of Example 3.1, where it is assumed that the numbers u and d are known. We choose the filtration equal to $\mathcal{F}_n = \sigma(S_0, \ldots, S_n)$, so that the (only) information available at time n consists of observation of the asset price process until that time.

In addition to the asset with price S we can save or borrow money at a fixed rate of interest r. We assume that $d \leq e^r \leq u$. This is a reasonable assumption, because if $e^r < d$ then the returns on the asset are with certainty bigger than the return on the savings account, whereas if $e^r > u$, then the returns are with certainty smaller. Then the riskless savings account is never or always preferable over the risky asset, respectively, and a reasonable portfolio will consist of only one type of investment.

We equip the branches of the binomial tree with the probability $q = (e^r - d)/(u - d)$, rather than a possible real world probability based on historical analysis. Example 3.11 shows that this gives the unique probability measure on the tree that renders the discounted asset process $\tilde{S}_0, \tilde{S}_1, \ldots$, where $\tilde{S}_n = e^{-rn} S_n$, into a martingale.

A *claim* is a nonnegative function of $C = C(S_0, \ldots, S_N)$, where N is the expiry time. Given a claim define the stochastic process

$$\tilde{V}_n = \mathrm{E}_q(e^{-rN} C \,|\, \mathcal{F}_n).$$

The index q on the expectation operator E indicates that we compute expectations under the martingale measure q. In view of Example 3.6 the process $\tilde{V}$ is a martingale. Therefore, by Theorem 3.12 there exists a predictable process ϕ such that

$$\Delta \tilde{V}_n = \phi_n \Delta \tilde{S}_n, \qquad n = 1, \ldots, N. \tag{4.1}$$

Given this process we define another process ψ by

$$\psi_n = \tilde{V}_{n-1} - \phi_n \tilde{S}_{n-1}. \tag{4.2}$$

From the facts that ϕ is predictable and $\tilde{V}$ and $\tilde{S}$ are adapted, it follows that the process ψ is predictable.

We now interpret (ϕ_n, ψ_n) as a portfolio at time n:

(i) ϕ_n is the number of assets held during the period $(n-1, n]$.

(ii) ψ_n is the number of units in the saving account during the period $(n-1, n]$.

Because both processes are predictable, the portfolio (ϕ_n, ψ_n) can be created at time $n-1$ based on information gathered up to time $n-1$, i.e. based on observation of $S_0, S_1, \ldots, S_{n-1}$. We shall think of the assets and savings changing value (from S_{n-1} to S_n and $e^{r(n-1)}$ to e^{rn}) exactly at time n, and of adapting our portfolio just after time n. Then the value of the portfolio at time n is

$$V_n = \phi_n S_n + \psi_n e^{rn}.$$

Just after time n we change the content of the portfolio; the value of the new portfolio is equal to

$$\phi_{n+1} S_n + \psi_{n+1} e^{rn}.$$

The following theorem shows that this amount is equal to V_n and hence the new portfolio can be formed without additional money: the portfolio process (ϕ, ψ) is *self-financing*.

Furthermore, the theorem shows that the value V_N of the portfolio is exactly equal to the value of the claim C at the expiry time N. As a consequence, we should be indifferent to owning the contract with claim C or the portfolio (ϕ_0, ψ_0) at time 0, and hence the "just price" of the contract is the value V_0 of the portfolio.

4.3 Theorem. *The portfolio process (ϕ, ψ) defined by (4.1)-(4.2) is self-financing. Furthermore, its value process V is nonnegative and satisfies $\tilde{V}_n = e^{-rn} V_n$ for every n. In particular $V_N = C$ with probability one.*

Proof. The equation (4.2) that defines ψ_{n+1} can be rewritten in the form $\phi_{n+1} S_n + \psi_{n+1} e^{rn} = e^{rn} \tilde{V}_n$, and $\tilde{V}_n = \tilde{V}_{n-1} + \Delta \tilde{V}_n$, where $\Delta \tilde{V}_n = \phi_n \Delta \tilde{S}_n$ by (4.1). Therefore,

$$\begin{aligned}
\phi_{n+1} S_n + \psi_{n+1} e^{rn} - V_n &= e^{rn} \tilde{V}_n - V_n \\
&= e^{rn} \tilde{V}_{n-1} + e^{rn} \phi_n \Delta \tilde{S}_n - (\phi_n S_n + \psi_n e^{rn}) \\
&= e^{rn} \tilde{V}_{n-1} + e^{rn} \phi_n (\Delta \tilde{S}_n - \tilde{S}_n) - \psi_n e^{rn} \\
&= e^{rn} \tilde{V}_{n-1} - e^{rn} \phi_n \tilde{S}_{n-1} - \psi_n e^{rn}.
\end{aligned}$$

The right side is zero by the definition of ψ_n in (4.2). Thus the portfolio is self-financing, as claimed.

It also follows from these equations that $e^{rn} \tilde{V}_n - V_n = 0$, whence $\tilde{V}_n = e^{-rn} V_n$ is the discounted value of the portfolio at time n, for every

n. Because $\tilde{V}_n = \mathrm{E}_q(e^{-rN}C|\,\mathcal{F}_n)$ is nonnegative, so is V_n. Furthermore, the value of the portfolio at time N is

$$V_N = e^{rN}\tilde{V}_N = e^{rN}\mathrm{E}_q(e^{-rN}C|\,\mathcal{F}_N) = C,$$

since C is a function of $S_0, S_1, \ldots, S_N$, by assumption. ■

Because $V_N = C$ with certainty the portfolio (ϕ, ψ) is said to *replicate* the claim C. The claim value C is a random variable that depends on $S_0, \ldots, S_N$. However, no matter which path the asset prices take in the binomial tree, the portfolio always ends up having the same value as the claim. This is achieved by reshuffling assets and savings at each time n, based on the available information at that time, a strategy called *hedging*.

The fact that the value process of the claim is nonnegative ensures that the portfolio management can be implemented in practice. If we have sufficient funds to form the portfolio at time 0, then we never run into debt when carrying out the hedging strategy.

We interpret the value of the portfolio at time 0, the amount of money needed to create the portfolio (ϕ_0, ψ_0), as the "just price" of the claim at time 0. This is

$$V_0 = \tilde{V}_0 = \mathrm{E}_q(e^{-rN}C|\,\mathcal{F}_0) = \mathrm{E}_q(e^{-rN}C).$$

The second equality follows from the fact that $\tilde{V}$ is a martingale under q. Note that the formula expresses the price in the claim C without intervention of the portfolio processes. These were only a means to formulate the economic argument.

4.4 Example (Forward). The claim of a forward with strike price K is $C = S_N - K$. The value at time 0 is equal to $\mathrm{E}_q e^{-rN}(S_N - K) = \mathrm{E}_q e^{-rN}S_N - e^{-rN}K$. Because the process $e^{-rn}S_n$ is a martingale under the (martingale) measure q the expected value $\mathrm{E}e^{-rn}S_n$ is constant in n. Hence the value of the forward contract at time 0 is equal to $\mathrm{E}_q e^{-r0}S_0 - e^{-rN}K$. The strike price K that makes the value equal to zero is $K = e^{rN}S_0$.

In Section 1.1 we obtained the same result by describing an explicit hedging strategy. □

4.5 Example (European call option). The claim of a European call option with strike price K is $C = (S_N - K)^+$. The fair price at time 0 is equal to $\mathrm{E}_q e^{-rN}(S_N - K)^+$. The variable S_N is distributed as $S_0 u^{X_N} d^{N-X_N}$, where X_N is the number of upward moves in the tree. Because the variable X_N is binomially distributed with parameters N and q (under the martingale measure), it follows that the value of the option at time 0 is equal to

$$e^{-rN}\sum_{x=0}^{N}(S_0 u^x d^{N-x} - K)^+ \binom{N}{x} q^x (1-q)^{N-x}.$$

This expression is somewhat complicated, but easy to evaluate on a computer. An alternative method of computation is backwards induction, as in Chapter 2. An approximation formula for large N is given in the next section. □

4.6 EXERCISE. Suppose that (ϕ_n, ψ_n) are predictable processes such that the process V defined by $V_n = \phi_n S_n + \psi_n e^{rn}$ satisfies $V_n = \phi_{n+1} S_n + \psi_{n+1} e^{rn}$. (In other words (ϕ, ψ) defines a self-financing portfolio process with value process V.) Show that the process $\tilde{V}$ defined by $\tilde{V}_n = e^{-rn} V_n$ satisfies $\Delta\tilde{V}_n = \phi_n \Delta\tilde{S}_n$, and conclude that it is a martingale under q. [This gives some motivation for the definitions used to prove the theorem.]

4.1 Towards Continuous Time

In the real world asset prices change almost continuously in time. The binomial tree model can approximate this if the number of steps N is large. Mathematically we can even compute limits as $N \to \infty$, in the hope that this gives a realistic model.

A limit exists only if we make special choices for the relative up and down moves u and d. Unless u and d tend to 1 as the number of moves N increases, the asset price will explode and our model does not tend to a limit. We shall think of the N moves in the binomial tree taking place in a fixed interval $[0, T]$, at the times $\delta, 2\delta, \ldots, N\delta$ for $\delta = T/N$. Then it is reasonable to redefine the interest rate in one time instant as $r\delta$, giving a total interest of r over the interval $[0, T]$. We also assume that, for given constants $\mu \in \mathbb{R}$ and $\sigma > 0$,

$$d = e^{\mu\delta - \sigma\sqrt{\delta}}, \qquad u = e^{\mu\delta + \sigma\sqrt{\delta}}.$$

These definitions satisfy that they approach 1 as the length δ of a time interval tends to zero. The exact definitions are somewhat special, but can be motivated by the fact that the resulting model tends to continuous time model considered later on.

The asset price at time N is equal to

$$S_N = S_0 u^{X_N} d^{N - X_N} = S_0 \exp\Bigl(\mu T + \sigma\sqrt{T}\frac{(2X_N - N)}{\sqrt{N}}\Bigr),$$

where X_N is the number of times the stock price goes up in the time span $1, 2, \ldots, N$.

In a standard model for the stock market the jumps up and down have equal probabilities. Then X_N is binomially $(N, \frac{1}{2})$-distributed and the "log returns" satisfy

$$\log \frac{S_N}{S_0} = \mu T + \sigma\sqrt{T}\frac{X_N - N/2}{\sqrt{N}/2} \rightsquigarrow N(\mu T, \sigma^2 T),$$

by the Central Limit theorem. Thus in the limit the log return at time T is normally distributed with drift μT and variance $\sigma^2 T$.

As we have seen the true distribution of the stock prices is irrelevant for pricing the option. Rather we need to repeat the preceding calculation using the martingale measure $q = q_N$. Under this measure X_N is binomially(N, q) distributed, for

$$\begin{aligned} q &= \frac{e^{rT/N} - e^{\mu T/N - \sigma\sqrt{T/N}}}{e^{\mu T/N + \sigma\sqrt{T/N}} - e^{\mu T/N - \sigma\sqrt{T/N}}} \\ &= \tfrac{1}{2} - \tfrac{1}{2}\sqrt{\frac{T}{N}}\Big(\frac{\mu + \frac{1}{2}\sigma^2 - r}{\sigma}\Big) + O\Big(\frac{1}{N}\Big), \end{aligned}$$

by a Taylor expansion. Then $q_N(1 - q_N) \to 1/4$ and

$$\begin{aligned} \log \frac{S_N}{S_0} &= \mu T + \sigma\sqrt{T}\Big(\frac{X_N - Nq_N}{\sqrt{N}/2} - \sqrt{T}\Big(\frac{\mu + \frac{1}{2}\sigma^2 - r}{\sigma}\Big)\Big) + O\Big(\frac{1}{\sqrt{N}}\Big) \\ &\rightsquigarrow N\big((r - \tfrac{1}{2}\sigma^2)T, \sigma^2 T\big). \end{aligned}$$

Thus, under the martingale measure, in the limit the stock at time T is log normally distributed with drift $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$.

Evaluating the (limiting) option price is now a matter of straightforward integration. For a claim $C = C(S_N)$ with expiry time T the "fair price" at time 0 is given by

$$e^{-rT}\mathrm{E}_{q_N} C(S_N) \approx e^{-rT}\mathrm{E}C(\bar{S}_T),$$

for $\log(\bar{S}_T/S_0)$ normally distributed with mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$.

4.7 Example (European call option). The (limiting) fair price of a European call option with expriry time T and strike price K is the expectation of $e^{-rT}(\bar{S}_T - K)^+$, where $\log(\bar{S}_T/S_0)$ possesses the log normal distribution with parameters $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$. This can be computed to be

$$S_0\Phi\Big(\frac{\log(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\Big) - Ke^{-rT}\Phi\Big(\frac{\log(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\Big).$$

This is the famous formula found by Black and Scholes in 1973 using a continuous time model. We shall recover it later in a continuous time setup. □

4.8 EXERCISE. Suppose that Z is standard normally distributed and $S > 0$, $\tau > 0$, μ and K are constants. Show that, for $z_0 = \tau^{-1}(\log(K/S) - \mu)$,

$$\mathrm{E}(Se^{\tau Z+\mu} - K)^+ = Se^{\mu+\frac{1}{2}\tau^2}\Phi(\tau - z_0) - K\Phi(-z_0).$$

5 Continuous Time Stochastic Processes

5.1 Stochastic Processes

A continuous-time *stochastic process* is an indexed collection of random variables $X = (X_t: t \geq 0)$, defined on a given probability space. Thus every X_t is a map $X_t: \Omega \to \mathbb{R}$ mapping outcomes $\omega \in \Omega$ into numbers $X_t(\omega)$. The functions $t \mapsto X_t(\omega)$ attached to the outcomes are called *sample paths*, and the index t is referred to as "time". The best way to think of a stochastic process is to view it as a "random function" on the domain $[0, \infty)$, with the sample paths as its realizations.

For any finite set $t_1 < t_2 < \cdots < t_k$ of time points the vector $(X_{t_1}, \ldots, X_{t_k})$ is an ordinary random vector in $\mathbb{R}^k$, and we can describe a great deal of the process by describing the distributions of all such vectors. On the other hand, qualitative properties such as continuity of differentiability of a sample path depend on infinitely many time points.

5.2 Brownian Motion

Brownian motion is a special stochastic process, which is of much interest by itself, but will also be used as a building block to construct other processes. It can be thought of as the "standard normal" process. A Brownian motion is often denoted by the letter W, after Wiener, who was among the first to study Brownian motion in a mathematically rigorous way. The distribution of Brownian motion is known as the "Wiener measure".

A stochastic process W is a *Brownian motion* if

(i) the *increment* $W_t - W_s$ is normally distributed with mean 0 and variance $t-s$, for any $0 \leq s < t$.
(ii) the *increment* $W_t - W_s$ is stochastically independent of $(W_u\colon u \leq s)$, for any $0 \leq s < t$.
(iii) $W_0 = 0$.
(iv) any sample path $t \mapsto W_t(\omega)$ is a continuous function.

It is certainly not clear from the definition that Brownian motion exists, in the sense that there exists a probability space with random variables W_t defined on it that satisfy the requirements (i)-(iv). However, it is a mathematical theorem that Brownian motion exists, and there are several constructive ways of exhibiting one. We shall take the existence for granted.

Properties (i) and (ii) can be understood in the sense that, given the sample path $(W_u\colon u \leq s)$ up to some point s, Brownian motion continues from its "present value" W_s by adding independent (normal) variables. In fact it can be shown that given $(W_u\colon u \leq s)$ the process $t \mapsto W_{s+t} - W_s$ is again a Brownian motion. Thus at every time instant Brownian motion starts anew from its present location, independently of its past.

The properties of Brownian motion can be motivated by viewing Brownian motion as the limit of the process in a binomial tree model, where starting from $S_0 = 0$ a process $S_1, S_2, \ldots$ is constructed by moving up or down 1 in every step, each with probability 1/2, i.e.

$$S_n = \sum_{i=1}^n X_i,$$

for an i.i.d. sequence $X_1, X_2, \ldots$ with $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$. For a given N we could place the values of the process $S_0, S_1, S_2, \ldots$ at the time points $0, 1/N, 2/N, \ldots$ and rescale the vertical axis so that the resulting process remains stable. This leads to the process $W^{(N)}$ given by

$$W_t^{(N)} = \frac{1}{\sqrt{N}} \sum_{i:i \leq tN} X_i.$$

By the Central Limit Theorem, as $N \to \infty$, with $\rightsquigarrow$ denoting convergence in distribution,

$$W_t^{(N)} - W_s^{(N)} = \frac{1}{\sqrt{N}} \sum_{i:sN<i\leq tN} X_i \rightsquigarrow N(0, t-s).$$

Furthermore, the variable on the left side is independent of the variables X_i with index i not contained in the sum, and hence of $W_s^{(N)}$. Thus in the limit as $N \to \infty$ the processes $W^{(N)}$ satisfy the properties (i)-(iii). It can indeed be shown that the sequence $W^{(N)}$ converges, in a suitable sense, to a Brownian motion process. The main challenge in proving existence of Brownian motion is the required continuity (iv) of the sample paths.

5.3 Filtrations

A *filtration* $(\mathcal{F}_t)_{t\geq 0}$ in continuous time is an increasing collection of σ-fields indexed by $[0,\infty)$. Thus $\mathcal{F}_s \subset \mathcal{F}_t$ for every $s < t$. A stochastic process $X = (X_t: t \geq 0)$ is *adapted* to a given filtration if $\sigma(X_t) \subset \mathcal{F}_t$ for every t. In such case, all events concerning the sample paths of an adapted process until time t are contained in $\mathcal{F}_t$. The *natural filtration* of a stochastic process $X = (X_t: t \geq 0)$ is

$$\mathcal{F}_t = \sigma(X_s: s \leq t).$$

This filtration corresponds exactly to observing the sample paths of X up to time t, and is the smallest filtration to which X is adapted.

5.4 Martingales

A *martingale* in continuous time relative to a given filtration $(\mathcal{F}_t)_{t\geq 0}$ is an adapted process X such that

$$\mathrm{E}(X_t|\,\mathcal{F}_s) = X_s, \qquad \text{every } s < t.$$

This property is equivalent to the increments $X_t - X_s$ having expected value 0 given the past and present: $\mathrm{E}(X_t - X_s|\,\mathcal{F}_s) = 0$.

5.1 Example (Brownian motion). A Brownian motion W is a martingale relative to its natural filtration. This follows since, the conditional expectation $\mathrm{E}(W_t - W_s|\,\mathcal{F}_s)$ is equal to the unconditional expectation $\mathrm{E}(W_t - W_s)$ by the independence property (ii), and the latter is zero by property (i). □

5.2 EXERCISE. Let W be a Brownian motion and $(\mathcal{F}_t)_{t\geq 0}$ its natural filtration. Show that the process $W_t^2 - t$ is a martingale.

5.5 Generalized Brownian Motion

When working with a Brownian motion it is sometimes useful to include more information into a filtration than given by observing the Brownian sample paths. Given a filtration $(\mathcal{F}_t)_{t\geq 0}$ we replace property (ii) of a Brownian motion by the alternative property

(ii') W is adapted to $(\mathcal{F}_t)_{t\geq 0}$ and the increment $W_t - W_s$ is stochastically independent of $\mathcal{F}_s$, for any $0 \leq s < t$.

