Continuous Time Finance

2017

Tomas Björk

Stockholm School of Economics

Contents

Textbook:
Björk, T: "Arbitrage Theory in Continuous Time"
Oxford University Press, 2009. (3:rd ed.)

Chapter 4

Stochastic Integrals

Tomas Björk

Typical Setup

Take as given the market price process, S, of some underlying asset.

$$S_t = \text{price, at } t, \text{ per unit of underlying asset}$$

Consider a fixed **financial derivative**, e.g. a European call option.

Main Problem: Find the arbitrage free price of the derivative.

We Need:

1. Mathematical model for the underlying price process. (The Black-Scholes model)

2. Mathematical techniques to handle the price dynamics. (The Itô calculus.)

Stochastic Processes

- We model the stock price S_t as a **stochastic process**, i.e. it **evolves randomly over time**.

- We model S as a **Markov process**, i.e. in order to predict the future only the present value is of interest. All past information is already incorporated into today's stock prices. (Market efficiency).

Stochastic variable
Choosing a **number** at random

Stochastic process
choosing a **curve** (trajectory) at random.

Notation

$$\begin{aligned} X_t &= \text{any random process,} \\ dt &= \text{small time step,} \\ dX_t &= X_{t+dt} - X_t \end{aligned}$$

- We often write $X(t)$ instead of X_t.
- dX_t is called the **increment** of X over the interval $[t, t + dt]$.
- For any fixed interval $[t, t + dt]$, the increment dX_t is a stochastic variable.
- If the increments dX_s and dX_t, over the disjoint intervals $[s, s + ds]$ and $[t, t + dt]$ are independent, then we say that X has **independent increments**.
- If every increment has a normal distribution we say that X is a **normal**, or **Gaussian** process.

The Wiener Process

A stochastic process W is called a **Wiener process** if it has the following properties

- The increments are normally distributed: For $s < t$:

$$W_t - W_s \sim N[0, t-s]$$

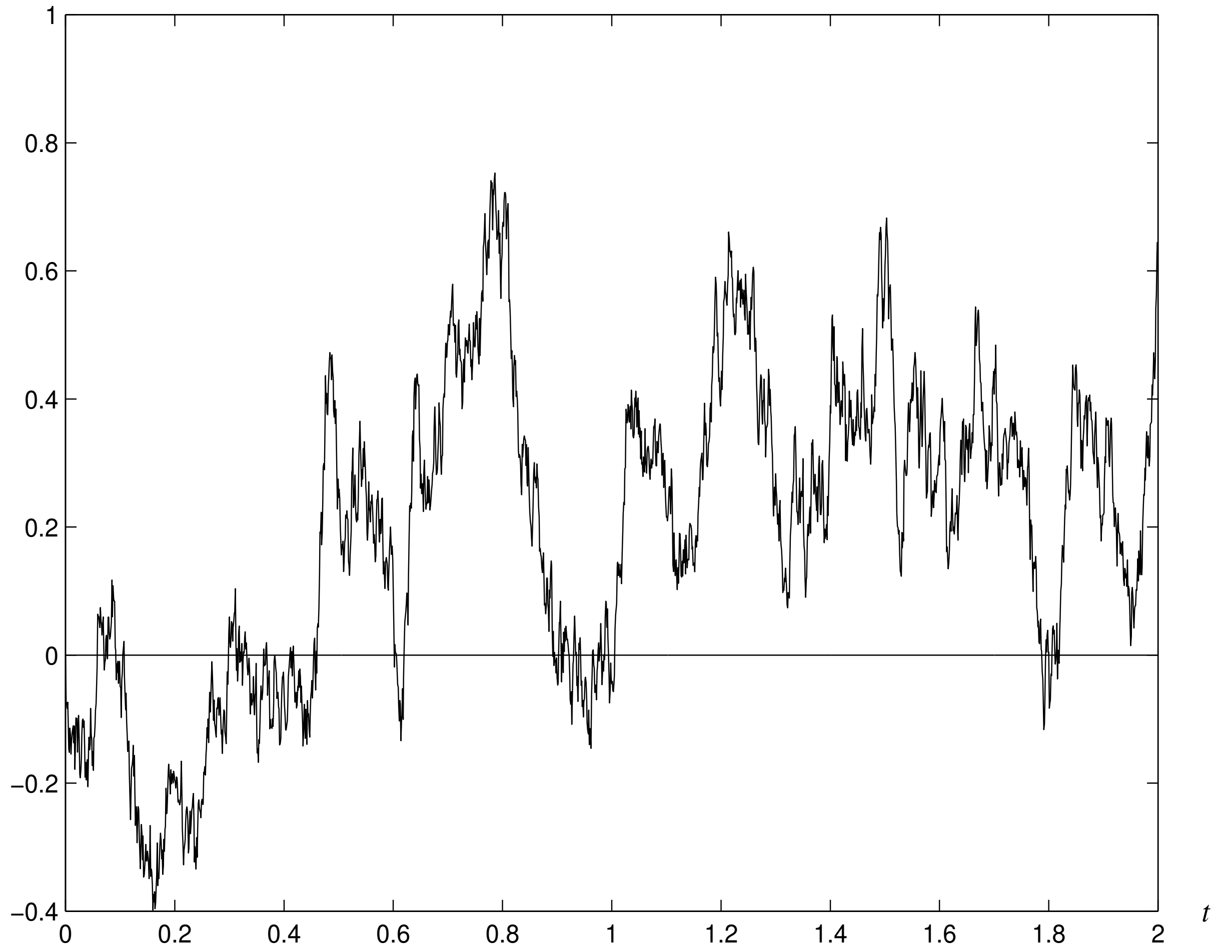
$$E[W_t - W_s] = 0, \quad Var[W_t - W_s] = t - s$$