By the requirement that W be adapted, the filtration $(\mathcal{F}_t)_{t\geq 0}$ is necessarily larger than the natural filtration of W. Therefore, the property (ii') requires more than the corresponding (ii).

It can be checked that a "generalized Brownian motion" is still a martingale.

5.6 Variation

Brownian motion has strange sample paths. They are continuous by assumption, but they are not differentiable. We can see this by studying the variation of the sample paths.

Let $0 = t_0^n < t_1^n < \cdots < t_{k_n}^n$ be a sequence of partitions of partitions of a given interval $[0, t]$ such that the meshwidth $\max_i (t_i^n - t_{i-1}^n)$ tends to zero as $n \to \infty$. Then for a continuously differentiable function $f\colon [0, t] \to \mathbb{R}$ we have, as $n \to \infty$,

$$\sum_{i=1}^{k_n} \bigl|f(t_i^n) - f(t_{i-1}^n)\bigr| \approx \sum_{i=1}^{k_n} \bigl|f'(t_{i-1}^n)\bigr|\bigl|t_i^n - t_{i-1}^n\bigr| \to \int_0^t |f'(s)|\, ds.$$

The left side of this equation is called the *variation* of f over the given partition. The approximation can be shown to be correct in the sense that the variation indeed converges to the integral on the right as the meshwidth of the partition tends to zero. We conclude that the variation of a continuously differentiable function is bounded if the meshwidth of the partition decreases to zero. As a consequence the *quadratic variation*

$$\sum_{i=1}^{k_n} \bigl|f(t_i^n) - f(t_{i-1}^n)\bigr|^2 \leq \max_i \bigl|f(t_i^n) - f(t_{i-1}^n)\bigr| \sum_{i=1}^{k_n} \bigl|f(t_i^n) - f(t_{i-1}^n)\bigr|$$

tends to zero as $n \to \infty$, since the maximum of the increments tends to zero by the continuity of f and the variation is bounded.

The sample paths of Brownian motion do not possess this property. In fact, the quadratic variation rather than the variation of the sample paths of Brownian motion tends to a nontrivial limit.

This is true in a stochastic sense. It will be convenient to use the notation $\stackrel{L_2}{\to}$ for "convergence in second mean": a sequence of random variables X_n is said to *converge in second mean* or *converge in L_2* to a random variable X, notation $X_n \stackrel{L_2}{\to} X$, if

$$\mathrm{E}(X_n - X)^2 \to 0.$$

Because the second moment of a random variable is the sum of its variance and the square of its expectation, convergence in L_2 is equivalent to $\mathrm{E}X_n \to \mathrm{E}X$ and $\mathrm{var}(X_n - X) \to 0$.

5.3 Lemma. *For any sequence of partitions of partitions* $0 = t_0^n < t_1^n < \cdots < t_{k_n}^n$ *of the interval* $[0,t]$ *with* $\max_i(t_i^n - t_{i-1}^n) \to 0$ *we have*

$$\sum_{i=1}^{k_n} \bigl|W_{t_i^n} - W_{t_{i-1}^n}\bigr|^2 \overset{L_2}{\to} t.$$

Proof. The increments $W_{t_i^n} - W_{t_{i-1}^n}$ of Brownian motion over the partition are independent random variables with $N(0, t_i^n - t_{i-1}^n)$-distributions. Therefore,

$$\begin{aligned}
\mathrm{E}\sum_{i=1}^{k_n} \bigl|W_{t_i^n} - W_{t_{i-1}^n}\bigr|^2 &= \sum_{i=1}^{k_n} (t_i^n - t_{i-1}^n) = t,\\
\mathrm{var}\sum_{i=1}^{k_n} \bigl|W_{t_i^n} - W_{t_{i-1}^n}\bigr|^2 &= \sum_{i=1}^{k_n} \mathrm{var}\bigl(|W_{t_i^n} - W_{t_{i-1}^n}|^2\bigr)\\
&= \sum_{i=1}^{k_n} 2(t_i^n - t_{i-1}^n)^2.
\end{aligned}$$

Here we use that the variance of the square Z^2 of a $N(0,\sigma^2)$-distributed random variable (i.e. σ^2 times a χ_1^2-variable) is equal to $2\sigma^4$. The expression in the last line of the display goes to 0, because it is the quadratic variation of the identity function. ∎

Because the quadratic variation of a continuously differentiable function tends to zero as the meshwidth of the partition tends to zero, the sample paths of Brownian motion cannot be continuously differentiable. Otherwise the limit in the lemma would have been 0, rather than t. This expresses that the sample paths of Brownian motion possess a certain roughness. This is nice if we want to use Brownian motion as a model for irregular processes, such as the Brownian motion of particles in a fluid or gas, or a financial process, but it complicates the use of ordinary calculus in connection to Brownian motion.

For instance, if W were the price of an asset and we would have $\phi_{t_{i-1}}$ assets in a our portfolio during the time interval $(t_{i-1}, t_i]$, then our increase in wealth due to changes in value of the asset during the full interval $(0,t]$ would be

$$\sum_i \phi_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}) = \sum_i \phi_{t_{i-1}} \Delta W_{t_i}.$$

We would like to extend this to portfolios ϕ_t that change continuously in time, and thus would like to be able to talk about something like

$$\int \phi_t \, dW_t.$$

For a continuously differentiable function f we would interpret $\int \phi(t)\,df(t)$ as $\int \phi(t)\,f'(t)\,dt$. Because the sample paths of Brownian motion do not have derivatives, we cannot define $\int \phi_t\,dW_t$ in this way. "Stochastic integrals" provide a way around this.

5.7 Stochastic Integrals

Let W be a Brownian motion relative to a given filtration $(\mathcal{F}_t)_{t\geq 0}$. We define an integral $\int_0^t X_s\,dW_s$ for given stochastic processes X defined on the same probability space as W in steps:

(a) If $X_t = 1_{(u,v]}(t)A$ for a random variable $A \in \mathcal{F}_u$, then $\int X_s\,dW_s$ is the random variable $\bigl(W_v - W_u\bigr)A$.

(b) If $X = \sum_i X^{(i)}$, then $\int X_s\,dW_s = \sum_i \int X_s^{(i)}\,dW_s$.

(c) If $\mathrm{E}\int (X_s^{(n)} - X_s)^2\,ds \to 0$ for some sequence $X^{(n)}$, then $\int X_s\,dW_s$ is the L_2-limit of the sequence $\int X_s^{(n)}\,dW_s$.

(d) $\int_0^t X_s\,dW_s = \int (1_{(0,t]}X_s)\,dW_s$.

The following lemma shows that the integral $\int_0^t X_s\,dW_s$ can be defined by this procedure for any adapted process X with $\mathrm{E}\int_0^t X_s^2\,ds < \infty$. If this integral is finite for all $t > 0$, then we obtain a stochastic process denoted by $X \cdot W$ and given by $X \cdot W = (\int_0^t X_s\,dW_s : t \geq 0)$.

5.4 Theorem. *Let X be adapted and satisfy $\mathrm{E}\int_0^t X_s^2\,ds < \infty$ for every $t \geq 0$. Then $\int_0^t X_s\,dW_s$ can be defined through steps (a)–(d) and*

(i) $\mathrm{E}\int_0^t X_s\,dW_s = 0$.

(ii) $\mathrm{E}\bigl(\int_0^t X_s\,dW_s\bigr)^2 = \mathrm{E}\int_0^t X_s^2\,ds$.

(iii) the process $(\int_0^t X_s\,dW_s : t \geq 0)$ *is a martingale.*

A sketch of the proof of the theorem is given in Section 5.15. The interested reader is also referred to e.g. the book by Chung and Williams. The intuition behind assertion (iii) is that the increments $X_{s_i}\Delta W_{s_i}$ of the integral $\int_0^t X_s\,dW_s$ satisfy

$$\mathrm{E}\bigl(X_{s_i}\Delta W_{s_i} \,|\, \mathcal{F}_{s_i}\bigr) = X_{s_i}\mathrm{E}(\Delta W_{s_i} \,|\, \mathcal{F}_{s_i}) = 0,$$

because the increments of Brownian motion have mean zero and are independent of the past. This reasoning is insightful, perhaps more so than the proof of the lemma. As a mathematical justification it is wrong, because the integral is a much more complicated object than a sum of (infinitesimal) increments.

5.8 Geometric Brownian Motion

The state W_t of a Brownian motion at time t is normally distributed with mean zero and hence is negative with probability 1/2. This is an embarassing property for a model of an asset price. One way out of this difficulty would be to model the asset prices as the sum $f(t) + W_t$ of a deterministic function and Brownian motion. Brownian motion with a *linear drift*, the process $\alpha + \beta t + W_t$, is a special example. This type of model uses Brownian motion as a noisy aberration of a deterministic asset price $f(t)$. If the deterministic function satisfies $f(t) \gg 0$, then the probability that $f(t) + W_t$ is negative is very small, but still positive.

Another way out is to model the asset price as a *geometric Brownian motion*, which is given by

$$e^{\sigma W_t + \alpha + \beta t}.$$

Putting the process in the exponential certainly keeps it positive.

5.9 Stochastic Differential Equations

A more general approach to modelling using Brownian motion is in terms of differential equations. An asset price process could be postulated to satisfy, for given stochastic processes μ and σ,

$$S_t = S_0 + \int_0^t \mu_s\, ds + \int_0^t \sigma_s\, dW_s. \tag{5.5}$$

This integral equation is usually written in differential form as

$$dS_t = \mu_t\, dt + \sigma_t\, dW_t. \tag{5.6}$$

In the case that $\sigma_t = 0$, this reduces to the ordinary differential equation $dS_t = \mu_t\, dt$. Adding the term $\sigma_t\, dW_t$ introduces a random perturbation of this differential equation. The infinitesimal change dS_t in S_t is equal to $\mu_t\, dt$ plus a noise term. Because the increments of Brownian motion are independent, we interpret the elements dW_t as "independent noise variables".

The integral $\int_0^t \sigma_s\, dW_s$ in (5.5) must of course be interpreted as a stochastic integral in the sense of Section 5.7, whereas the integral $\int_0^t \mu_s\, ds$ is an ordinary integral, as in calculus. The stochastic differential equation (5.6), or *SDE*, is merely another way of writing (5.5), the latter integral equation being its only mathematical interpretation. The understanding of dW_t as a random noise variable is helpful for intuition, but does not make mathematical sense.

For the integral $\int_0^t \sigma_s\, dW_s$ to be well defined, the process σ must be adapted.

In many examples the process μ_t and σ_t are defined in terms of the process S. For instance, a *diffusion equation* takes the form

$$dS_t = \mu(t, S_t)\, dt + \sigma(t, S_t)\, dW_t, \tag{5.7}$$

for given functions μ and σ on $[0,\infty) \times \mathbb{R}$. Then the stochastic differential equation is recursive and the process S_t is only implicitly defined, and in fact there is no guarantee that it exists. Just as for ordinary differential equations, existence of solutions for stochastic differential equations is an important subject of study. There are several general theorems that guarantee the existence of solutions under certain conditions, but we omit a discussion.

5.10 Markov Processes

A *Markov Process* X is a stochastic process with the property that for every $s < t$ the conditional distribution of X_t given $(X_u: u \leq s)$ is the same as the conditional distribution of X_t given X_s. In other words, given the "present" X_s the "past" $(X_u: u \leq s)$ gives no additional information about the "future" X_t.

5.8 Example (Brownian Motion). Because $W_t = W_t - W_s + W_s$ and $W_t - W_s$ is normal $N(0, t-s)$ distributed and independent of $(W_u: u \leq s)$, the conditional distribution of W_t given $(W_u: u \leq s)$ is normal $N(W_s, t-s)$ and hence depends on W_s only. Therefore, Brownian motion is a Markov process. □

5.9 Example (Diffusions). A diffusion process S, as given by the SDE (5.7), does not possess independent increments as Brownian motion. However, in an infinitesimal sense the increments dS_t depend only on S_t and the infinitesimal increment dW_t, which is independent of the past. This intuitive understanding of the evolution suggests that a diffusion process may be Markovian. This is indeed the case, under some technical conditions. □

5.11 Quadratic variation - revisited

In Section 5.6 we have seen that the quadratic variation of Brownian motion converges to a limit as the meshwidth of the partitions tends to zero. This is true for general solutions to SDEs, except that in general the convergence must be interpreted "in probability". We say that a sequence of random variables X_n *converges in probability* to a random variable X if

$$\mathbb{P}\big(|X_n - X| > \varepsilon\big) \to 0, \qquad \text{for every } \varepsilon > 0.$$

This is denoted by $X_n \xrightarrow{\text{P}} X$.

5.10 Lemma. *Consider a stochastic process S that satisfies the SDE (5.6) for adapted processes μ and σ, then for any sequence of partitions $0 = t_0^n < t_1^n < \cdots < t_{k_n}^n$ of the interval $[0,t]$ with $\max_i(t_i^n - t_{i-1}^n) \to 0$ we have*

$$\sum_{i=1}^{k_n} \big|S_{t_i^n} - S_{t_{i-1}^n}\big|^2 \xrightarrow{\text{P}} \int_0^t \sigma_s^2\, ds.$$

Proof. (Sketch.) The increments can be written

$$S_{t_i^n} - S_{t_{i-1}^n} = \int_{t_{i-1}^n}^{t_i^n} \mu_s\, ds + \int_{t_{i-1}^n}^{t_i^n} \sigma_s\, dW_s.$$

The first term on the right is an ordinary integral and gives no contribution to the quadratic variation, since the sum of its squares is of the order $\sum_i (t_i^n - t_{i-1}^n)^2 \to 0$. The second term is approximately $\sigma_{t_{i-1}^n}(W_{t_i^n} - W_{t_{i-1}^n})$. Now

$$\mathrm{E}\Big(\sum_{i=1}^{k_n} \sigma_{t_{i-1}^n}^2 \big((W_{t_i^n} - W_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)\big)\Big) = 0,$$

$$\begin{aligned}
\mathrm{var}\Big(&\sum_{i=1}^{k_n} \sigma_{t_{i-1}^n}^2 \big((W_{t_i^n} - W_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)\big)\Big)\\
&= \sum_{i=1}^{k_n} \mathrm{var}\Big(\sigma_{t_{i-1}^n}^2 \big((W_{t_i^n} - W_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)\big)\Big)\\
&= \sum_{i=1}^{k_n} \mathrm{E}\sigma_{t_{i-1}^n}^4\, \mathrm{E}\big((W_{t_i^n} - W_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)\big)^2\\
&= \sum_{i=1}^{k_n} \mathrm{E}\sigma_{t_{i-1}^n}^4 (t_i^n - t_{i-1}^n)^2.
\end{aligned}$$

Here we used the independence of the increments of Brownian motion from the past several times, for instance to see that the terms of the sum are uncorrelated. Together these equations suggest that

$$\mathrm{E}\Big(\sum_{i=1}^{k_n} \sigma_{t_{i-1}^n}^2 \big((W_{t_i^n} - W_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)\big)\Big)^2 \to 0.$$

Combined with the convergence

$$\sum_{i=1}^{k_n} \sigma^2_{t^n_{i-1}} (t^n_i - t^n_{i-1}) \to \int_0^t \sigma_s^2 \, ds$$

this would give the result.

This proof can be made precise without much difficulty if the process σ is bounded and left-continuous. For a complete proof we need to use a truncation argument involving stopping times. ∎

The limit of the sums of squares $\sum_{i=1}^{k_n} \bigl|S_{t^n_i} - S_{t^n_{i-1}}\bigr|^2$ is called the *quadratic variation* of the process S. It is denoted by $[S]_t$, and also known as the "square bracket process". For a solution S to the SDE (5.6) we have $[S]_t = \int_0^t \sigma_s^2 \, ds$.

Besides the quadratic variation of a single process, there is also a cross quadratic variation of a pair of processes R and S, defined as the limit (in probability)

$$[R, S]_t = \lim_{n\to\infty} \sum_{i=1}^{k_n} \bigl(R_{t^n_i} - R_{t^n_{i-1}}\bigr)\bigl(S_{t^n_i} - S_{t^n_{i-1}}\bigr).$$

5.11 EXERCISE. Suppose that the processes R and S both satisfy an SDE (5.6), but with different functions μ and σ. Guess $[R, S]$ if
(i) R and S depend in (5.6) on the same Brownian motion.
(ii) the SDEs for R and S are driven by independent Brownian motions.

5.12 Itô Formula

The geometric Brownian motion is actually a special case of the SDE approach. By a celebrated formula of Itô it can be shown that geometric Brownian motion satisfies a SDE.

Itô's formula is a chain rule for stochastic processes, but due to the special nature of stochastic integrals it takes a surprising form. The version of Itô's formula we present here says that a transformation $f(S_t)$ of a process that satisfies an SDE by a smooth function f again satisfies an SDE, and gives an explicit expression for it.

Recall that for a stochastic process S as in (5.6), the quadratic variation is the process $[S]$ such that $d[S]_t = \sigma_t^2 \, dt$.

5.12 Theorem (Itô's formula). *If the stochastic process S satisfies the SDE (5.6) and $f: \mathbb{R} \to \mathbb{R}$ is twice continuouly differentiable, then*

$$df(S_t) = f'(S_t)\,dS_t + \tfrac{1}{2} f''(S_t)\,d[S]_t.$$

Proof. (Sketch.) For a sequence of sufficiently fine partitions $0 = t_0^n < t_1^n < \cdots < t_{k_n}^n = t$ of the interval $[0, t]$ with $\max_i (t_i^n - t_{i-1}^n) \to 0$ we have

$$\begin{aligned} f(S_t) - f(S_0) &= \sum_{i=1}^{k_n} \bigl(f(S_{t_i^n}) - f(S_{t_{i-1}^n})\bigr) \\ &\approx \sum_{i=1}^{k_n} f'(S_{t_{i-1}^n})(S_{t_i^n}) - S_{t_{i-1}^n}) + \tfrac{1}{2} \sum_{i=1}^{k_n} f''(S_{t_{i-1}^n})(S_{t_i^n} - S_{t_{i-1}^n})^2. \end{aligned}$$

The first term on the far right tends to the stochastic integral $\int_0^t f'(S_s)\,dS_s$. By the same arguments as used in Section 5.11 the second sum tends to $\frac{1}{2} \int_0^t f''(S_s) \sigma_s^2\,ds$. ■

We have written Itô's formula in differential form, but as usually it should be mathematically interpreted as a statement about integrals.

The striking aspect of Itô's formula is the second term $\frac{1}{2} f''(S_t)\,d[S]_t$, which would not appear if the sample path $t \mapsto S_t$ were a differentiable function. As the proof shows it does appear, because the variation of the sample paths of S is not finite, whereas the quadratic variation tends to a nontrivial limit.

5.13 Example. Brownian motion W itself certainly satisfies a stochastic differential equation: the trivial one $dW_t = dW_t$.