- W has independent increments.
- $W_0 = 0$
- W has continuous trajectories.

Continuous random walk

Note: In Hull, a Wiener process is typically denoted by Z instead of W.

A Wiener Trajectory



Important Fact

Theorem:
A Wiener trajectory is, with probability one, a continuous curve which is **nowhere differentiable**.

Proof. Hard.

Wiener Process with Drift

A stochastic process X is called a Wiener process with **drift** μ and **diffusion coefficient** σ if it has the following dynamics

$$dX_t = \mu dt + \sigma dW_t,$$

where μ and σ are constants.

Summing all increments over the interval $[0, t]$ gives us

$$X_t - X_0 = \mu \cdot t + \sigma \cdot (W_t - W_0),$$

$$X_t = X_0 + \mu t + \sigma W_t$$

Thus

$$X_t \sim N[X_0 + \mu t, \ \sigma^2 t]$$

Itô processes

We say, losely speaking, that the process X is an **Itô process** if it has dynamics of the form

$$dX_t = \mu_t dt + \sigma_t dW_t,$$

where μ_t and σ_t are random processes.

Informally you can think of dW_t as a random variable of the form

$$dW_t \sim N[0, dt]$$

To handle expressions like the one above, we need some **mathematical theory**.

First, however, we present an important example, which we will discuss informally.

Example: The Black-Scholes model

Price dynamics: (Geometrical Brownian Motion)

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

Simple analysis:
Assume that $\sigma = 0$. Then

$$dS_t = \mu S_t dt$$

Divide by dt!

$$\frac{dS_t}{dt} = \mu S_t$$

This is a simple ordinary differential equation with solution

$$S_t = s_0 e^{\mu t}$$

Conjecture: The solution of the SDE above is a randomly disturbed exponential function.

Intuitive Economic Interpretation

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

Over a small time interval $[t, t + dt]$ this means:

$$\begin{aligned} \text{Return} &= \text{(mean return)} \\ &+ \sigma \times \text{(Gaussian random disturbance)} \end{aligned}$$