Applied with the function $f(x) = x^2$ Itô's formula gives $dW_t^2 = 2W_t\,dW_t + \frac{1}{2} 2\,dt$, because $[W]_t = t$. We conclude that $W_t^2 = 2\int_0^t W_s\,dW_s + t$. Compare this to the formula $f^2(t) = 2\int_0^t f(s)\,df(s)$ for a continuously differentiable function f with $f(0) = 0$. □

5.14 Example (Geometric Brownian motion). As a consequence of Itô's formula, the geometric Brownian motion $S_t = \exp(\sigma W_t + \alpha + \beta t)$ satisfies the SDE

$$dS_t = (\beta + \tfrac{1}{2}\sigma^2) S_t\,dt + \sigma S_t\,dW_t.$$

To see this, apply Itô's formula with the process $X_t = \alpha + \beta t + \sigma W_t$ and the function $f(x) = \exp(x)$. □

Itô's theorem is also valid for functions of more than one process. For instance, consider a process $f(t, S_t)$ for a function $f: [0, \infty) \times \mathbb{R} \to \mathbb{R}$ of two arguments. Write f_t, f_s and f_{ss} for the partial derivatives $\partial/\partial t f(t, s)$, $\partial/\partial s f(t, s)$ and $\partial^2/\partial s^2 f(t, s)$, respectively.

5.15 Theorem (Itô's formula). *If the stochastic process S satisfies the SDE (5.6) and $f\colon [0,\infty)\times\mathbb{R}\to\mathbb{R}$ is twice continuouly differentiable, then*

$$df(t,S_t) = f_t(t,S_t)\,dt + f_s(t,S_t)\,dS_t + \tfrac{1}{2}f_{ss}(t,S_t)\,d[S]_t.$$

As a second example consider a process $f(R_t,S_t)$ of two stochastic processes R and S. If an index r or s denotes partial differentiation with respect to r or s, then we obtain the following formula.

5.16 Theorem (Itô's formula). *If the stochastic processes R and S satisfy the SDE (5.6) and $f\colon\mathbb{R}^2\to\mathbb{R}$ is twice continuouly differentiable, then*

$$\begin{aligned}df(R_t,S_t) &= f_r(R_t,S_t)\,dR_t + f_s(R_t,S_t)\,dS_t + \tfrac{1}{2}f_{rr}(R_t,S_t)\,d[R]_t\\ &\quad + \tfrac{1}{2}f_{ss}(t,S_t)\,d[S]_t + f_{rs}(R_t,S_t)\,d[R,S]_t.\end{aligned}$$

5.13 Girsanov's Theorem

The stochastic integral $\int_0^t X_s\,dW_s$ of an adapted process relative to Brownian motion is a (local) martingale. Thus the solution S to the SDE (5.6) is the sum of a local martingale $\int_0^t \sigma_s\,dW_s$ and the process $A_t = \int_0^t \mu_s\,ds$. The sample paths of the process A are the primitive functions, in the sense of ordinary calculus, of the sample paths of the process μ, and are therefore differentiable. They are referred to as "drift functions". The presence of a drift function destroys the martingale property: a solution of an SDE can be a martingale only if the drift is zero.

5.17 Lemma. *The process S defined by $S_t = \int_0^t \mu_s\,ds + \int_0^t \sigma_s\,dW_s$ is a local martingale if and only if $\mu = 0$.*

Proof. (Sketch.) If S is a local martingale, then so is the process $A_t = \int_0^t \mu_s\,ds$, because the process $\int_0^t \sigma_s\,dW_s$ is a local martingale and the difference of two local martingales is a local martingale. Because the sample paths of A are differentiable functions, the rules of ordinary calculus apply, and yield that $d(A_t^2) = 2A_t\,dA_t$, or

$$A_t^2 = \int_0^t 2A_s\,dA_s.$$

The local martingale property of A carries over to every process of the form $\int_0^t X_s\,dA_s$ for an adapted process X. This can be proved by first considering simple adapted processes, and next limits, along the same lines as the martingale property of stochastic integrals was proved. In particular, we may choose $X = 2A$, and we see that the process A^2 is a local martingale. If it is a martingale with finite second moments, then we can conclude that $\mathrm{E}A_t^2 = \mathrm{E}A_0^2 = 0$, whence $A = 0$. The general case can be handled by a stopping time argument. ∎

The martingale property refers to the underlying probability distribution on the outcome space. Therefore a process may well be a martingale relative to a probability measure $\mathbb{Q}$, whereas it is not a martingale if the outcome space is equipped with another probability measure $\mathbb{P}$. If the process is given by an SDE under $\mathbb{P}$, then this somehow means that the drift of the process can be made to "disappear" by changing the probability distribution on the space of outcomes. This observation turns out to be crucial in the pricing theory.

It will be sufficient to consider this for the case that $\sigma = 1$, i.e. $S_t = W_t + \int_0^t \mu_s\,ds$. If W is a Brownian motion under the probability measure $\mathbb{P}$, then W is a martingale under $\mathbb{P}$ and hence S cannot be a martingale under $\mathbb{P}$, (unless $\mu = 0$). Girsanov's theorem shows that for "most" processes μ, there exists another probability measure $\mathbb{Q}$ such that S is a martingale, and even a Brownian motion, under $\mathbb{Q}$.

5.18 Theorem (Girsanov). *If $(W_t: 0 \leq t \leq T)$ is a Brownian motion under the probability measure $\mathbb{P}$ and μ is an adapted process with $\mathrm{E}\exp(\frac{1}{2}\int_0^T \mu_s^2\,ds) < \infty$, then there exists a probability measure $\mathbb{Q}$ such that the process $(W_t + \int_0^t \mu_s\,ds: 0 \leq t \leq T)$ is a Brownian motion under $\mathbb{Q}$.*

There is even a constructive formula for finding the "martingale measure" $\mathbb{Q}$ from $\mathbb{P}$, given by

$$\mathbb{Q}(A) = \mathrm{E}\Big(1_A e^{-\int_0^T \mu_s\,dW_s - \frac{1}{2}\int_0^T \mu_s^2\,ds}\Big),$$

where the expectation on the right is computed under the probability measure $\mathbb{P}$. The condition $\mathrm{E}\exp(\frac{1}{2}\int_0^T \mu_s^2\,ds) < \infty$ ensures that the formula in the preceding display indeed defines a probability measure $\mathbb{Q}$. If the process μ is bounded (e.g. constant), then the condition is clearly satisfied. In general, the condition says that μ "should not grow too big".

5.14 Brownian Representation

Let $(\mathcal{F}_t)_{t\geq 0}$ be the natural filtration of a given Brownian motion W. Stochastic processes defined on the same outcome space that are martingales relative to this "Brownian filtration" are referred to as *Brownian* *martingales*. Brownian motion itself is an example, and so are all stochastic integrals $X \cdot B$ for adapted processes X.

The following theorem shows that these are the only Brownian martingales.

5.19 Theorem. *Let $\{\mathcal{F}_t\}$ be the (completion of the) natural filtration of a Brownian motion process W. If M is a (cadlag) local martingale relative to $\{\mathcal{F}_t\}$, then there exists a predictable process X with $\int_0^t X_s^2\,ds < \infty$ almost surely for every $t \geq 0$ such that $M_t = M_0 + \int_0^t X_s\,dW_s$.*

This Brownian representation theorem remains true if the filtration is generated by multiple, independent Brownian motion processes $W^{(1)}, W^{(2)}, \ldots, W^{(d)}$. Then an arbitrary (cadlag) local martingale can be written as $M_t = M_0 + \sum_{i=1}^d \int_0^t X_s^{(i)}\,dW_s^{(i)}$.

* 5.15 Proof of Theorem 5.4

In this section we provide details for the construction of the stochastic integral in Section 5.7. Because this material is mathematically quite involved, we do not give a full proof of Theorem 5.4, but we indicate the most essential steps.

A process of the type $X_t = 1_{(u,v]}(t)A$ for a random variable $A \in \mathcal{F}_u$ as in (a) is adapted and hence so is process of the type

$$X_t = \sum_i 1_{(u_i,v_i]}(t)A_{u_i}, \qquad u_i < v_i, A_{u_i} \in \mathcal{F}_{u_i}.$$

A process of this type is called *simple adapted*. By splitting up sets if necessary, it is always possible to represent such a simple adapted process with disjoint intervals $(u_i, v_i]$. For X as in the preceding display we define

$$\int X_s\,dW_s = \sum_i A_{u_i}(W_{v_i} - W_{u_i}). \tag{5.20}$$

Because the representation of X in terms of the intervals $(u_i, v_i]$ and A_{u_i} is not unique (we could for instance split up the intervals further), it must be verified that this definition is consistent, but we omit this part of the proof.

We next verify property (ii) for simple adapted processes X and $t = \infty$. If, as we assume, the intervals $(u_i, v_i]$ are disjoint, then $X_t^2 = \sum_i 1_{(u_i,v_i]}(t)A_{u_i}^2$. Therefore, the right side of (ii) with $t = \infty$ is equal to

$$\mathrm{E}\int X_s^2\,ds = \mathrm{E}\int \sum_i 1_{(u_i,v_i]}(t)A_{u_i}^2\,ds = \sum_i \mathrm{E}A_{u_i}^2(v_i - u_i).$$

The left side of (ii) with $t = \infty$ is given by, in view of (5.20),

$$\mathrm{E}\Bigl(\int X_s\,dW_s\Bigr)^2 = \mathrm{E}\sum_i\sum_j A_{u_i}A_{u_j}(W_{v_i}-W_{u_i})(W_{v_j}-W_{u_j}).$$

Because the intervals $(u_i, v_i]$ are disjoint, we have that $\mathrm{E}(W_{v_i}-W_{u_i})(W_{v_j}-W_{u_j}) = 0$ for $i \neq j$, by the independence and the zero means of the increments of Brownian motion. It follows that the diagonal terms in the double sum vanish, whence the preceding display is equal to

$$\mathrm{E}\sum_i A_{u_i}^2(W_{v_i}-W_{u_i})^2 = \sum_i \mathrm{E}A_{u_i}^2\mathrm{E}(W_{v_i}-W_{u_i})^2 = \sum_i A_{u_i}^2(v_i-u_i),$$

where in the second step we use the independence of the increment $W_{v_i} - W_{u_i}$ of $\mathcal{F}_{u_i}$ and hence of A_{u_i}. Thus we have verified that for simple adapted processes X

$$\mathrm{E}\Bigl(\int X_s\,dW_s\Bigr)^2 = \mathrm{E}\int X_s^2\,ds.$$

In words we have shown that the integral is a "linear isometry". A *linear isometry* between normed spaces $\mathbb{X}$ and $\mathbb{Y}$ is a linear map $I\colon \mathbb{X} \to \mathbb{Y}$ such that $\|I(x)\|_{\mathbb{Y}} = \|x\|_{\mathbb{X}}$ for every $x \in \mathbb{X}$. This isometry is the basis for the extension of the integral to general adapted processes, by way of the following result from analysis.

Any linear isometry $I\colon \mathbb{X}_0 \subset \mathbb{X} \to \mathbb{Y}$ from a linear subspace $\mathbb{X}_0$ of a normed space $\mathbb{X}$ into a complete normed space $\mathbb{Y}$ possesses a unique extension to an isometry defined on the closure $\bar{\mathbb{X}}_0 = \bigl\{X \in \mathbb{X}\colon \exists\{X_n\} \subset \mathbb{X}_0 \text{ with } \|X_n - X\| \to 0\bigr\}$ of $\mathbb{X}_0$ in $\mathbb{X}$.

In our situation we take the space $\mathbb{X}$ equal to all adapted processes X with $\|X\|_{\mathbb{X}}^2 = \mathrm{E}\int_0^t X_s^2\,ds < \infty$, and $\mathbb{X}_0$ equal to the collection of all simple adapted processes. We have seen that the map $I\colon X \mapsto \int X_s\,dW_s$ is an isometry into the set $\mathbb{Y}$ of random variables with finite second moments, with $\|Y\|_{\mathbb{Y}}^2 = \mathrm{E}Y^2$. Thus the integral can be extended to the closure of the set of simple adapted processes. That this closure is the set of all adapted processes with $\mathrm{E}\int_0^t X_s^2\,ds < \infty$ can be shown by approximation by step functions. We omit the details of this part of the proof.

Thus the integral is defined. The verification of its properties (i)-(iii) proceeds by first verifying that these assertions hold on the set $\mathbb{X}_0$ of simple processes and next showing that these properties are preserved under taking limits.

For a simple adapted process of the form $X_t = 1_{(u,v]}(t)A$ with $\sigma(A) \subset \mathcal{F}_u$ and $s < t$ we have

$$\int_0^t X_r\,dW_r - \int_0^s X_r\,dW_r = \begin{cases} A(W_{t\wedge v} - W_{s\vee u}), & \text{if } t\wedge v > s \vee u,\\ 0, & \text{otherwise.}\end{cases}$$

Because A is known at time $u \leq s \vee u$ and Brownian motion is a martingale we have $\mathrm{E}(A(W_{t\wedge v} - W_{s\vee u})|\,\mathcal{F}_{s\vee u}) = A\mathrm{E}(W_{t\wedge v} - W_{s\vee u}|\,\mathcal{F}_{s\vee u}) = 0$. Therefore, with the help of the tower property of conditional expectation, it follows that

$$\mathrm{E}\Big(\int_0^t X_r\, dW_r - \int_0^s X_r\, dW_r|\,\mathcal{F}_s\Big) = 0.$$

Thus the stochastic integral $\int_0^t X_s\, dW_s$ is a martingale for X of this form. Because the sum of two martingales is again a martingale, this conclusion extends to all simple adapted processes. Because the martingale property is preserved under taking L_2-limits, it next extends to the stochastic integral in general.

* 5.16 Stopping

Stopping times are intuitively meaningful objects that have interest on their own, and are also essential for extension of the definition of stochastic integrals, as given in the next section. However, we shall not need the material in this and the following section in later chapters.

A *stopping time* relative to a filtration $(\mathcal{F}_t)_{t\geq 0}$ is a random variable T with values in $[0,\infty]$ such that $\{T \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$. A stopping time formalizes a strategy to play (or invest) on the market until a given time, which need not be predetermined, but may be based on observing the market. The requirement that the event $\{T \leq t\}$ is known at time t says that the decision to stop trading must be made based on information collected in past and present. If the filtration is generated by a process X, then this requirement implies that the decision to stop at time t must be based on the sample path of X until t.

5.21 Example (Hitting time). If X is an adapted process with continuous sample paths, then $T = \inf\{t \geq 0 \colon X_t \in B\}$ is a stopping time for every (Borel) set B. This is known as the *hitting time* of B. □

Stopping times are important tools in the theory of stochastic processes, but are also crucial to evaluate *American options.* These are contracts that give the holder the right to collect a certain payment at a time t in a given interval $[0,T]$ of his own choosing. The amount of the payment depends on the history of an asset price up to the time of payment. For instance, an American call option on an asset with price process S gives the right to buy the asset at a predetermined price K at any time t in an interval $[0,T]$. This corresponds to a payment of $(S_t - K)^+$ at the chosen time t. The financial problem is to determine an optimal stopping time for the payment, and to evaluate the value of the resulting contract.

Given a stopping time T and a stochastic process the *stopped process* X^T is defined as the stochastic process such that

$$(X^T)_t = X_{T\wedge t}.$$

The sample paths of the stopped process are identical to the sample paths X up to time T and take the constant value X_T for $t \geq T$.

5.22 Theorem. *If X is a martingale, then so is X^T.*

More explicitly, the theorem says that, if X is a martingale, then

$$\mathrm{E}(X_{T\wedge t}|\,\mathcal{F}_s) = X_{T\wedge s}, \qquad s < t.$$

In particular, we have $\mathrm{E}X_{T\wedge t} = \mathrm{E}X_{T\wedge s} = \mathrm{E}X_0$, because we can choose $s = 0$. If T is a bounded stopping time, then we may choose $t \geq T$ and we find

$$\mathrm{E}X_T = \mathrm{E}X_0.$$

This says that stopping does not help if the pay-off process X is a martingale. No matter how clever the stopping strategy T, the expected pay-off $\mathrm{E}X_T$ is $\mathrm{E}X_0$.

5.23 Example. The process $W_t^2 - t$ is a martingale. It can be shown that $T = \inf\{t \geq 0\colon |W_t| = a\}$ is finite almost surely, whence $W_T^2 = a^2$. The identity $\mathrm{E}(W_T^2 - T) = \mathrm{E}(W_0^2 - 0) = 0$ reduces to $\mathrm{E}T = a^2$.

However, it is not permitted to apply this identity directly, as T is not a bounded stopping time. A way around this is to apply the identity with $T \wedge n$ for a given n and next take limits as $n \to \infty$. Because $T \wedge n$ is bounded we find $\mathrm{E}W_{T\wedge n}^2 = \mathrm{E}(T \wedge n)$. Because $W_{T\wedge n}^2 \leq a^2$, we have $\mathrm{E}W_{T\wedge n}^2 \to \mathrm{E}W_T^2 = a$ as $n \to \infty$ by the dominated convergence theorem. Also we have $\mathrm{E}T \wedge n \uparrow \mathrm{E}T$ by the monotone convergence theorem. Thus the formula $\mathrm{E}T = a^2$ is correct. □

5.24 EXERCISE. For given $a > 0$, let $T = \inf\{t \geq 0\colon W_t = a\}$.

(i) Show that $Y_t = e^{\theta W_t - \frac{1}{2}\theta^2 t}$ is a martingale, for every $\theta \in \mathbb{R}$.
(ii) Show that $\mathrm{E}\exp(-\theta T) = \exp(-\sqrt{2\theta}a)$.
(iii) Show that $\mathrm{E}T = \infty$.

* 5.17 Extended Stochastic Integrals

Using stopping times we can define a useful extension of the definition of the stochastic integral. We have already defined the stochastic integral $\int_0^t X_s\,dW_s$ for any adapted process X with $\mathrm{E}\int_0^t X_s^2\,ds < \infty$. We shall now extend this to all adapted processes X with

$$\int X_s^2\,ds < \infty, \qquad \text{a.s..} \tag{5.25}$$

This is a larger set of adapted processes, as finiteness of the expected value of a positive random variable implies finiteness of the variable with probability one, but not the other way around.

We "truncate" a given adapted process by stopping it appropriately. For a given n we define the stopping time

$$T_n = \inf\{t \geq 0 \colon \int_0^t X_s^2\,ds \geq n\}.$$

The finiteness (5.25) of the (nondecreasing) process $\int_0^t X_s^2\,ds$ implies that $T_n \uparrow \infty$ as $n \to \infty$. From the definition of T_n it follows immediately that $\int_0^t X_s^2\,ds \leq n$ if $t \leq T_n$. Consequently $\mathrm{E}\int_0^t (X_s 1_{s\leq T_n})^2\,ds = \mathrm{E}\int_0^{t\wedge T_n} X_s^2\,ds \leq \mathrm{E}n = n < \infty$. We can therefore define, for every n and t,

$$\int_0^t (X_s 1_{s\leq T_n})\,dW_s.$$

We define $\int_0^t X_s\,dW_s$ as the limit of these variables, in the almost sure sense, as $n \to \infty$. It can be shown that this limit indeed exists.