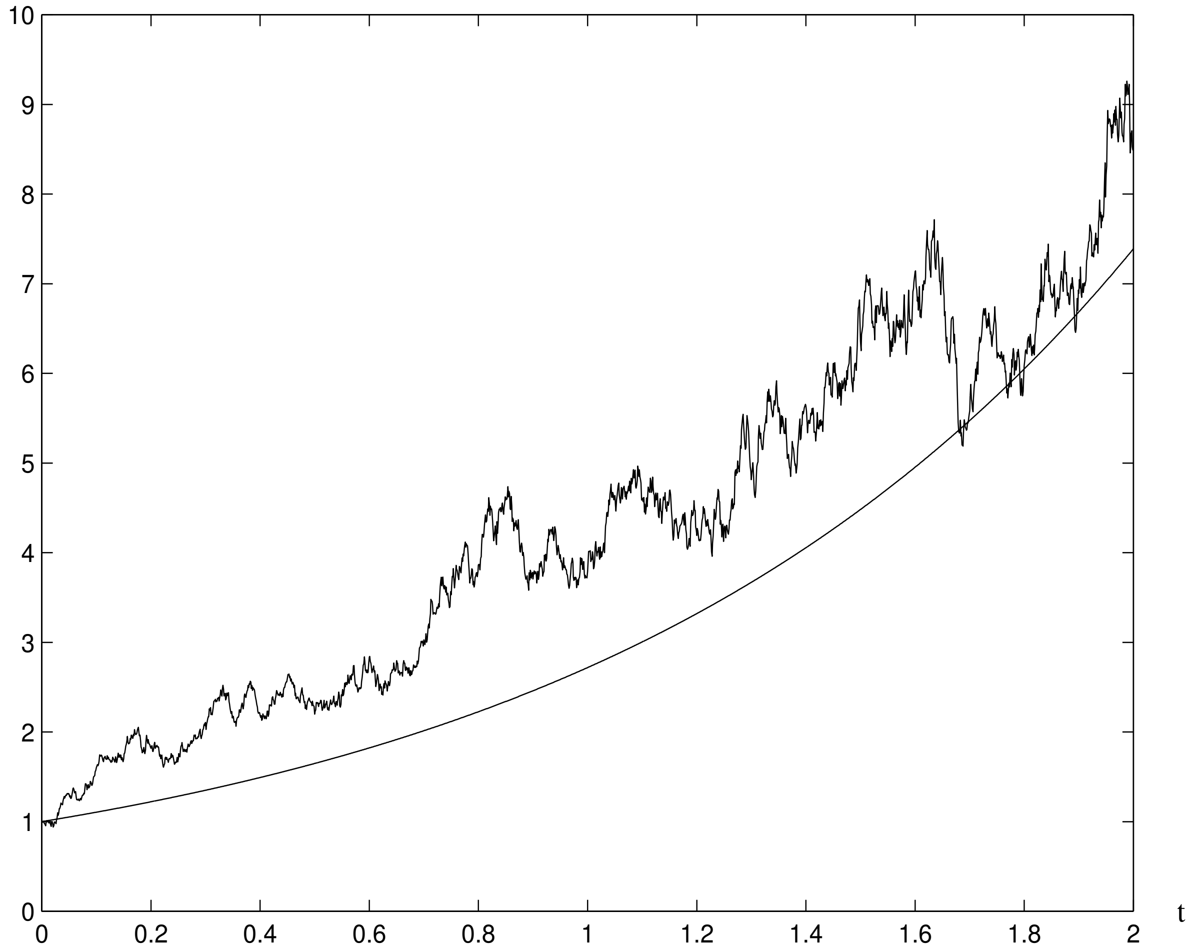
- The asset **return** is a random walk (with drift).
- $\mu =$ mean rate of return per unit time
- $\sigma =$ volatility per unit time

Large $\sigma =$ large random fluctuations

Small $\sigma =$ small random fluctuations

- The returns are normal.
- The stock price is lognormal.

A GBM Trajectory



Stochastic Differentials and Integrals

Consider an expression of the form

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t dW_t, \\ X_0 &= x_0 \end{aligned}$$

Question: What **exactly** do we mean by this?

Answer: Write the equation on integrated form as

$$X_t = x_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$$

How is this interpreted?

Recall:

$$X_t = x_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$$

Two terms:

- $$\int_0^t \mu_s ds$$
  This is a standard Riemann integral for each μ-trajectory.

- $$\int_0^t \sigma_s dW_s$$
  Stochastic integral. This can **not** be interpreted as a Stieljes integral for each trajectory. We need a new theory for this **Itô integral**.

Information

Consider a process X.

Def:

$$\mathcal{F}_t^W = \text{“The information generated by } X \text{ over the interval } [0,t]\text{”}$$

Def: Let Z be a stochastic variable. If the value of Z is completely determined by $\mathcal{F}_t^X$, we write

$$Z \in \mathcal{F}_t^X$$

Ex:
For the stochastic variable Z, defined by

$$Z = \int_0^5 X_s ds,$$

we have $Z \in \mathcal{F}_5^X$.

We do **not** have $Z \in \mathcal{F}_4^X$.

Adapted Processes

Let X be a random process.

Definition:
A process Y is **adapted** to the filtration $\left\{\mathcal{F}_t^X :\ t \geq 0\right\}$ if

$$Y_t \in \mathcal{F}_t^X, \quad \forall t \geq 0$$

"**An adapted process does not look into the future**"

Adapted processes are nice integrands for stochastic integrals.

- The process

$$Y_t = \int_0^t X_s ds,$$

is adapted.

- The process

$$Y_t = \sup_{s \le t} X_s$$

is adapted.

- The process

$$Y_t = \sup_{s \le t+1} X_s$$

is **not** adapted.