Each of the processes in the preceding display is a martingale. The stochastic integral $Y_t = \int_0^t X_s\,W_s$ is the limit of these martingales, but need not be a martingale itself. (The limit is only in an almost sure sense, and this is not strong enough to preserve the martingale property.). However, the stopped process Y^{T_n} is exactly the the integral $\int_0^t (X_s 1_{s\leq T_n})\,dW_s$ and hence is a martingale. This has gained the stochastic integral $\int_0^t X_s\,dW_s$ the name of being a *local martingale*.

6
Black-Scholes Model

In this chapter we assume that we can trade continuously in a (riskless) bond and some risky asset, for instance a stock. We assume that the bond-price B evolves as

$$B_t = e^{rt},$$

where r is the riskless interest rate. The price process S of the risky asset is assumed to be a geometric Brownian motion, i.e.

$$S_t = S_0 e^{\mu t + \sigma W_t}.$$

Here W is a Brownian motion, $\mu \in \mathbb{R}$ is called the *drift* of the process, and σ the *volatility*. We denote by $(\mathcal{F}_t)$ the filtration generated by the price process S. Observe that $(\mathcal{F}_t)$ is also the natural filtration of the Brownian motion W, since both processes generate the same flow of information.

For some fixed $T > 0$, let $C \in \mathcal{F}_T$ be a non-negative random variable whose value is determined by the information up till time T. We think of C as the pay-off at time T of some contingent claim. For technical reasons, we assume that $\mathrm{E}C^2 < \infty$. We want to answer the same question as in the discrete-time setup: What is the fair price of the claim C at time zero?

To answer this question we follow the same route as in Chapter 4. We first use Girsanov's theorem to change the underlying probability measure in such a way that the discounted asset price $\tilde{S}_t = e^{-rt}S_t$ becomes a martingale under the new measure $\mathbb{Q}$. Then we consider the $\mathbb{Q}$-martingale $\tilde{V}_t = \mathrm{E}_{\mathbb{Q}}(e^{-rT}C|\,\mathcal{F}_t)$ and use the representation theorem to write it as an integral of a predictable process ϕ with respect to $\tilde{S}$. This leads to the construction of a self-financing trading strategy that replicates the pay-off C. By an arbitrage argument, the value of this trading portfolio at time zero must be the fair price of the claim. As in the binomial model, the fair price will turn out to be $\mathrm{E}_{\mathbb{Q}}e^{-rT}C$, i.e. the expectation under the martingale measure of the discounted pay-off.

6.1 Portfolios

Before we can carry out the programme outlined in the preceding section we have to give a mathematically precise definition of a self-financing portfolio in the present continuous-time setting. A *portfolio* is just a pair of predictable processes (ϕ_t, ψ_t). We interpret ϕ_t as the number of risky assets held at time t, and ψ_t as the number of bonds. Predictability roughly means that to determine the positions ϕ_t and ψ_t, only the information available before time t is used. For technical reasons we assume that almost surely,

$$\int_0^T |\phi_t|^2\, dt + \int_0^T |\psi_t|\, dt < \infty.$$

With the portfolio (ϕ, ψ) we associate the *value process* V defined by

$$V_t = \phi_t S_t + \psi_t B_t.$$

For hedging strategies we need the notion of a self-financing portfolio. Such a portfolio is created using some starting capital at time zero, and after time zero the portfolio is only changed by rebalancing, i.e. by replacing bonds by the risky asset or vice versa. No additional injections or withdrawals of money are allowed. Loosely speaking, such a portfolio has the property that in an infinitesimally small time interval $[t, t+dt]$, the changes in the portfolio value are only caused by changes in the price processes S and B, and not by changes in ϕ_t and ψ_t which are due to injections or withdrawals of money. Therefore, we call a portfolio (ϕ, ψ) *self-financing* if its price process V satisfies the SDE

$$dV_t = \phi_t\, dS_t + \psi_t\, dB_t.$$

A *replicating*, or *hedging portfolio* for the claim C is a self-financing portfolio (ϕ, ψ) with a value process V which satisfies $V_T = C$. If such a portfolio exists, then an arbitrage argument shows that the "fair price" of the claim at time $t \in [0, T]$ equals the value V_t of the portfolio.

Of course, the arbitrage argument is an economic one, and not a mathematical argument. When we use the phrase "fair price" in mathematical theorems below, the "fair price" or "value" will always be understood to be defined as the value process of a replicating portfolio. (We shall be a bit careless about the still open trap that there may be more than one replicating portfolios, with different value processes.)

6.2 The Fair Price of a Derivative

Let us now derive the pricing formula for the derivative C announced in the first section.

For the discounted asset price we have $\tilde{S}_t = e^{-rt}S_t = S_0\exp((\mu - r)t + \sigma W_t)$, whence it is a geometric Brownian motion with drift $\mu - r$ and volatility σ. By Example 5.14 it satisfies the SDE

$$d\tilde{S}_t = (\mu - r + \tfrac{1}{2}\sigma^2)\tilde{S}_t\,dt + \sigma\tilde{S}_t\,dW_t.$$

If we define $\tilde{W}_t = W_t + t(\mu - r + \frac{1}{2}\sigma^2)/\sigma$ this simplifies to

$$d\tilde{S}_t = \sigma\tilde{S}_t\,d\tilde{W}_t. \tag{6.1}$$

By Girsanov's theorem, there exists a new underlying probability measure $\mathbb{Q}$ such that $\tilde{W}$ is a Brownian motion under $\mathbb{Q}$. Hence, the preceding SDE implies that the process $\tilde{S}$ is a $\mathbb{Q}$-martingale.

Now consider the process $\tilde{V}_t = \mathrm{E}_{\mathbb{Q}}(e^{-rT}C|\,\mathcal{F}_t)$. By the tower property of conditional expectations this is a $\mathbb{Q}$-martingale relative to the filtration $(\mathcal{F}_t)$. It is obvious that the natural filtration $(\mathcal{F}_t)$ of W is also the natural filtration of the process $\tilde{W}$ (the processes generate the same flow of information). Hence, by the Brownian representation theorem, Theorem 5.19, there exists a predictable process $\tilde{\phi}$ such that $d\tilde{V}_t = \tilde{\phi}_t\,d\tilde{W}_t$. So if we define $\phi_t = \tilde{\phi}_t/\sigma\tilde{S}_t$, we obtain

$$d\tilde{V}_t = \phi_t\,d\tilde{S}_t. \tag{6.2}$$

Next we define the process $\psi_t = \tilde{V}_t - \phi_t\tilde{S}_t$.

We claim that (ϕ, ψ) is a hedging portfolio for the derivative C. To prove this, consider the value process V of the portfolio (ϕ, ψ). Then by construction, we have

$$V_t = e^{rt}\tilde{V}_t.$$

In particular $V_T = \mathrm{E}_{\mathbb{Q}}(C|\,\mathcal{F}_T) = C$, so indeed the portfolio has the value C at time T. To prove that it is self-financing, the result of the following exercise is useful.

6.3 EXERCISE. Use Itô's formula to show that if X satisfies an SDE and F is a differentiable function, then $d(F(t)X_t) = F(t)\,dX_t + X_t\,dF(t)$.

Now we can compute dV_t. By the result of the exercise we have

$$dV_t = d(e^{rt}\tilde{V}_t) = \tilde{V}_t\,de^{rt} + e^{rt}\,d\tilde{V}_t.$$

If we use the definition of ψ_t to rewrite the first term on the right-hand side, use (6.2) to rewrite the second term and recall that $e^{rt} = B_t$, we find that

$$dV_t = (\psi_t + \phi_t\tilde{S}_t)\,dB_t + \phi_t B_t\,d\tilde{S}_t = \phi_t(\tilde{S}_t\,dB_t + B_t\,d\tilde{S}_t) + \psi_t\,dB_t.$$

So by the result of the exercise again, we indeed have the relation

$$dV_t = \phi_t\,dS_t + \psi_t\,dB_t,$$

which shows that the portfolio (ϕ, ψ) is self-financing.

In view of the standard arbitrage argument the fair price of the claim C at time t is given by $V_t = e^{rt}\tilde{V}_t = e^{rt}\mathrm{E}_{\mathbb{Q}}(e^{-rT}C|\,\mathcal{F}_t)$. Hence, we have proved the following theorem.

6.4 Theorem. *The value of the claim $C \in \mathcal{F}_T$ at time $t \in [0, T]$ is given by $e^{rt}\mathrm{E}_{\mathbb{Q}}(e^{-rT}C|\,\mathcal{F}_t)$, where $\mathbb{Q}$ is the measure under which the discounted price process $e^{-rt}S_t$ is a martingale. In particular, the price at time $t = 0$ is given by $\mathrm{E}_{\mathbb{Q}}e^{-rT}C$.*

6.3 European Options

If the claim C is *European*, meaning that it is of the form $C = f(S_T)$ for some function f, then we can derive a more explicit formula for its fair price.

Recall that under the martingale measure $\mathbb{Q}$ we have that $d\tilde{S}_t = \sigma\tilde{S}_t\,d\tilde{W}_t$, where $\tilde{W}$ is a $\mathbb{Q}$-Brownian motion. By the preceding exercise it holds that $d\tilde{S}_t = e^{-rt}\,dS_t + S_t\,de^{-rt}$, and it follows that

$$dS_t = rS_t\,dt + \sigma S_t\,d\tilde{W}_t.$$

This is the SDE of a geometric Brownian motion. By Example 5.14 it holds that

$$S_t = S_0 e^{(r-\frac{1}{2}\sigma^2)t+\sigma\tilde{W}_t}.$$

So under $\mathbb{Q}$ the asset price S is also a geometric Brownian motion, with drift $r - \sigma^2/2$ and volatility σ. In particular we have, under $\mathbb{Q}$,

$$S_T = S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z},$$

where Z is a standard Gaussian random variable. It follows that for the price $\mathrm{E}_{\mathbb{Q}}e^{-rT}C$ of the claim we have the expression

$$\begin{aligned}\mathrm{E}_{\mathbb{Q}}e^{-rT}f(S_T) &= e^{-rT}\mathrm{E}f\Big(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z}\Big)\\ &= \frac{e^{-rT}}{\sqrt{2\pi}}\int_{\mathbb{R}} f\Big(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}z}\Big)e^{-\frac{1}{2}z^2}\,dz.\end{aligned}$$

Thus, we have proved the following theorem.

6.5 Theorem. *The fair price of a European claim with pay-off* $C = f(S_T)$ *is given by*

$$\frac{e^{-rT}}{\sqrt{2\pi}} \int_{\mathbb{R}} f\Big(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}z}\Big) e^{-\frac{1}{2}z^2}\, dz.$$

For a given choice of the function f it is typically not possible to evaluate this integral analytically and one has to resort to numerical integration. For the prices of European calls and puts however, we can derive explicit expressions. The European call option with strike K and maturity T corresponds to the function $f(x) = (x-K)^+$. For this special choice of f the preceding formula can be simplified further and yields the expression

$$S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2)$$

for the price of the option, where Φ is the distribution function of the standard Gaussian distribution and the constants are given by

$$d_1 = \frac{\log(S_0/K) + (r+\sigma^2/2)T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}.$$

This is the celebrated *Black-Scholes formula.*

6.6 EXERCISE. Use the call-put parity to derive the Black-Scholes formula for the price of a European put option.

6.4 The Black-Scholes PDE and Hedging

The derivation of the pricing formula for a European claim $C = f(S_T)$ given in Theorem 6.5 can easily be extended to a formula for the price at any time $t \le T$. From the absence of arbitrage it follows that this price equals the value V_t of the hedging portfolio at time t. It holds that $V_t = F(t, S_t)$, where the function F is given by

$$F(t,x) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{\mathbb{R}} f\Big(x e^{(r-\frac{1}{2}\sigma^2)(T-t)+\sigma z\sqrt{T-t}}\Big) e^{-\frac{1}{2}z^2}\, dz.$$

Observe that if we substitute $t = 0$, we indeed recover the result of Theorem 6.5.

The pricing function F can also be obtained as the solution of the so-called Black-Scholes partial differential equation (PDE). This provides a second method for finding the price of the claim. For a given function f this PDE can usually not be solved analytically, and one has to resort to numerical methods.

6.7 Theorem. *The value of a European claim $C = f(S_T)$ at time $t \le T$ is given by $V_t = F(t, S_t)$, where F is the solution of the partial differential equation*

$$F_t(t,x) + rxF_x(t,x) + \tfrac{1}{2}\sigma^2 x^2 F_{xx}(t,x) - rF(t,x) = 0,$$

subject to the boundary condition $F(T,x) = f(x)$.

Proof. The function F is smooth in both arguments. Therefore, we can apply Itô's formula to the the value process $V_t = F(t, S_t)$ to see that this satisfies the SDE

$$dV_t = F_t(t,S_t)\,dt + F_x(t,S_t)\,dS_t + \tfrac{1}{2}F_{xx}(t,S_t)\,d[S]_t.$$

By Example 5.14 we have $dS_t = (\mu + \sigma^2/2)S_t\,dt + \sigma S_t\,dW_t$. In particular its quadratic variation satisfies $d[S]_t = \sigma^2 S_t^2\,dt$. Substituting these identities in the preceding display we see that the SDE for V_t reduces to

$$dV_t = \Bigl(F_t(t,S_t) + (\mu + \tfrac{1}{2}\sigma^2)F_x(t,S_t)S_t + \tfrac{1}{2}\sigma^2 F_{xx}(t,S_t)S_t^2\Bigr)\,dt \\ + \sigma F_x(t,S_t)S_t\,dW_t.$$

Using the definition of $\tilde{W}$ we can also write this equation in the form

$$dV_t = \Bigl(F_t(t,S_t) + rF_x(t,S_t)S_t + \tfrac{1}{2}\sigma^2 F_{xx}(t,S_t)S_t^2\Bigr)\,dt + \sigma F_x(t,S_t)S_t\,d\tilde{W}_t.$$

On the other hand, equations (6.1) and (6.2) imply that $d\tilde{V}_t = \sigma\phi_t\tilde{S}_t\,d\tilde{W}_t$. By the exercise above, it follows that $V_t = e^{rt}\tilde{V}_t$ satisfies

$$dV_t = rF(t,S_t)\,dt + \sigma\phi_t S_t\,d\tilde{W}_t.$$

Comparison of the dt-terms of the last two equations for dV_t yields the PDE for the function F.

The boundary condition follows from the fact that $f(S_T) = V_T = F(T, S_T)$. ∎

It should be noted that the PDE for the value process is the same for every type of European option. The type of option is only important for the boundary condition.

In the proof of the preceding theorem we only compared the dt-terms of the two SDE's that we obtained for the value process V of the claim. By comparing the $d\tilde{W}_t$-terms we obtain the following explicit formulas for the hedging portfolio of the claim.

6.8 Theorem. *A European claim $C = f(S_T)$ with value process $V_t = F(t, S_t)$ can be hedged by a self-financing portfolio consisting at time t of ϕ_t risky assets and ψ_t bonds, where*

$$\phi_t = F_x(t, S_t),$$
$$\psi_t = e^{-rt}\Big(F(t, S_t) - F_x(t, S_t)S_t\Big).$$

Proof. The formula for ϕ_t follows from the comparison of the $d\tilde{W}_t$-terms of the two SDE's for V that we obtained in the proof of the preceding theorem. Recall that $\psi_t = \tilde{V}_t - \phi_t \tilde{S}_t$. Substituting $V_t = F(t, S_t)$ and $\phi_t = F_x(t, S_t)$ yields the formula for ψ_t. ∎

The hedging strategy exhibited in the preceding theorem is called the *delta hedge* for the claim. Note that in general, the numbers of stocks and bonds in the hedging portfolio change continuously. In practice it is of course not possible to trade continuously. Moreover, very frequent trading will not always be sensible in view of transaction costs. However, the delta hedge can be used in practice to indicate what a hedging portfolio should look like.

6.5 The Greeks

Parties which are trading a claim with associated value function $V_t = F(t, S_t)$ are often interested in the sensitivity of the price of the claim with respect to the price of the underlying risky asset, and also with respect to time, volatility, etc. Reasonable measures for these sensitivities are the derivatives of the function $F(t, x)$. These derivatives have special names. The quantities

$$\Delta = F_x, \quad \Gamma = F_{xx}, \quad \Theta = F_t, \quad \mathbb{V} = F_\sigma$$

are called the delta, gamma, theta and vega of the claim, respectively. Together they are called the *Greeks.*

For instance, the delta of a claim measures the first order dependence of the price of the claim relative to the price of the underlying asset. A very high delta means that small changes in the asset price cause relatively large changes in the value of the claim. Observe that the delta is precisely the number of stocks in the hedging portfolio of Theorem 6.8.

6.9 EXERCISE. Calculate the delta for the European call option and give the delta hedging strategy for the claim.

6.6 General Claims

If the claim $C \in \mathcal{F}_T$ is not of European type we typically have no nice closed-form expression for its price. In that case, one can use simulation to find the approximate price. The price of the claim is given by $\mathrm{E}_{\mathbb{Q}} e^{-rT} C$ and in the preceding section we saw that under the martingale measure $\mathbb{Q}$

$$S_t = S_0 e^{(r-\frac{1}{2}\sigma^2)t+\sigma\tilde{W}_t},$$

where $\tilde{W}$ is a Brownian motion. To approximate the price, the following procedure can be followed:

1) Simulate a large number, say n, of realizations of the process S under $\mathbb{Q}$.
2) For each realization, compute the corresponding pay-off of the claim, yielding n numbers $C_1, \ldots, C_n$. Compute the average

$$\overline{c}_n = \frac{1}{n}\sum_{i=1}^n C_i.$$

3) Then by the law of large numbers, the discounted average $e^{-rT}\overline{c}_n$ is a good approximation for the price $\mathrm{E}_{\mathbb{Q}} e^{-rT} C$ if n is large enough.

We can quantify the quality of the approximation by obtaining a confidence interval. For this we also need the sample standard deviation s_n, which is defined by

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^n (C_i - \overline{c}_n)^2.$$

By the central limit theorem and the law of large numbers we have the convergence in distribution

$$\sqrt{n}\Big(\frac{\overline{c}_n - \mathrm{E}_{\mathbb{Q}} C}{s_n}\Big) \rightsquigarrow N(0,1)$$

as $n \to \infty$. Hence, for large n, we have the approximation

$$\mathbb{Q}\Big(\sqrt{n}\Big|\frac{\overline{c}_n - \mathrm{E}_{\mathbb{Q}} C}{s_n}\Big| > 1.96\Big) \leq 0.05.$$

It follows that $[e^{-rT}(\overline{c}_n - 1.96 s_n/\sqrt{n}), e^{-rT}(\overline{c}_n + 1.96 s_n/\sqrt{n})]$ is an approximate 95%-confidence interval for the price of the claim C. The length of the interval tends to zero as $n \to \infty$, which means that our simulation scheme can achieve arbitrary accuracy if we simulate long enough. In practice we shall be limited by computation time.

6.7 Exchange Rate Derivatives

Companies who do business in a country with a different currency are often interested in reducing the risk due to uncertainty in the exchange rate. One possibility to reduce this risk is to buy a suitable "exchange rate derivative". For instance, a Dutch company that will place a large order in the US one month from now may want to have an option to buy a large number of dollars for a specified price (in euros) at that time. In this section we use the developed Black-Scholes theory to derive the fair price of such a derivative.

We assume that there exist dollar bonds in the US and we can trade in euro bonds in The Netherlands. The prices of these bonds (in their respective currencies) are supposed to be given by

$$D_t = e^{qt}, \quad B_t = e^{rt}$$

respectively, where r is the European interest rate and q is the US interest rate. The exchange rate E_t, i.e. the euro value of one dollar, is modelled as a geometric Brownian motion,

$$E_t = E_0 e^{\nu t + \sigma W_t},$$

for certain parameters ν, σ and a Brownian motion W.

From the Dutch perspective, we can now trade in two assets: the riskless euro bond and the "risky" US bond, which have (euro) price processes B and $S = ED$, respectively. The process S is given by

$$S_t = E_t D_t = S_0 e^{(q+\nu)t + \sigma W_t}.$$

In other words, S is a geometric Brownian motion with drift $q + \nu$ and volatility σ. From the point of view of a Dutch trader this is just a standard Black-Scholes market and we know how to price derivatives.

Consider for instance a contract giving a Dutch trader the right to buy one US dollar for K euros at time $T > 0$. The pay-off at time T of this contract in euros is $(E_T - K)^+$. By the standard theory the fair euro price of the contract is $e^{-rT}\mathrm{E}_{\mathbb{Q}}(E_T - K)^+$, where $\mathbb{Q}$ is the martingale measure, under which the discounted price process $\tilde{S}_t = e^{-rt}S_t$ is a martingale. Note that

$$e^{-rT}\mathrm{E}_{\mathbb{Q}}(E_T - K)^+ = e^{-qT}\mathrm{E}_{\mathbb{Q}} e^{-rT}(S_T - Ke^{qT})^+.$$

This is $\exp(-qT)$ times the standard Black-Scholes price of a European call option with maturity T and strike $K\exp(qT)$. For the latter we have an explicit formula.

7
Extended Black-Scholes Models

The classical Black-Scholes model that we considered in the preceding chapter can be extended in several directions. So far we only considered markets in which a single bond and one risky asset are traded. We can also study the more complex situation that there are several risky assets with price processes that do not evolve independently. This allows the pricing of derivatives which depend on the behaviour of several assets. The assumption of a constant drift μ and volatility σ can also be relaxed. They can be replaced by arbitrary, predictable stochastic processes.

In general we can consider a market in which a bond is traded with price process B and n risky assets with price processes $S^1, \ldots, S^n$. We assume that the bond price is of the form $B_t = \exp(\int_0^t r_s\,ds)$ for r_t the "interest rate" at time t, so that it satisfies the ordinary diffential equation

$$dB_t = B_t r_t\,dt, \qquad B_0 = 1.$$

The interest rate r may be an arbitrary predictable process and hence depend on all information before time t. We assume that the asset price processes satisfy the system of stochastic differential equations

$$dS_t^i = \mu_t^i S_t^i\,dt + S_t^i \sum_{j=1}^d \sigma_t^{ij}\,dW_t^j, \qquad i = 1, \ldots, n, \tag{7.1}$$

where $W^1, \ldots, W^d$ are d independent Brownian motions, and the μ^i and σ^{ij} are predictable processes. Then the processes μ^i model the drift, and the σ^{ij} model both the volatility and the dependence structure of the price processes.

Under certain conditions such general market models are also free of arbitrage and have the property that each claim that is a function of the asset prices can be hedged by a self-financing trading strategy. Explicit pricing formulas are usually not available in such general models. However, if the model is free of arbitrage and complete, the general fact that "price is expectation of discounted pay-off under a martingale measure" is still true. The SDE's satisfied by the price processes under the martingale measure are typically easily obtained, so the simulation method can be used to approximate claim prices. This requires the simulation of solutions of multi-dimensional SDE's.

7.1 Market Price of Risk

The key structural condition needed to push through the theory is the existence of a predictable, vector-valued process $\theta = (\theta^1, \ldots, \theta^d)$, called the *market price of risk*, such that

$$\sum_{j=1}^{d} \sigma_t^{ij} \theta_t^j = r_t - \mu_t^i, \qquad i = 1, 2, \ldots, n. \tag{7.2}$$

We can write this system of equations in vector form as $\sigma_t \theta_t = r_t 1 - \mu_t$, and hence the existence of the "market price of risk" process requires that the vector $r_t 1 - \mu_t$ is contained in the range space of the $(n \times d)$-matrix σ_t. This is immediate if the rank of σ_t is equal to the number n of stocks in the economy, as the range of σ_t is all of $\mathbb{R}^n$ in that case. If the rank of σ_t is smaller than the number of stocks, then existence of the market price of risk process requires a relationship between the three parameters σ, r and μ. This situation is certain to arise if the number of components of the driving Brownian motion is smaller than the number of risky assets, i.e. $d < n$. Hence we can interpret the condition of existence of a process θ as in the preceding display as implying that the "random inputs $W^{(i)}$ to the market should be at least as numerous as the (independent) risky assets". We shall see a somewhat different interpretation when discussing models for the term structure of interest rates, where the market price of risk assumption will come back in the natural, intuitive form that "a market cannot have two different interest rates".

7.2 Fair Prices

In the present extended situation a portfolio is still a pair (ψ, ϕ) of a predictable process ψ_t, giving the number of bonds, and a vector-valued predictable process $\phi_t = (\phi_t^1, \ldots, \phi_t^n)$, giving the numbers of assets of the various types. To make the integrals well defined we assume that

$$\sum_{i=1}^{n} \int_0^T |\phi_t^i|^2 \, dt + \int_0^T |\psi_t| \, dt < \infty.$$

The portfolio is called *self-financing* if its value process $V_t = \psi_t B_t + \sum_{i=1}^n \phi_t^i S_t^i$ satisfies

$$dV_t = \sum_{i=1}^{n} \phi_t^i \, dS_t^i + \psi_t \, dB_t. \tag{7.3}$$

By definition the fair price at time t of a claim C is the value V_t of a replicating strategy at time t, where a "replicating strategy" is exactly as before a self-financing strategy whose value at T is equal to $V_T = C$.

In the present situation we discount using the process B rather than the exponential factors e^{rt}. Thus the discounted stock processes are

$$\tilde{S}_t^i = B_t^{-1} S_t^i = e^{-\int_0^t r_s \, ds + \int_0^t \left(\mu_s^i - \frac{1}{2} \sum_{j=1}^d (\sigma_s^{ij})^2\right) ds + \sum_{j=1}^d \sigma_s^{ij} \, dW_s^j}.$$

The second equality follows from the definition of B, the SDE (7.1) for the asset prices and Itô's formula, applied as in Example 5.14. In other words, by another application of Itô's formula,

$$d\tilde{S}_t^i = \tilde{S}_t^i (\mu_t^i - r_t) \, dt + \tilde{S}_t^i \sum_{j=1}^{d} \sigma_t^{ij} \, dW_t^j.$$

If there exists a market price of risk process θ, then this can be rewritten in the form

$$d\tilde{S}_t^i = \tilde{S}_t^i \sum_{j=1}^{d} \sigma_t^{ij} \, d\tilde{W}_t^j,$$

where $\tilde{W}_t = W_t - \int_0^t \theta_s \, ds$. Unless $\theta = 0$, the process $\tilde{W}$ will not be a $\mathbb{P}$-Brownian motion in view of Lemma 5.17. However, according to Girsanov's theorem there exists a probability measure $\mathbb{Q}$ under which $\tilde{W}$ is a Brownian motion (if θ is appropriately integrable). Under this "martingale measure" $\mathbb{Q}$ the discounted stock prices are local martingales.

We can now follow roughly the reasoning in Section 6.2 to construct a replicating portfolio for a claim that pays an amount C at time T. A key element in this construction is to find a process ϕ_t such that the martingale $\tilde{V}_t = \mathrm{E}_{\mathbb{Q}}\big(B_T^{-1} C |\, \mathcal{F}_t\big)$ is representable as $d\tilde{V}_t = \phi_t \, d\tilde{S}_t$. In the present vector-valued situation this is to be understood as

$$d\tilde{V}_t = \sum_{i=1}^{n} \phi_t^i \, d\tilde{S}_t^i.$$

If the matrices σ_t are square and invertible, then this representation can be easily obtained from the vector-valued version of the Brownian representation theorem, Theorem 5.19, by the same arguments as in Section 6.2. More generally, the desired representation is typically possible if the filtration $\mathcal{F}_t$ is generated by the asset processes S_t^i. In the following theorem we refer to this assumption by assuming that "the stock price processes possess the representation property".

7.4 Theorem. *Assume that there exists a predictable process θ satisfying (7.2), and that the stock price processes possess the representation property. Furthermore, assume that* $\mathrm{E}\exp\frac{1}{2}\int_0^T \|\sigma_s\|^2\,ds < \infty$ *and that* $\mathrm{E}\exp\frac{1}{2}\int_0^T \|\theta_s\|^2\,ds < \infty$. *Then the value of the claim $C \in \mathcal{F}_T$ at time $t \in [0,T]$ is given by $B_t \mathrm{E}_{\mathbb{Q}}(B_T^{-1}C|\,\mathcal{F}_t)$, where $\mathbb{Q}$ is the measure under which the discounted price processes $B_t^{-1}S_t^i$ are local martingales.*

7.3 Arbitrage

In the preceding it was seen that existence of the market-price-of-risk process is essential for the construction of a martingale measure $\mathbb{Q}$ under which the discounted stock price processes are local martingales. To underline the necessity of the existence of the market-price-of-risk we shall now show that without it, the market allows arbitrage.

The value process of a self-financing strategy (ψ,ϕ) can be written as (cf. (7.3))

$$\begin{aligned} V_t - V_0 &= \int_0^t \psi_s B_s r_s\,ds + \sum_{i=1}^n \int_0^t \phi_s^i\,dS_s^i \\ &= \int_0^t V_s r_s\,ds + \sum_{i=1}^n \int_0^t \phi_s^i\,(dS_s^i - S_s^i r_s\,ds). \end{aligned}$$

By the partial integration formula and the fact that $dB_t = r_t B_t\,dt$,

$$d\Big(\frac{V}{B}\Big)_t = -V_t\frac{1}{B_t^2}\,dB_t + \frac{1}{B_t}\,dV_t = \frac{1}{B_t}\sum_{i=1}^n \phi_t^i\,(dS_t^i - S_t^i r_t\,dt), \qquad (7.5)$$

by the preceding display. Hence the discounted value process takes the form, in view of (7.1),

$$\tilde{V}_t = \frac{V_t}{B_t} = V_0 + \sum_{i=1}^n \int_0^t \frac{S_s^i\phi_s^i}{B_s}\Big(\sum_{j=1}^d \sigma_s^{ij}\,dW_s^j - (r_s - \mu_s^i)\,ds\Big). \qquad (7.6)$$

This formula does not make explicit reference to the amount ψ invested in the bond, which has been eliminated. A "partial strategy" ϕ defines a value process through the preceding displays (7.5)-(7.6), and given ϕ we can define a process ψ from the equation $V_t = \psi_t B_t + \sum_{i=1}^n \phi_t^i S_t^i$. By retracing the calculations the resulting strategy (ψ, ϕ) can be seen to be self-financing and to possess value process V_t. Thus to see which value processes are possible it suffices to construct the stock portfolio ϕ.

Nonexistence of a market price of risk process implies that the vector $r_t 1 - \mu_t$ is not contained in the range of σ_t, for a positive set of times t. Then there exists a vector ϕ_t such that the vector $(S_t^1 \phi_t^1, \ldots, S_t^n \phi_t^n)$ is orthogonal to this range such that the inner product with the vector $r_t 1 - \mu_t$ is strictly negative for a positive set of times t:

$$\sum_{i=1}^n S_t^i \phi_t^i \sigma_t^{ij} = 0, \qquad j = 1, \ldots, d,$$
$$\sum_{i=1}^n \mathcal{S}_t^i \phi_t^i (r_t - \mu_t^i) < 0.$$

We can arrange it so that the latter inner product is never positive and hence, by (7.6), the corresponding discounted gain process will be zero at time 0 and strictly positive at time T. This is an example of arbitrage.

On the other hand, if the market price of risk process θ exists, then the discounted gains process in (7.6) can be written as a stochastic integral relative to the process $\sigma \cdot \tilde{W}$, for $\tilde{W}_t = W_t - \int_0^t \theta_s \, ds$. Under the martingale measure $\mathbb{Q}$ the process $\tilde{W}$ is a Brownian motion, and hence the discounted gains process will be a local martingale. Under the integrability assumptions of Theorem 7.4 it is a $\mathbb{Q}$-martingale, and hence cannot become strictly positive as its mean must remain zero. Thus existence of the market price of risk is intimately connected to the nonexistence of arbitrage.

7.4 PDEs

Under the conditions of Theorem 7.4 the process $\tilde{W} = (\tilde{W}^1, \ldots, \tilde{W}^d)$ defined by $\tilde{W}_t = W_t - \int_0^t \theta_s \, ds$ is a Brownian motion under the martingale measure $\mathbb{Q}$. Because option prices can be written as expectations under $\mathbb{Q}$, it is useful to rewrite the stochastic differential equation (7.1) in terms of the process $\tilde{W}$. If we also assume that the processess r and σ take the forms $r_t = r(t, B_t, S_t)$ and $\sigma_t = \sigma(t, B_t, S_t)$, then the equations describing the asset prices become

$$\begin{aligned} dB_t &= B_t\, r(t, B_t, S_t)\, dt, \\ dS_t^i &= S_t^i\, r(t, B_t, S_t)\, dt + S_t^i \sum_{j=1}^d \sigma^{ij}(t, B_t, S_t)\, d\tilde{W}_t^j, \quad i = 1, \ldots, n. \end{aligned} \tag{7.7}$$

As usual we assume that (B, S) is adapted to the augmented natural filtration $\mathcal{F}_t^{\tilde{W}}$ of $\tilde{W}$. Then, under regularity conditions on r and σ, the process (B, S) will be Markovian relative to this filtration. If we assume in addition that σ is invertible, then $\tilde{W}$ can be expressed in (B, S) by inverting the second equation, and hence the filtrations $\mathcal{F}_t$ and $\mathcal{F}_t^{\tilde{W}}$ generated by (B, S) and $\tilde{W}$ are the same. The process (B, S) is then Markovian relative to its own filtration $\mathcal{F}_t$. In that case a conditional expectation of the type $\mathrm{E}_{\mathbb{Q}}(X|\, \mathcal{F}_t)$ of a random variable X that is a function of $(B_s, S_s)_{s \geq t}$ can be written as $F(t, B_t, S_t)$ for a function F.

This observation can be used to characterize the value processes of certain options through a partial differential equation. The value process of a claim that is a function $C = g(S_T)$ of the final value S_T of the stocks takes the form

$$V_t = B_t \mathrm{E}_{\mathbb{Q}}\Bigl(\frac{g(S_T)}{B_T}|\, \mathcal{F}_t\Bigr) = \mathrm{E}_{\mathbb{Q}}\Bigl(e^{-\int_t^T r(s, B_s, S_s)\, ds} g(S_T)|\, \mathcal{F}_t\Bigr).$$

If the process (B, S) is Markovian as in the preceding paragraph, then we can write $V_t = F(t, B_t, S_t)$ for a function F. We assume that this function possesses continuous partial derivatives up to the second order. For simplicity of notation we also assume that S is one-dimensional. Then, by Itô's formula,

$$dV_t = F_t\, dt + F_b\, dB_t + F_s\, dS_t + \tfrac{1}{2} F_{ss}\, d[S]_t.$$

Here F_t, F_b, F_s are the first order partial derivatives of F relative to its three arguments, F_{ss} is the second order partial derivative relative to its third argument, and for brevity we have left off the argument (t, B_t, S_t) of these functions. A second application of Itô's formula and substitution of the diffusion equation for (B, S) yields

$$d\Bigl(\frac{V_t}{B_t}\Bigr) = \frac{1}{B_t}\Bigl(-Fr + F_t + F_b B_t r + F_s S_t r + \tfrac{1}{2} F_{ss} S_t^2 \sigma^2\Bigr)\, dt + \frac{1}{B_t} F_s \mathcal{S}_t \sigma\, d\tilde{W}_t.$$

The process V_t/B_t was seen previously to be a $\mathbb{Q}$-local martingale. Because the process $\tilde{W}$ is a Brownian motion, this can only be true if the drift term on the right side of the preceding display is zero, i.e.

$$\begin{aligned} &-(Fr)(t, b, s) + F_t(t, b, s) + (rF_b)(t, b, s)b \\ &\qquad + (rF_s)(t, b, s)s + \tfrac{1}{2}(\sigma^2 F_{ss})(t, b, s)s^2 = 0. \end{aligned}$$

This is an extension of the *Black-Scholes partial differential equation.*

This partial differential equation is useful for the numerical computation of option prices. Even though the equation is rarely explicitly solvable, a variety of numerical methods permit to approximate the solution F. The equation depends only on the functions r and σ defining the stochastic differential equation (7.7). Hence it is the same for every option with a claim of the type $C = g(S_T)$, the form of the claim only coming in to determine the boundary condition. Because $C = g(S_T) = F(T, B_T, S_T)$, this takes the form

$$F(T, r, s) = g(s).$$

For instance, for a European call option on the stock S, this becomes $F(T, r, s) = (s - K)^+$.

8
Interest Rate Models

8.1 The Term Structure of Interest Rates

In the classical Black-Scholes model the interest rate is a deterministic constant. In reality the situation is much more complicated of course. In general, it is not even possible to talk about *the* interest rate, since short term and long term rates are usually different. Moreover, the time evolution of interest rates typically has a random component.

In this chapter we introduce interest rate models that capture these properties of the time value of money. Such models are necessary for the pricing of so-called interest rate derivatives. These are financial contracts that are designed to trade and manage the risk that is caused by the uncertainty about the time value of money.

8.1.1 Discount Bonds

Pure discount bonds are simple financial contracts that capture the time value of money. A *discount bond* which matures at time $T > 0$, also called a T-*bond*, is a contract which guarantees a pay-off of 1 euro at time T. The price of a T-bond at time $t \leq T$ is denoted by $P(t, T)$. It is the amount we are willing to pay at time t to receive 1 euro at time T. The collection $\{P(0,T) \colon T > 0\}$ of all bond prices at time $t = 0$ completely determines the time-value of money at time 0. It is called the *term structure of interest rates.*

For fixed t, the function $T \mapsto P(t, T)$ is typically smooth, since, for instance, the price of a bond that matures 9 years from now will be close to the price of a bond that matures 10 years from now. For a fixed maturity $T > 0$ however, the function $t \mapsto P(t, T)$ will appear to fluctuate randomly.

By construction it holds that $P(T,T) = 1$.

8.1.2 Yields

If we have 1 euro at time t, we can use it to buy $1/P(t,T)$ T-bonds. At time T we then receive $1/P(t,T)$ euros. Hence, a euro at time t grows to $1/P(t,T)$ euros at time T. It the interest rate over the interval $[t,T]$ had been constant, say r, a euro at time t would have grown to $\exp(r(T-t))$ at time T. If we compare this, we see that buying the T-bonds at time t leads to a "constant interest rate" over the time interval $[t,T]$ of

$$Y(t,T) = -\frac{\log P(t,T)}{T-t}. \tag{8.1}$$

We call this the *yield* over $[t,T]$. The collection of all yields of course contains exactly the same information as the collection of all bond prices. However, the yields have a somewhat easier interpretation in terms of interest rates.