The Itô Integral

Consider a Wiener process W. We will define the Itô integral

$$\int_a^b g_s dW_s$$

for processes g satisfying

- The process g is adapted to $\mathcal{F}_t^W$
- The process g satisfies

$$\int_a^b E\left[g_s^2\right] ds < \infty$$

This will be done in two steps.

Simple Integrands

Definition: The process g is **simple**, if

- g is adapted.
- There exists deterministic points $t_0 \dots, t_n$ with $a = t_0 < t_1 < \dots < t_n = b$ such that g is piecewise constant, i.e.

$$g(s) = g(t_k), \quad s \in [t_k, t_{k+1})$$

For simple g we define

$$\int_a^b g_s dW_s = \sum_{k=0}^{n-1} g(t_k) \left[W(t_{k+1}) - W(t_k) \right]$$

FORWARD INCREMENTS!

Properties of the Integral

Theorem: For simple g the following relations hold

- The expected value is given by

$$E\left[\int_a^b g_s dW_s\right] = 0$$

- The second moment is given by

$$E\left[\left(\int_a^b g_s dW_s\right)^2\right] = \int_a^b E\left[g_s^2\right] ds$$

- We have

$$\int_a^b g_s dW_s \in \mathcal{F}_b^W$$

General Case

For a general g we do as follows.

1. Approximate g with a sequence of simple g_n such that
$$\int_a^b E\left[\{g_n(s) - g(s)\}^2\right] ds \to 0.$$

2. For each n the integral
$$\int_a^b g_n(s) dW(s)$$
is a well defined stochastic variable Z_n.

3. One can show that the Z_n sequence converges to a limiting stochastic variable.

4. We define $\int_a^b g dW$ by
$$\int_a^b g(s) dW(s) = \lim_{n \to \infty} \int_a^b g_n(s) dW(s).$$

Properties of the Integral

Theorem: For general g following relations hold

- The expected value is given by

$$E\left[\int_a^b g_s dW_s\right] = 0$$

- We do in fact have

$$E\left[\int_a^b g_s dW_s \middle| \mathcal{F}_a\right] = 0$$

- The second moment is given by

$$E\left[\left(\int_a^b g_s dW_s\right)^2\right] = \int_a^b E\left[g_s^2\right] ds$$

- We have

$$\int_a^b g_s dW_s \in \mathcal{F}_b^W$$

Martingales

Definition: An adapted process X is a **martingale** if

$$E\left[X_t \middle| \mathcal{F}_s\right] = X_s, \quad \forall s \leq t$$

"A martingale is a process without drift"

Proposition: For any g (sufficiently integrable) the process

$$X_t = \int_0^t g_s dW_s$$

is a martingale.

Proposition: If X has dynamics

$$dX_t = \mu_t dt + \sigma_t dW_t$$

then X is a martingale **iff** $\mu = 0$.

Chapters 4-5

Stochastic Calculus

Tomas Björk

Stochastic Calculus

General Model:

$$dX_t = \mu_t dt + \sigma_t dW_t$$

Let the function $f(t, x)$ be given, and define the stochastic process Z_t by

$$Z_t = f(t, X_t)$$

Problem: What does $df(t, X_t)$ look like?

The answer is given by the **Itô formula**.

We provide an intuitive argument. The formal proof is very hard.

A close up of the Wiener process

Consider an "infinitesimal" Wiener increment

$$dW_t = W_{t+dt} - W_t$$

We know:

$$dW_t \quad \sim \quad N[0, dt]$$

$$E[dW_t] \quad = \quad 0, \quad Var[dW_t] = dt$$

From this one can show

$$E[(dW_t)^2] = dt, \quad Var[(dW_t)^2] = 2(dt)^2$$

Recall

$$E[(dW_t)^2] = dt, \quad Var[(dW_t)^2] = 2(dt)^2$$

Important observation:

1. Both $E[(dW_t)^2]$ and $Var[(dW_t)^2]$ are very small when dt is small .

2. $Var[(dW_t)^2]$ is negligable compared to $E[(dW_t)^2]$.

3. Thus $(dW_t)^2$ is **deterministic**.

We thus conclude, at least intuitively, that

$$(dW_t)^2 = dt$$

This was only an intuitive argument, but it can be proved rigorously.

Multiplication table.