8.1.3 Short Rate

Although *the* interest rate does not exist, we can construct an object that can be interpreted in this way. We just saw that the yield $Y(t,T)$ can be interpreted as the constant interest rate valid in the time interval $[t,T]$. The number

$$r_t = \lim_{T\downarrow t} Y(t,T) = -\frac{\partial}{\partial T} \log P(t,T)\Big|_{T=t}$$

can therefore be viewed as the interest rate at time t (or in the infinitesimal interval $[t, t+dt]$). We call r_t the *short rate* at time t. From its definition it is clear that in general, the short rate does not contain all information about the time value of money.

8.1.4 Forward Rates

Let $t < S < T$ and consider the following strategy. At time t, we sell one S-bond, giving us $P(t,S)$ euros. We immediately use this money to buy $P(t,S)/P(t,T)$ T-bonds. At time S the S-bond matures, which means we have to pay one euro to its holder. At time T the T-bond matures, and we receive $P(t,S)/P(t,T)$ euros.

If we follow this strategy, the net effect is that one euro at time S grows to $P(t,S)/P(t,T)$ euros at time T. If the interest rate were a constant r over the time interval $[S,T]$, one euro at time S would grow to $\exp(r(T-S))$ at time T. Hence, the "constant interest rate over $[S,T]$ determined at time t" is

$$-\frac{\log P(t,T) - \log P(t,S)}{T-S}.$$

This number is called the *forward rate for* $[S, T]$, *contracted at time* t. If we let $S \uparrow T$ we get

$$f(t, T) = -\frac{\partial}{\partial T} \log P(t, T), \tag{8.2}$$

which is the *forward rate at time* T, *contracted at time* t. Note that the short rate is a particular forward rate, we have $f(t, t) = r_t$. Moreover, it is easy to see that

$$P(t, T) = e^{-\int_t^T f(t,s)\,ds},$$

so the collection of all forward rates contains all information about the term structure of interest rates.

8.2 Short Rate Models

The classical approach to interest rate models is to specify a stochastic model for the short rate r_t and to assume that the bond price $P(t, T)$ is some smooth function of r_t. A model of this type is called a *short rate model*.

So let us suppose that under the "real world" probability measure $\mathbb{P}$, the short rate satisfies the SDE

$$dr_t = \mu(t, r_t)\,dt + \sigma(t, r_t)\,dW_t, \tag{8.3}$$

where W is a $\mathbb{P}$-Brownian motion, and μ and σ are certain functions on $[0, \infty) \times \mathbb{R}$. Let $(\mathcal{F}_t)$ be the filtration generated by the process r. We assume that we can put money in a bank which pays the interest rate r_t, in the sense that one euro at time zero grows to B_t euros at time t, where $B_t = \exp(\int_0^t r_s\,ds)$. In differential notation, the process B satisfies

$$dB_t = r_t B_t\,dt.$$

For the bond prices we assume that $P(t, T) = F^T(t, r_t)$, where F^T is some smooth function on $[0, \infty) \times \mathbb{R}$ which may depend on the time to maturity T. Clearly, the functions should satisfy $F^T(T, r) = 1$ for all T and r. In the preceding section we noted that the short rate does not determine the whole term structure of interest rates, so we can expect that we have some freedom in choosing the functions F^T. On the other hand we do not want to allow arbitrage opportunities in the bond market. It is intuitively clear that this implies certain restrictions on the relation between the prices of the T-bonds for various T, leading to restrictions on the functions F^T. In the remainder of this section we explain how we can construct arbitrage free short rate models.

The first step is the observation that by the absence of arbitrage there cannot be banks with different rates of interest.

8.4 Lemma. *Suppose there exists a self-financing portfolio with value process V which satisfies $dV_t = q_t V_t\,dt$ for some adapted process q. Then $q_t = r_t$ for all $t \geq 0$.*

Proof. We sketch the proof. Suppose for simplicity that q and r are constant and that $q > r$. Then we can borrow 1 euro at rate r and invest it in the portfolio which pays "interest" q. At time T, say, we sell the portfolio, giving us $\exp(qT)$ euros. We pay back our loan, which is now $\exp(rT)$, and are left with a risk-free profit of $\exp(qT) - \exp(rT)$ euros. This is clearly an arbitrage, which is not allowed.

The general case of random, nonconstant processes q and r can be handled similarly. ∎

By assumption the price $P(t,T)$ of a T-bond is given by $P(t,T) = F^T(t,r_t)$. This is a smooth function of t and a process which satisfies an SDE. Hence, by Itô's formula, we have that

$$dP(t,T) = F_t^T(t,r_t)\,dt + F_r^T(t,r_t)\,dr_t + \tfrac{1}{2}F_{rr}^T(t,r_t)\,d[r]_t.$$

If we combine this with the SDE (8.3) for the short rate r_t we obtain

$$dP(t,T) = \alpha^T(t,r_t)P(t,T)\,dt + \sigma^T(t,r_t)P(t,T)\,dW_t, \tag{8.5}$$

where the functions α^T and σ^T are given by

$$\alpha^T = \frac{F_t^T + \mu F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T}{F^T}, \tag{8.6}$$

$$\sigma^T = \frac{\sigma F_r^T}{F^T}. \tag{8.7}$$

Below we write α_t^T and σ_t^T instead of $\alpha^T(t,r_t)$ and $\sigma^T(t,r_t)$.

To study the relation between the prices of bonds with different maturities we are now going to consider a self-financing portfolio consisting of S-bonds and T-bonds, for $S < T$. Suppose we are given such a portfolio, consisting at time $t < S$ of ϕ_t^T T-bonds and ϕ_t^S S-bonds, and let V denote its value process. Since the portfolio is self-financing we have

$$\begin{aligned} dV_t &= \phi_t^T\,dP(t,T) + \phi_t^S\,dP(t,S) \\ &= u_t^T V_t \frac{dP(t,T)}{P(t,T)} + u_t^S V_t \frac{dP(t,S)}{P(t,S)}, \end{aligned}$$

where u^T and u^S are the fractions of the portfolio consisting respectively of T-bonds and S-bonds, defined by

$$u_t^T = \frac{\phi_t^T P(t,T)}{V_t}, \quad u_t^S = \frac{\phi_t^S P(t,S)}{V_t}.$$

If we combine this with the SDE (8.5) for $P(t,T)$ we get

$$dV_t = \left(u_t^T\alpha_t^T + u_t^S\alpha_t^S\right)V_t\,dt + \left(u_t^T\sigma_t^T + u_t^S\sigma_t^S\right)V_t\,dW_t.$$

This SDE holds for every self-financing portfolio consisting of S-bonds and T-bonds. Conversely, we can construct a particular portfolio by specifying fractions u^T and u^S satisfying $u_t^T + u_t^S = 1$. The choice

$$u_t^T = -\frac{\sigma_t^S}{\sigma_t^T - \sigma_t^S},$$

$$u_t^S = \frac{\sigma_t^T}{\sigma_t^T - \sigma_t^S}$$

leads to a self-financing portfolio with value process V satisfying

$$dV_t = \left(\frac{\alpha_t^S\sigma_t^T - \alpha_t^T\sigma_t^S}{\sigma_t^T - \sigma_t^S}\right)V_t\,dt.$$

The dW_t-term has disappeared, so by Lemma 8.4 it must hold that

$$\frac{\alpha_t^S\sigma_t^T - \alpha_t^T\sigma_t^S}{\sigma_t^T - \sigma_t^S} = r_t$$

for all $t \geq 0$. We can rewrite this relation as

$$\frac{\alpha_t^S - r_t}{\sigma_t^S} = \frac{\alpha_t^T - r_t}{\sigma_t^T}.$$

In other words, the ratio $(\alpha_t^T - r_t)/\sigma_t^T$ must be independent of T. Thus, we have proved the following lemma.

8.8 Lemma. *There exists a function λ on $[0,\infty)\times\mathbb{R}$, independent of T, such that, for all t,T,*

$$\lambda(t,r_t) = \frac{\alpha^T(t,r_t) - r_t}{\sigma^T(t,r_t)}$$

Recall that α_t^T and σ_t^T are the local rate of return and volatility of the T-bond, respectively (cf. (8.5)). Hence, the difference $\alpha_t^T - r_t$ can viewed as a *risk premium*. It is the excess return that we get if we invest in the risky T-bond instead of putting our money in the bank. The quantity $(\alpha_t^T - r_t)/\sigma_t^T$, i.e. the *risk premium per unit of volatility*, is called the *market price of risk* of the T-bond. In this terminology the preceding lemma states that in an arbitrage free bond market, all bonds have the same market price of risk.

If we combine the result of Lemma 8.8 with the definitions (8.6) and (8.7) of the processes α^T and σ^T, we arrive at a PDE for the pricing functions of T-bonds, called the *term structure equation*.

8.9 Theorem. *Let $\lambda(t, r_t)$ denote the market price of risk. Then for every $T > 0$ the function F^T satisfies the PDE*

$$F_t^T + (\mu - \lambda\sigma)F_r^T + \tfrac{1}{2}\sigma^2 F_{rr}^T - rF^T = 0,$$

subject to the boundary condition $F^T(T, r) = 1$.

Under certain regularity conditions the term structure equation has a unique solution for every $T > 0$, so the bond prices $P(t, T) = F^T(t, r_t)$ are completely determined by the functions μ, σ and λ.

It is now clear how we can construct a short rate model leading to an arbitrage free bond market:

1) Specify the drift μ and volatility σ for the short rate r_t (under $\mathbb{P}$) and assume that r_t satisfies the SDE (8.3), where W is a $\mathbb{P}$-Brownian motion.
2) Choose a function λ on $[0, \infty) \times \mathbb{R}$ and for $T > 0$, let F^T be the solution of the term structure equation corresponding to μ, σ and λ.
3) Finally, define the price of a T-bond as $P(t, T) = F^T(t, r_t)$.

Observe that the term structure equation for the price of a T-bond is very similar to the Black-Scholes PDE for the pricing function of a European claim, cf. Theorem 6.7. In the preceding chapter we saw that the price of a European claim also equals the expectation of the discounted pay-off under a new measure $\mathbb{Q}$. We have the following analogous theorem for the price of a T-bond in a short rate model.

8.10 Theorem. *If μ and σ are the drift and volatility of the short rate under $\mathbb{P}$ and λ is the market price of risk, the price of a T-bond at time t is given by*

$$P(t, T) = B_t \mathrm{E}_{\mathbb{Q}}\left(B_T^{-1} \mid \mathcal{F}_t\right) = \mathrm{E}_{\mathbb{Q}}\left(e^{-\int_t^T r_s\, ds} \mid \mathcal{F}_t\right),$$

where $\mathbb{Q}$ is the measure under which the short rate satisfies the SDE

$$dr_t = \Big(\mu(t, r_t) - \sigma(t, r_t)\lambda(t, r_t)\Big)\, dt + \sigma(t, r_t)\, d\tilde{W}_t,$$

and $\tilde{W}$ is a $\mathbb{Q}$-Brownian motion.

Note that for every $T > 0$ it holds that the discounted price $\tilde{P}(t, T) = B_t^{-1} P(t, T)$ of a T-bound satisfies

$$B_t^{-1} P(t, T) = \mathrm{E}_{\mathbb{Q}}\left(B_T^{-1} \mid \mathcal{F}_t\right),$$

so for every $T > 0$ the process $(\tilde{P}(t, T))_{t \leq T}$ is a martingale under $\mathbb{Q}$. Therefore the measure $\mathbb{Q}$ appearing in the statement of the theorem is called the *martingale measure* of the model. Observe that the formula

$$P(0, T) = \mathrm{E}_{\mathbb{Q}} B_T^{-1}$$

for the current price of a T-bond is a statement of the usual form "price of a derivative is expectation of the discounted pay-off under the martingale measure", since a T-bond can be viewed as a claim which pays off 1 euro at time T.

Theorem 8.10 gives us a second method for the construction of a model for an arbitrage free bond market:

1) Specify an SDE for the short rate r_t under the martingale measure $\mathbb{Q}$ and let $(\mathcal{F}_t)$ be the natural filtration of the process r.
2) Define the price $P(t,T)$ of a T-bond by

$$P(t,T) = \mathrm{E}_{\mathbb{Q}}\Big(e^{-\int_t^T r_s\,ds} \,|\, \mathcal{F}_t\Big).$$

This second procedure for the construction of short rate models is known as *martingale modeling* and has the obvious advantage that we do not have to specify the market price of risk explicitly. In the next section we introduce a popular short rate model that is constructed in this way.

8.3 The Hull-White Model

The Hull-White model for the term structure of interest rates assumes that under the martingale measure $\mathbb{Q}$ the short rate r_t satisfies the SDE

$$dr_t = (\theta(t) - ar_t)\,dt + \sigma\,dW_t, \tag{8.11}$$

where a and σ are certain numbers, θ is a deterministic function and W is a $\mathbb{Q}$-Brownian motion. The natural filtration of r (and W) is denoted by $(\mathcal{F}_t)$ and the price $P(t,T)$ of a T-bond at time t is defined by

$$P(t,T) = \mathrm{E}_{\mathbb{Q}}\Big(e^{-\int_t^T r_s\,ds} \,|\, \mathcal{F}_t\Big).$$

By the preceding section this defines an arbitrage free model for the bond market.

It is possible to obtain concrete formulas for the bond prices in this model. The main reason is that we have an explicit expression for the solution of the SDE (8.11). This allows us to calculate the conditional distribution of the integral $\int_t^T r_s\,ds$ given $\mathcal{F}_t$, which we need to calculate $P(t,T)$.

8.12 Lemma. *Given $\mathcal{F}_t$ the integral $\int_t^T r_s\,ds$ possesses a Gaussian distribution with mean*

$$B(t,T)r_t + \int_t^T B(u,T)\theta(u)\,du$$

and variance

$$\sigma^2 \int_t^T B^2(u,T)\,du,$$

where

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a}. \tag{8.13}$$

Proof. First we apply the Itô formula to calculate $d(\exp(at)r_t)$ and use (8.11) to find that

$$r_s = e^{-as} r_0 + e^{-as} \int_0^s \theta(u) e^{au}\,du + \sigma e^{-as} \int_0^s e^{au}\,dW_u. \tag{8.14}$$

Integrating this from t to T and interchanging integrals gives

$$\begin{aligned}\int_t^T r_s\,ds =& r_0 e^{-at} B(t,T) + e^{-at} B(t,T) \int_0^t e^{au}\theta(u)\,du + \int_t^T B(u,T)\theta(u)\,du \\ &+ B(t,T)\sigma e^{-\alpha t} \int_0^t e^{\alpha u}\,dW_u + \sigma \int_t^T B(u,T)\,dW_u \\ =& B(t,T) r_t + \int_t^T B(u,T)\theta(u)\,du + \sigma \int_t^T B(u,T)\,dW_u.\end{aligned}$$

Given $\mathcal{F}_t$ the first two terms on the right-hand side are known. The third one is independent of $\mathcal{F}_t$ and is Gaussian with mean zero and variance

$$\sigma^2 \int_t^T B^2(u,T)\,du.$$

This completes the proof. ∎

We can now derive the bond price formula for the Hull-White model.

8.15 Theorem. *In the Hull-White model the price of a T-bond is given by*

$$P(t,T) = e^{A(t,T) - B(t,T) r_t}, \tag{8.16}$$

where B is defined by (8.13) and

$$A(t,T) = \int_t^T \Big(\tfrac{1}{2}\sigma^2 B^2(u,T) - \theta(u) B(u,T)\Big)\,du.$$

Proof. We have to calculate the expectation of $\exp(-\int_t^T r_s\,ds)$ given $\mathcal{F}_t$. By the preceding lemma this boils down to computing the expectation of the exponential of a Gaussian random variable. If Z is a Gaussian random variable with mean m and variance s^2 it holds that $\mathrm{E}\exp(Z) = \exp(m + s^2/2)$. Together with the lemma this yields the statement of the theorem. ∎

Short rate models in which the bond price is of the form (8.16) are called *affine models*. The reason for this name is that the yields and forward rates are affine in r_t in that case. The yield $Y(t,T)$ (see (8.1) is given by

$$Y(t,T) = \frac{B(t,T)}{T-t} r_t - \frac{A(t,T)}{T-t}$$

and for the forward rate (see (8.2) we have

$$f(t,T) = B_T(t,T) r_t - A_T(t,T).$$

Now consider a specific bond market in which bonds of all maturities are traded. Then at time zero, we can observe the bond prices and forward rates with all maturities. We denote the *observed* prices and rates in the market by $P^*(0,T)$ and $f^*(0,T)$, respectively. On the other hand, the Hull-White model gives the formula

$$f(0,T) = B_T(0,T) r_0 - A_T(0,T) \tag{8.17}$$

for the forward rates. Obviously, we would like to match the theoretical rates $f(0,T)$ with the observed rates $f^*(0,T)$. We will now show that this is possible by choosing an appropriate function θ in (8.11). This procedure is called *fitting the model to the term structure of interest rates.*

8.18 Theorem. *Let the parameters a, σ in (8.11) be given. Then with the choice*

$$\theta(T) = af^*(0,T) + f_T^*(0,T) + \sigma^2 B(0,T)(e^{-aT} + \tfrac{1}{2} a B(0,T)) \tag{8.19}$$

the theoretical Hull-White bond prices and forward rates coincide with the observed prices and rates. The price of a T-bond is then given by

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\Big(B(t,T) f^*(0,t) - \frac{\sigma^2}{4a} B^2(t,T)(1-e^{-2at}) - B(t,T) r_t\Big).$$

Proof. If we insert the expressions for A and B into (8.17) we see that we have to solve the equation

$$\begin{aligned} f^*(0,T) &= e^{-aT} r_0 + \int_0^T e^{-a(T-u)} \theta(u)\, du - \frac{\sigma^2}{2a^2}(1-e^{-aT})^2 \\ &= g(T) - h(T), \end{aligned}$$

where g is the solution of the differential equation $g' + ag = \theta, \quad g(0) = r_0$ and $h(t) = \sigma^2 B^2(0,t)/2$. Then $g(T) = f^*(0,T) + h(T)$, which shows that the solution of the equation for θ is given by

$$\theta(T) = g'(T) + ag(T) = f_T^*(0,T) + h_T(T) + a(f^*(0,T) + h(T)).$$

This proves the first part of the theorem.

The second statement can be obtained by inserting the expression for θ in formula (8.16). ■

So after fitting the term structure of interest rates there are only two free parameters left in the Hull-White model, a and σ. In practice these are determined by matching the theoretical prices for certain interest rate derivatives with observed prices. This procedure is called the *calibration* of the model. The pricing of interest rate derivatives is the topic of the next section.

8.4 Pricing Interest Rate Derivatives

The result of Theorem 8.10 can be viewed as a pricing formula for the simplest possible claim which has a pay-off of one euro at time T. Using the same arguments as above it can be extended to a general claim which pays some random amount $C \in \mathcal{F}_T$ at time T.

8.20 Theorem. *Let $C \in \mathcal{F}_T$ be a claim. Its value at time $t \leq T$ is given by*

$$V_t = \mathrm{E}_{\mathbb{Q}}\Big(e^{-\int_t^T r_s\,ds} C | \, \mathcal{F}_t\Big),$$

where $\mathbb{Q}$ is the martingale measure.

Note that for $C \equiv 1$ we indeed recover the formula for the bond price $P(t,T)$.

Many interest rate derivatives do not only have a payment at the time T of maturity, but also at certain fixed intermediate times. Holding such a product is equivalent to holding a portfolio of derivatives with different maturities. Hence, Theorem 8.20 implies the following.

8.21 Theorem. *Let $0 < T_1 < \cdots < T_n = T$ and $C_i \in \mathcal{F}_{T_i}$ for $i = 1, \ldots, n$. Consider a derivative with a payment of C_i at time T_i for $i = 1, \ldots, n$. Its value at time 0 is given by*

$$V_0 = \mathrm{E}_{\mathbb{Q}} \sum_{i=1}^{n} e^{-\int_0^{T_i} r_s\,ds} C_i,$$

where $\mathbb{Q}$ is the martingale measure.

The result of Theorem 8.21 can now be used to determine the price of a given derivative by simulation methods, just as we discussed in Section 6.6 for the Black-Scholes model. The procedure is as follows:

1) Simulate a large number of realizations, say n, of the short rate process r under the martingale measure $\mathbb{Q}$.

2) For realization number j, compute the pay-off of the claim and determine an approximation c_j for the discounted pay-off given by $\sum_i \exp(-\int_0^{T_i} r_s\, ds)C_i$.
3) Then by the law of large numbers,

$$\bar{c}_n = \frac{1}{n}\sum_{j=1}^{n} c_j$$

is a good approximation of the claim price, provided that n is large enough.
4) By the central limit theorem the interval $[\bar{c}_n - 1.96 s_n/\sqrt{n}, \bar{c}_n + 1.96 s_n/\sqrt{n}]$ is an approximate 95% confidence interval for the price of the derivative, where

$$s_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(c_j - \bar{c}_n)^2.$$

The usual approach to simulating realizations of the short rate is to discretize the SDE for r_t under $\mathbb{Q}$. Suppose that under $\mathbb{Q}$ the short rate satisfies (8.3), where μ and σ are given functions. Then for small $\Delta > 0$ we have the approximation

$$r_{(k+1)\Delta} - r_{k\Delta} \approx \mu(k\Delta, r_{k\Delta})\Delta + \sigma(k\Delta, r_{k\Delta})(W_{(k+1)\Delta} - W_{k\Delta})$$

for $k = 0, 1, 2, \ldots$. The increments $W_{(k+1)\Delta} - W_{k\Delta}$ are independent Gaussian random variables with mean zero and variance Δ. So the approximation can be written as

$$r_{(k+1)\Delta} - r_{k\Delta} \approx \Delta\mu(k\Delta, r_{k\Delta}) + \sqrt{\Delta}\sigma(k\Delta, r_{k\Delta})Z_k,$$

where $Z_0, Z_1, Z_2, \ldots$ are independent, standard Gaussian random variables.

This so-called *Euler approximation* of the SDE (8.3) can be used to simulate sample paths and to determine the corresponding approximation of the discounted pay-off. A single realization is constructed as follows:

1) Partition the interval $[0, T]$ into m intervals of length $\Delta = T/m$ and simulate m i.i.d. standard Gaussian random variables $Z_0, Z_1, \ldots, Z_{m\Delta}$.
2) The short rate r_0 at time 0 is known. The future rates $r_\Delta, r_{2\Delta}, \ldots, r_{m\Delta}$ are computed recursively by the formula

$$r_{(k+1)\Delta} = r_{k\Delta} + \Delta\mu(k\Delta, r_{k\Delta}) + \sqrt{\Delta}\sigma(k\Delta, r_{k\Delta})Z_k.$$

3) For this realization the pay-off is calculated and the discount factors $\exp(-\int_0^T r_s\, ds)$ are approximated by $\exp(-\Delta\sum_k r_{k\Delta})$.

The outlined procedure works for any short rate model. In the special case of the Hull-White model we have in fact an *exact* recurrence formula for the discretized short rate process. This means that realizations can be simulated without introducing approximation errors due to discretization of the SDE.

8.22 Theorem. *Suppose the short rate satisfies (8.11) under the martingale measure $\mathbb{Q}$, where θ is given by (8.19). Then $r_t = \alpha(t) + y_t$, where*

$$\alpha(t) = e^{-\alpha t} r_0 + f^*(0,t) + \tfrac{1}{2}\sigma^2 B^2(0,t)$$

and y_t satisfies $y_0 = 0$ and for $\Delta > 0$ and $k = 0, 1, \ldots$,

$$y_{(k+1)\Delta} - e^{-a\Delta} y_{k\Delta} \sim \sqrt{\tfrac{1}{2}\sigma^2 B(0, 2\Delta)}\, Z_k,$$

with $Z_0, Z_1, \ldots$ i.i.d. standard Gaussian and B given by (8.13).

Proof. It follows from (8.14) that $r_t = \alpha(t) + y_t$, where

$$\alpha(t) = e^{-at}\Big(r_0 + \int_0^t \theta(u) e^{au}\, du\Big)$$

and

$$y_t = \sigma e^{-at} \int_0^t e^{au}\, dW_u.$$

The expression for α in the statement of the theorem now follows after inserting (8.19) and some straightforward calculations. To prove the recurrence formula for the process y, observe that the random variables

$$e^{a(k+1)\Delta} y_{(k+1)\Delta} - e^{ak\Delta} y_{k\Delta} = \sigma \int_{k\Delta}^{(k+1)\Delta} e^{au}\, dW_u$$

are independent, Gaussian, have zero mean and variance

$$\sigma^2 \int_{k\Delta}^{(k+1)\Delta} e^{2au}\, du = \sigma^2 e^{2ak\Delta} \frac{e^{2a\Delta} - 1}{2a}.$$

Hence

$$e^{a(k+1)\Delta} y_{(k+1)\Delta} - e^{ak\Delta} y_{k\Delta} \sim e^{ak\Delta} \sqrt{\sigma^2 \frac{e^{2a\Delta} - 1}{2a}}\, Z_k,$$

with $Z_0, Z_1, \ldots$ i.i.d., standard Gaussian. The proof is completed by dividing this by $\exp(a(k+1)\Delta)$. ∎

8.5 Examples of Interest Rate Derivatives

In this section we discuss some common interest rate products, and their valuation.

8.5.1 Bonds with Coupons

In practice pure discount bonds are not often traded. Instead, bonds typically do not only have a pay-off at maturity, the so-called *principal value*, but also make smaller regular payments before maturity. Such a bond is called a *coupon bond.* A 10-year, 5% coupon bond with a principle value of 100 euros for instance, pays 5 euros every year until maturity and 100 euros at maturity (after ten year).

More generally, suppose that a bond makes a payment of k euros at times $T_1 < \cdots < T_n = T$, and pays off its principle value of 1 euro at time T. Then holding this coupon bond is equivalent to holding k pure discount bonds with maturity T_i for $i = 1, \ldots, n$, and one T-bond. Hence, the value of the coupon bond is

$$P(0,T) + k \sum_{i=1}^{n} P(0,T_i).$$

We remark that conversely, the prices of pure discount bonds may be expressed in terms of the prices of coupon bonds. In practice this is the usual way in which the prices of discount bonds are inferred from market data.

8.5.2 Floating Rate Bonds

There also exist bonds with intermediate payments that depend on the interest rates at the time of the payment. The *LIBOR rate for the time interval* $[S,T]$*, set at time* S is defined as

$$L(S,T) = -\frac{P(S,T) - 1}{(T-S)P(S,T)}.$$

This is simply the return per time unit on an investment at time S in a T-bond. A *floating rate bond* is a bond with additional payments at times $T_1 < \cdots < T_n = T$. The payment C_i at time T_i is

$$(T_i - T_{i-1})L(T_{i-1}, T_i) = \frac{1}{P(T_{i-1}, T_i)} - 1.$$

This is precisely the gain we would have had at time T_i if we had bought one euro worth of T_i-bonds at time T_{i-1}. The principle value of one euro is payed at time T.

By Theorem 8.21, the price of this asset at time 0 is given by

$$P(0,T) + \sum_i \mathrm{E}_{\mathbb{Q}} B_{T_i}^{-1} (P(T_{i-1}, T_i)^{-1} - 1).$$

By the tower property of conditional expectation and Theorem 8.10 the ith term in the sum equals

$$\begin{aligned}
&\mathrm{E}_{\mathbb{Q}}\mathrm{E}_{\mathbb{Q}}\Big(B_{T_i}^{-1}(P(T_{i-1},T_i)^{-1}-1)|\,\mathcal{F}_{T_{i-1}}\Big)\\
&\quad= \mathrm{E}_{\mathbb{Q}}(P(T_{i-1},T_i)^{-1}-1)B_{T_{i-1}}^{-1}P(T_{i-1},T_i)\\
&\quad= \mathrm{E}_{\mathbb{Q}}B_{T_{i-1}}^{-1} - \mathrm{E}_{\mathbb{Q}}B_{T_{i-1}}^{-1}P(T_{i-1},T_i)\\
&\quad= P(0,T_{i-1}) - P(0,T_i).
\end{aligned}$$

Hence, the value of the bond is

$$P(0,T) + \sum_{i=1}^{n}\Big(P(0,T_{i-1}) - P(0,T_i)\Big) = P(0,0) = 1$$

So the price of a floating rate bond equals its principal value.

8.23 EXERCISE. Derive the pricing formula for the floating rate bond by showing that there exists a self-financing portfolio with initial value of one euro which has the same pay-off as the bond.

8.5.3 Swaps

A *swap* is contract that exchanges a stream of varying, interest rate dependent payments for a stream of fixed amount payments, or vice versa. Consider for example time points $0 < T_1 < \cdots < T_n$. At time T_i we make a payment of $(T_i - T_{i-1})L(T_{i-1},T_i)$ euros and we receive k euros. In other words, we swap the gain obtained from a one euro investment in T_i-bonds at time T_{i-1} for the constant "gain" k. Buying this contract is equivalent to selling a floating rate bond and buying a coupon bond which pays k euros at each time T_i. Hence, the price of the swap at time 0 is

$$P(0,T_n) + k\sum_{i=1}^{n} P(0,T_i) - 1.$$

8.5.4 Swaptions

A *swaption* is a contract giving the holder the right to enter into a swap at a future date. Suppose for instance that the contract gives the right to enter at time $T_0 > 0$ into a swap described in the preceding section. Then the pay-off at time T_0 of the option is

$$\Big(P(T_0,T_n) + k\sum_{i=1}^{n} P(T_0,T_i) - 1\Big)^+.$$

By Theorem 8.20, the price at time 0 of the swaption is therefore given by

$$\mathrm{E}_{\mathbb{Q}} e^{\int_0^{T_0} r_s\,ds} \Big(P(T_0, T_n) + k \sum_{i=1}^{n} P(T_0, T_i) - 1 \Big)^{+}.$$

In general this expectation can not be evaluated analytically and one has to resort to numerical methods. Note by the way that the latter formula shows that a swaption can also be viewed as a call option on a coupon bond.

9
Risk Measurement

Financial institutions deal in risk of various types. *Market risk* is the exposure to the changing prices of assets on the market, and can be limited by using appropriate portfolios that include instruments such as options and swaps. Managing risk is important for:

- Internal management, e.g. optimizing profit subject to restrictions on risk.
- To fulfill the requirements of regulatory authorities, such as national banks.
- Credit ratings.

The management of risk requires that risk be measured. In this chapter we discuss the most popular measure of risk: *Value-at-Risk*, abbreviated VaR.

9.1 Value-At-Risk

Let V_t be the value of a portfolio at time t, and $\mathcal{F}_t$ the information available at time t. We fix a given time $t+\delta t$ in the future and confidence level $1-\alpha$. The variable $V_t - V_{t+\delta t}$ is the loss that we shall incur in the period $[t, t+\delta t]$. VaR is defined as a number such that the loss is with high probability smaller than this number.

9.1 Definition. VaR *is the upper α-quantile of the conditional distribution of $V_t - V_{t+\delta t}$ given $\mathcal{F}_t$, i.e.*

$$\mathbb{P}(V_t - V_{t+\delta t} < \mathrm{VaR}|\, \mathcal{F}_t) \leq 1 - \alpha \leq \mathbb{P}(V_t - V_{t+\delta t} \leq \mathrm{VaR}|\, \mathcal{F}_t).$$

The Value-at-Risk depends of course on both δt and α. For regulators a period of 10 days and $\alpha = 1\%$ is usual, whereas for other purposes periods from one day to a year or $\alpha = 0.05\%$ may be considered appropriate. VaR as defined here refers to the absolute value of the portfolio, but for some purposes it can be useful to consider the "Value-at-Risk per capital" VaR/V_t instead.

The conditioning on the past information $\mathcal{F}_t$ in our definition appears natural, but is not always made explicit. Although we shall consider VaR in the following only at one fixed time t, Value-at-Risk is in our definition a stochastic process in time. For larger lags δt the dependence on t will typically be small.

One criticism to the use of VaR for risk measurement is that it may say little about expected losses. In particular, it says nothing about the sizes of the losses bigger than VaR, which occur with frequency α, and which could be arbitrarily big. Related to this is that VaR is in general not subconvex under combination of portfolios. If two subunits of a financial institution both control their VaR, then there is no guarantee that the VaR of the financial institution as a whole is also under control. This is illustrated in Example 9.2. Other measures of risk, which are not open to these criticisms are the conditional variance (or "volatility") and the expected excess loss, given by

$$\mathrm{var}(V_t - V_{t+\delta t}|\,\mathcal{F}_t), \qquad \text{and} \qquad \mathrm{E}\big((V_t - V_{t+\delta t} - c)^+|\,\mathcal{F}_t\big).$$

Here c is some given threshold. The volatility is a classical measure, dating back to Markowitz, who first proposed to optimize profit under the side condition of bounded volatility. It is somewhat unstable for heavy-tailed distributions and perhaps can be critized for being symmetric.

9.2 Example (Nonconvexity). Suppose that the portfolio consists of two parts, with values X_t and Y_t at time t, so that the total value is $V_t = X_t + Y_t$. Then $V_t - V_{t+\delta t} = \Delta X_t + \Delta Y_t$ for ΔX_t and ΔY_t the losses on the two subportfolios in the interval $[t, t+\delta t]$, and the relative contributions of the two subportfolios in the total are $w_X = X_t/V_t$ and $w_Y = Y_t/V_t$. If $\mathrm{VaR}(V)$, $\mathrm{VaR}(X)$, $\mathrm{VaR}(Y)$ are the Value-at-Risks of the three portfolios, then it may happen that the $\mathrm{VaR}(V)$ is bigger than the convex combination $w_X \mathrm{VaR}(X) + w_Y \mathrm{VaR}(Y)$.

For an example choose some fixed α and let the vector $(\Delta X_t, \Delta Y_t)$ be distributed (conditionally on $\mathcal{F}_t$) on the set of points $\{(0,0),(c,0),(0,c)\}$ according to the probabilities $1-2\alpha, \alpha, \alpha$, for a given $c > 0$. Then $\mathrm{VaR}(X) = \mathrm{VaR}(Y) = 0$, but $\mathrm{VaR}(V) = c$. We can make the discrepancy c between the total Value-at-Risk and the convex combination arbitrarily large. □

To determine VaR we need a model for the conditional distribution of $V_t - V_{t+\delta t}$ given $\mathcal{F}_t$. There are many possibilities, such as the ARMA and GARCH models from time series, or the Black-Scholes or Hull-White models from derivative pricing. It is important to note that we need the distribution of the value process under the real-world measure, not the martingale measure. Thus given a model the parameters are estimated from time series' giving the actual prices of the assets over time, so-called "historical analysis". Because some of the parameters may be common to the real-world measure and martingale measure, some parameters could also be calibrated using observed derivative prices.

In rare cases, depending on the model, it is possible to determine an analytic expression for the conditional distribution of $V_{t+\delta t} - V_t$ given $\mathcal{F}_t$. More often the VaR must be calculated numerically, for instance by stochastic simulation. If we can generate a large number N (e.g. at least $N = 10000$ if $\alpha = 1\%$) of realizations from the conditional distribution of $V_t - V_{t+\delta t}$ given $\mathcal{F}_t$, then VaR is approximately the $(1-\alpha)N$ largest value among the realizations. Remember that, unlike when using simulation to determine a derivative price, this time we must simulate from the real-world measure.

9.2 Normal Returns

In practice it is often assumed that the return $V_{t+\delta t}/V_t - 1$ is conditionally normally distributed given $\mathcal{F}_t$. (It may even be assumed that the returns are independent of the past and form an i.i.d. sequence if restricted to a discrete time grid, but that is even more unrealistic and unimportant for the following.) If the conditional mean and variance are μ_t and σ_t^2, then the conditional distribution of $V_t - V_{t+\delta t}$ is normal with mean $-V_t\mu_t$ and standard deviation $V_t\sigma_t$, and the Value-at-Risk is given by

$$\text{VaR} = V_t\big(\sigma_t\Phi^{-1}(1-\alpha) - \mu_t\big). \tag{9.3}$$

Note that it is proportional to the current capital V_t and linearly increasing in the volatility σ_t. A positive drift μ_t decreases VaR.

If a portfolio consists of n assets or subportfolios, with value processes $V^1, \dots, V^n$, then it is often assumed that the vector of returns $(V^1_{t+\delta t}/V^1_t - 1, \dots, V^n_{t+\delta t}/V^n_t - 1)$ is conditionally multivariate-normally distributed given $\mathcal{F}_t$. The value of the whole portfolio $V_t = \sum_{i=1}^n V_t^i$ can be written as

$$V_{t+\delta t} - V_t = V_t \sum_{i=1}^n w_t^i \Big(\frac{V^i_{t+\delta t}}{V^i_t} - 1\Big),$$

where $w_t^i = V_t^i/V_t$ is the relative contribution of asset i to the whole portfolio. The sum is a linear combination of a Gaussian vector and hence is normally distributed. If the return vector possesses conditional mean vector $(\mu_t^1, \dots, \mu_t^n)$ and covariance matrix $(\sigma_t^{i,j})$, then the conditional distribution of $V_t - V_{t+\delta t}$ given $\mathcal{F}_t$ is normal with mean $-V_t\mu_t$ and standard deviation

$V_t\sigma_t$, for

$$\mu_t = \sum_{i=1}^n w_t^i \mu_t^i, \qquad (9.4)$$
$$\sigma_t^2 = \sum_{i=1}^n \sum_{j=1}^n w_t^i w_t^j \sigma_t^{i,j}.$$

The Value-at-Risk again takes the same form (9.3), but with the new values of μ_t and σ_t substituted.

The Cauchy-Schwarz inequality says that the covariances satisfy $|\sigma_t^{i,j}| \leq \sigma_t^i \sigma_t^j$, where $\sigma_t^i = \sigma_t^{i,i}$ are the variances of the components. This shows that

$$\sigma_t \leq \sum_{i=1}^n w_t^i \sigma_t^i.$$

Because the VaR is linear in the standard deviation, this shows that the combined portfolio has a smaller VaR than a similar single portfolio of the same volatility. This expresses the well-known guideline that *diversification* of a portfolio helps to control the risk.

As Example 9.2 shows, diversification is not always useful to control VaR, but the preceding shows that for portfolios with normal returns it is.

Empirical studies have now well established that economic time series are not Gaussian random walks, an assumption of the past that still lives on in many VaR-methods. Returns are not i.i.d. and their marginal distributions deviate from normal distributions in that they are typically heavier tailed and sometimes skewed. Conditional normality of the returns given the past, as is assumed in this section, is also debatable, but not always rejected by statistical tests on empirical data. In particular, GARCH models are often combined with normal conditional distribution, which automatically leads to heavier-tailed unconditional distributions.

9.3 Equity Portfolios

The value process of a portfolio of one stock with price S_t is equal to $V_t = S_t$. If we adopt the Black-Scholes model, then $S_t = S_0 \exp(\mu t + \sigma W_t)$ for a Brownian motion process W, and hence the log returns satisfy

$$R_t := \log \frac{S_{t+\delta t}}{S_t} = \mu\delta t + \sigma(W_{t+\delta t} - W_t).$$

Because the increments of Brownian motion are independent of the past, it follows that the log returns are conditionally normally distributed with mean $\mu\delta t$ and variance $\sigma^2\delta t$. The loss can be expressed in the log returns as $V_t - V_{t+\delta t} = S_t(1 - e^{R_t})$. Solving the equation $\mathbb{P}(V_t(1-e^R) \le v) = 1 - \alpha$ for a $N(\mu\delta t, \sigma^2\delta t)$ distributed variable R and a fixed value V_t, yields the Value-at-Risk

$$\text{VaR} = V_t\Big(1 - e^{\sigma\sqrt{\delta t}\Phi^{-1}(\alpha)+\mu\delta t}\Big). \tag{9.5}$$

This has similar features as the equation (9.3): the risk is proportional to the current value V_t, increasing in the volatility σ and decreasing in the drift μ.

Because $\log x \approx x - 1$ for $x \approx 1$, the log return R_t is close to the "ordinary" return $S_{t+\delta t}/S_t - 1$, if δ is small. If we make this approximation and still assume the normal model for the return, then we are in the situation of Section 9.2, with $\mu_t = \mu\delta t$ and $\sigma_t = \sigma\sqrt{\delta t}$. The resulting formula (9.3) is identical to the formula obtained by replacing the exponential in (9.5) by its linear approximation.

The value process of a combined portfolio consisting of ϕ_t^i assets of price S_t^i is $(i = 1, 2)$ is given by $V_t = \phi_t^1 S_t^1 + \phi_t^2 S_t^2$. If we assume that the numbers ϕ_i^t do not change in the interval $[t, t + \delta t]$, then the gain in this interval is given by

$$V_{t+\delta t} - V_t = \phi_t^1(S_{t+\delta t}^1 - S_t^1) + \phi_t^2(S_{t+\delta t}^2 - S_t^2).$$

To determine VaR we need a model for the conditional distribution of the vector $(S_{t+\delta t}^1 - S_t^1, S_{t+\delta t}^2 - S_t^2)$. There are many possibilities.

A natural generalization of the Black-Scholes model would be to assume that both asset price processes follow Black-Scholes models $S_t^i = S_0^i \exp(\mu^i t + \sigma^i W_t^i)$. Here W^1 and W^2 are Brownian motions, of which it would be natural to assume that they are also jointly Gaussian with some correlation. Then the joint returns (R_t^1, R_t^2), for $R_t^i = \log S_{t+\delta t}^i / S_t^i$, will be bivariate Gaussian, and we can compute the VaR in terms of the parameters μ^i, σ^i and the correlation of R_t^1 and R_t^2, at least by computer simulation.

If, as before, we simplify by assuming that the returns, and not the log returns, are bivariate Gaussian, then we shall be in the situation of Section 9.2. The VaR is then given by (9.3) with μ_t and σ_t given by (9.4), where $w_t^i = \phi_t^i S_t^i / V_t$.

Alternatively, we may apply more realistic, but more complicated models. The conditional distribution of the loss will then typically be non-Gaussian, and not analytically tractable, but the Value-at-Risk can often be obtained easily by simulation.

Deriving VaR of portfolios of more than two stocks does not cause conceptual difficulties. However, making realistic models for the joint distribution of many equities is difficult. Gaussian models are a possibility, but unrealistic. Other standard models may include many parameters, that may be difficult to estimate.

9.4 Portfolios with Stock Options

In the Black-Scholes model a European call option with strike K and expiry time T has value

$$C_t = S_t \Phi\Big(\frac{\log(S_t/K) + (r+\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}\Big) \\ - Ke^{-r(T-t)}\Phi\Big(\frac{\log(S_t/K) + (r-\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}\Big).$$

The distribution of this random variable, or, more appropriately, the conditional distribution of $C_{t+\delta t}$ given $\mathcal{F}_t$, is not easily obtainable by analytic methods, but it is easy to simulate, being an explicit function of the stock price S_t (or $S_{t+\delta t}$) and a number of nonrandom quantities. Thus the VaR of a portfolio consisting of a European call option can be easily obtained numerically.

The value of some other European options on a given stock with price process S_t can also be written as an explicit function $C_t = F(t, S_t)$ of the stock price. The Value-at-Risk can then also be obtained by one round of computer simulation. Given $\mathcal{F}_t$ the stock price S_t is known and hence the gain $C_{t+\delta t} - C_t = F(t+\delta t, S_{t+\delta t}) - F(t, S_t)$ is stochastic only through $S_{t+\delta t}$. We can simulate the conditional distribution of the gain by simulating a sample from the conditional distribution of $S_{t+\delta t}$ given $\mathcal{F}_t$. In the Black-Scholes model we have that

$$S_{t+\delta t} = S_t e^{\mu\delta t + \sigma(W_{t+\delta t} - W_t)},$$

and hence an appropriate simulation scheme is to simulate a standard normal variable Z and next compute $S_t e^{\mu\delta t + \sigma\sqrt{\delta t} Z}$.

Even though this seems easy enough, in practice one often uses approximations of the form

$$C_{t+\delta t} - C_t = F(t+\delta t, S_{t+\delta t}) - F(t, S_t) \\ \approx F_t(t, S_t)\delta t + F_s(t, S_t)(S_{t+\delta t} - S_t) + \tfrac{1}{2} F_{ss}(t, S_t)(S_{t+\delta t} - S_t)^2.$$

The three partial derivatives on the right side are exactly the "Greeks" Θ, Δ, and Γ, already encountered in Section 6.5. For small δt, the increment $S_{t+\delta t} - S_t$ is typically of the order $O(\sqrt{\delta t})$ in probability, and hence the middle term on the right dominates. If we neglect the other two terms, then we arrive in the pleasant situation that the gain $C_{t+\delta t} - C_t$ is a linear transformation $\Delta\, S_t R_t$ of the return $R_t = S_{t+\delta t}/S_t - 1$. If we also assume that this return is conditionally normally distributed, then we are back in the situation of Section 9.2, and VaR takes a familiar form.

Options for which there is no explicit pricing formula are more difficult to incorporate. According to the general pricing theory, the value of an option with payment C at time T in the Black-Scholes model is equal to

$$C_t = \mathrm{E}_{\mathbb{Q}}\big(e^{-r(T-t)} C \,|\, \mathcal{F}_t\big).$$

For complicated claims C this could be determined numerically by simulation, this time under the martingale measure $\mathbb{Q}$. Combined with the preceding this leads to a double (nested) simulation scheme, sometimes referred to as the "full monte". We shall consider the most complicated case, that of a claim that depends on the full path $(S_t: 0 \leq t \leq T)$ of the stock price.

Given the information $\mathcal{F}_t$, the beginning $(S_s: 0 \leq s \leq t)$ of the sample path is known and hence given $\mathcal{F}_t$ the claim can be written as a function $C = h_t(S_s: t < s \leq T)$ of the future of the path. Therefore we can simulate the value of the option given $\mathcal{F}_t$ by simulating many times the future path $(S_s: t < s \leq T)$ from its conditional distribution given $\mathcal{F}_t$, evaluating each time $h_t(S_s: t < s \leq T)$, and finally taking the average of these values. In this round of simulations we simulate from the martingale measure. If the stock price process S is Markovian, then the conditional distribution of $(S_s: t < s \leq T)$ given $\mathcal{F}_t$ is the same as the conditional distribution of this process given S_t. For instance, for the Black-Scholes model we have

$$S_s = S_t e^{(r-\sigma^2/2)(s-t)+\sigma(\tilde{W}_s - \tilde{W}_t)},$$

for a Brownian motion $\tilde{W}$. The conditional distribution of the right side is determined from the conditional distribution of $\tilde{W}_s - \tilde{W}_t$ given $\mathcal{F}_t$, which is simply that of a Brownian motion. Note that we have taken the drift equal to $r - \sigma^2/2$, because we must simulate under the martingale measure.

We can now perform a nested sequence of simulations to determine the Value-at-Risk of a portfolio consisting of options and stocks as follows. We denote the true-world and martingale measures by $\mathbb{P}$ and $\mathbb{Q}$, respectively, and abbreviate the value of the claim C given the initial path $(S_s: 0 \leq s \leq t)$ by $h_t(S_s: t < s \leq T)$.

FOR (i in 1:MANY)
    {
    SIMULATE $(S_s^i: 0 \leq s \leq t + \delta t)$ ACCORDING TO $\mathbb{P}$
    FOR (j in 1:MANY)
        {
        GIVEN S_t^i SIMULATE $(S_s^j: t < s \leq T)$ ACCORDING TO $\mathbb{Q}$
        GIVEN $S_{t+\delta t}^i$ SIMULATE $(S_s^j: t + \delta t < s \leq T)$ ACCORDING TO $\mathbb{Q}$
        }
    COMPUTE C_t^i AS AVERAGE $h_t(S_s^j: t < s \leq T)$ OVER j
    COMPUTE $C_{t+\delta t}^i$ AS AVERAGE $h_{t+\delta t}(S_s^j: t + \delta t < s \leq T)$ OVER j
    }
VaR IS $1 - \alpha$ LARGEST OF THE VALUES $C_t^i - C_{t+\delta t}^i$.

This scheme is sufficiently involved that it will pay to use special computational techniques to make the simulations more efficient and more accurate.

9.5 Bond Portfolios

The value of a bond portfolio, consisting of bonds of different maturity, with or without coupons, is a linear combination $V_t = \sum_{i=1}^n \phi_t^i P_{t,T_i}$ of discount bond prices P_{t,T_i}. A term structure model gives exactly the joint distribution of the zero-coupon bonds. Thus in principle every term structure model allows to calculate the VaR of the portfolio, if necessary by simulation. In the present situation we need the term structure model under the true world measure $\mathbb{P}$!

As a particular example, consider the Hull-White model. In this model the bond prices are given by an explicit formula of the form

$$P_{t,T} = e^{A(t,T) - B(t,T)r_t}.$$

Thus we can simulate the bond price at time t by simulating the short rate r_t. Under the martingale measure $\mathbb{Q}$ the short rate is the sum of a deterministic function and an Ornstein-Uhlenbeck process. Unfortunately, to compute VaR we need to simulate under the true-world measure $\mathbb{P}$, which may add a random drift to the Ornstein-Uhlenbeck process. This may destroy the Gaussianity and other nice properties, which the short rate process possesses under the martingale measure.

More specifically, in the Hull-White model the short rate satisfies formula (8.14), which can be written in the form

$$r_t = \alpha(t) + \sigma e^{-at} \int_0^t e^{as}\, dW_s,$$

for α the deterministic function $\alpha(t) = e^{-at} r_0 + e^{-at} \int_0^t \theta(s) e^{as}\, ds$. The function θ in this expression can be found by calibration on option prices observed in the market, and so can the parameters a and σ. However, the process W in the preceding display is a Brownian motion under the martingale measure $\mathbb{Q}$, and not under $\mathbb{P}$. In agreement with Girsanov's theorem, under $\mathbb{P}$ the process W is a Brownian motion with drift, and $W_t - \int_0^t \lambda(s, r_s)\, ds$ for λ the "market price of risk" is a $\mathbb{P}$-Brownian motion. There is no way the market price of risk can be calibrated from derivative prices alone, but it can be determined by historical analysis. If $\lambda(t, r_t)$ does not depend on r_t, then the change of drift only changes the deterministic function α in the preceding display, and we can use the simulation scheme for Ornstein-Uhlenbeck processes discussed in Section 8.4 and Theorem 8.22 to generate r_t, the bond prices, and hence the VaR. If $\lambda(t, r_t)$ is random, then the drift of the short rate process under $\mathbb{P}$ is random, and we must fall back on the more complicated simulation schemes for diffusion processes, such as the Euler scheme discussed in Section 8.4. Rather than calibrate a and θ from observed derivative prices on the market, we may then also choose to fit the diffusion model

$$dr_t = \mu(t, r_t)\, dt + \sigma\, dW_t,$$

with W a $\mathbb{P}$-Brownian motion, directly to historical data. Note that the volatility parameter σ is common to both the $\mathbb{P}$ and $\mathbb{Q}$ models and hence can both be calibrated and estimated.

This discussion of the Hull-White model extends to any model where the bond prices are simple functions $P_{t,T} = F^T(t, r_t)$ of the short rate, and more generally to *multi-factor models* in which the bond prices $P_{t,T} = F^T(t, X_t)$ can be written as a simple function of multivariate diffusion process X.

In practice one often uses simpler approaches based on approximations and the assumption that the yields are multivariate normal. Given a bond portfolio with value process $V_t = \sum_{i=1}^n \phi_t^i P_{t,T_i}$, where ϕ_t^i is assumed constant in $[t, t+\delta t]$, we first approximate

$$\begin{aligned}
V_{t+\delta t} - V_t &= \sum_{i=1}^n \phi_t^i (P_{t+\delta t,T_i} - P_{t,T_i}) \\
&= \sum_{i=1}^n \phi_t^i \Big(e^{-(T_i - t - \delta t) Y_{t+\delta t, T_i}} - e^{-(T_i - t) Y_{t,T_i}} \Big) \\
&\approx \sum_{i=1}^n \phi_t^i P_{t,T_i} \Big[-(T_i - t)(Y_{t+\delta t,T_i} - Y_{t,T_i}) + \delta t Y_{t+\delta t,T_i} \\
&\qquad\qquad + \tfrac{1}{2}(T_i - t)^2 (Y_{t+\delta t,T_i} - Y_{t,T_i})^2 \Big].
\end{aligned}$$

In the last step we use the approximation $e^y - e^x \approx e^x(y - x + \frac{1}{2}(y-x)^2)$ for $y \approx x$. The derivatives in the linear and quadratic parts are known as the *duration* and the *convexity*, respectively. The conditional distribution of the right side given $\mathcal{F}_t$ can be evaluated once we have a model for the conditional joint distribution of the vector of yield increments $(Y_{t+\delta t,T_i} - Y_{t,T_i})$ given $\mathcal{F}_t$. In practice this is often modelled by a mean zero multivariate normal distribution. If we neglect the quadratic term ("use only duration"), then we fall back in the situation of Section 9.2. VaR then takes the form as given in (9.3):

$$\mathrm{VaR} = \Phi^{-1}(1-\alpha) \sqrt{\sum_{i=1}^n \sum_{j=1}^n \phi_t^i \phi_t^j P_{t,T_i} P_{t,T_j} (T_i - t)(T_j - t) \sigma_t^{i,j}}.$$

VaR based on using both duration and convexity in the approximation can be determined by simulation.

9.6 Portfolios of Bonds and Swaptions

A swaption relative to the swap times $T_1 < T_2 < \cdots < T_n$ pays the amount $\big(P_{T_0,T_n} + K \sum_{i=1}^n P_{T_0,T_i} - 1\big)^+$ at time $T_0 < T_1$. Depending on the term structure model used, there may or may not be an explicit formula for the value of a swaption at time $t < t_0$. If there is, then determining the VaR does not present great difficulties.

In general, the value at time $t < T_0$ can be evaluated as

$$C_t = \mathrm{E}_{\mathbb{Q}}\Big(e^{-\int_t^{T_0} r_s\,ds}\big(P_{T_0,T_n} + K\sum_{i=1}^{n} P_{T_0,T_i} - 1\big)^+ \,|\, \mathcal{F}_t\Big).$$

In short rate models, such as the Hull-White model, the variable inside the expectation can be written as a function of the short rate process $(r_s\colon 0 \leq t \leq T_n)$. Given $\mathcal{F}_t$ the initial path $(r_s\colon 0 \leq s \leq t)$ is known and hence we can evaluate the price C_t by computing an expectation under $\mathbb{Q}$ of a function $h_t(r_s\colon t \leq s \leq T)$ of the future sample path, given its present state r_t. For instance, in the Hull-White model C_t is given by

$$\mathrm{E}_{\mathbb{Q}}\Big[e^{-\int_t^{T_0} r_s\,ds}\Big(e^{A(T_0,T_n)-B(T_0,T_n)r_{T_0}} + K\sum_{i=1}^{n} e^{A(T_0,T_i)-B(T_0,T_i)r_{T_0}} - 1\Big)^+ \,|\, r_0\Big].$$

In this particular case it is possible to evaluate the expectation analytically (at least for $n = 1$), using the approach of Lemma 8.12. In general, we need to use simulation or approximations.

To compute the VaR of a swaption portfolio, we need a double, nested round of simulations, one under the true-world measure and one under the martingale measure. If the bond price P_{t,T_i} can be expressed in r_{T_i} and the value of the swaption at time t can be written as

$$\mathrm{E}_{\mathbb{Q}}\big(h_t(r_s\colon t < s \leq T_0) \,|\, r_t\big),$$

then a simple (but computationally inefficient) scheme is as follows.

```
FOR (i in 1:MANY)
    {
    SIMULATE (r_s: 0 ≤ s ≤ t + δt) ACCORDING TO ℙ
    FOR (j in 1:MANY)
        {
        GIVEN r_t^i SIMULATE (r_s^j: t < s ≤ T_0) ACCORDING TO ℚ
        GIVEN r_{t+δt}^i SIMULATE (r_s^j: t + δt < s ≤ T_n) ACCORDING TO ℚ
        }
    COMPUTE C_t^i AS AVERAGE h_t(r_s^j: t < s ≤ T_0) OVER j
    COMPUTE C_{t+δt}^i AS AVERAGE h_{t+δt}(r_s^j: t + δt < s ≤ T_0) OVER j
    }
VaR IS 1 − α LARGEST OF THE VALUES C_t^i − C_{t+δt}^i.
```

Rather than using this double simulation scheme we may here also apply approximations. However, in general it is not easy to compute partial derivatives of the value process relative to e.g. the yield ("duration" and "convexity"), and we may need to use numerical (i.e. discretized) derivatives instead.

9.7 Diversified Portfolios

In the preceding section we have considered a variety of portfolios. The balance sheet of a large financial institution will typically include a combination of the assets considered so far. To compute the Value-at-Risk we can use the same arguments, with the important complication that we need models for the joint distribution of the various assets. For instance, the joint distribution of stocks on Philips and IBM, and bonds of various maturities. There is no standard approach to this.