Theorem: We have the following multiplication table

$$
\begin{aligned}
(dt)^2 &= 0 \\
dW_t \cdot dt &= 0 \\
(dW_t)^2 &= dt
\end{aligned}
$$

Deriving the Itô formula

$$
\begin{aligned}
dX_t &= \mu_t dt + \sigma_t dW_t \\
Z_t &= f(t, X_t)
\end{aligned}
$$

We want to compute $df(t, X_t)$

Make a Taylor expansion of $f(t, X_t)$ including second order terms:

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

Plug in the expression for dX, expand, and use the multiplication table!

$$
\begin{aligned}
df &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}[\mu dt + \sigma dW] + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 \\
&+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}[\mu dt + \sigma dW]^2 + \frac{\partial^2 f}{\partial t \partial x}dt \cdot [\mu dt + \sigma dW] \\
&= \frac{\partial f}{\partial t}dt + \mu\frac{\partial f}{\partial x}dt + \sigma\frac{\partial f}{\partial x}dW + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 \\
&+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}[\mu^2(dt)^2 + \sigma^2(dW)^2 + 2\mu\sigma dt \cdot dW] \\
&+ \mu\frac{\partial^2 f}{\partial t \partial x}(dt)^2 + \sigma\frac{\partial^2 f}{\partial t \partial x}dt \cdot dW
\end{aligned}
$$

Using the multiplikation table this reduces to:

$$
\begin{aligned}
df &= \left\{\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}\right\}dt \\
&+ \sigma\frac{\partial f}{\partial x}dW
\end{aligned}
$$

The Itô Formula

Theorem: With X dynamics given by

$$dX_t = \mu_t dt + \sigma_t dW_t$$

we have

$$\begin{aligned} df(t, X_t) &= \left\{ \frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} \right\} dt \\ &+ \sigma \frac{\partial f}{\partial x} dW_t \end{aligned}$$

Alternatively

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

where we use the multiplication table.

Example: GBM

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

We smell something exponential!

Natural Ansatz:

$$\begin{aligned} S_t &= e^{Z_t}, \\ Z_t &= \ln S_t \end{aligned}$$

Itô on $f(t,s) = \ln(s)$ gives us

$$\frac{\partial f}{\partial s} = \frac{1}{s}, \quad \frac{\partial f}{\partial t} = 0, \quad \frac{\partial^2 f}{\partial s^2} = -\frac{1}{s^2}$$

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

Recall

$$dZ_t = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t$$

Integrate!

$$\begin{aligned} Z_t - Z_0 &= \int_0^t \left(\mu - \frac{1}{2}\sigma^2\right) ds + \sigma \int_0^t dW_s \\ &= \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t \end{aligned}$$

Using $S_t = e^{Z_t}$ gives us

$$S_t = S_0 e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t}$$

Since W_t is $N[0, t]$, we see that S_t has a lognormal distribution.

A Useful Trick

Problem: Compute $E\left[Z(T)\right]$.

- Use Itô to get

$$dZ(t) = \mu_Z(t)dt + \sigma_Z(t)dW_t$$

- Integrate.

$$Z(T) = z_0 + \int_0^T \mu_Z(t)dt + \int_0^T \sigma_Z(t)dW_t$$

- Take expectations.

$$E\left[Z(T)\right] = z_0 + \int_0^T E\left[\mu_Z(t)\right] dt + 0$$

- The problem has been reduced to that of computing $E\left[\mu_Z(t)\right]$.

The Connection SDE $\sim$ PDE

Given: $\mu(t,x), \quad \sigma(t,x), \quad \Phi(x), \quad T$

Problem: Find a function F solving the Partial Differential Equation (PDE)

$$
\begin{aligned}
\frac{\partial F}{\partial t}(t,x) + \mathcal{A}F(t,x) &= 0, \\
F(T,x) &= \Phi(x).
\end{aligned}
$$

where $\mathcal{A}$ is defined by

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

Assume that F solves the PDE.

Fix the point (t, x).

Define the process X by

$$\begin{aligned} dX_s &= \mu(s, X_s)dt + \sigma(s, X_s)dW_s, \\ X_t &= x, \end{aligned}$$

Apply Ito to the process $F(t, X_t)$!

$$\begin{aligned} F(T, X_T) &= F(t, X_t) \\ &+ \int_t^T \left\{ \frac{\partial F}{\partial t}(s, X_s) + \mathcal{A}F(s, X_s) \right\} ds \\ &+ \int_t^T \sigma(s, X_s) \frac{\partial F}{\partial x}(s, X_s) dW_s. \end{aligned}$$

By assumption $\frac{\partial F}{\partial t} + \mathcal{A}F = 0$, and $F(T, x) = \Phi(x)$

Thus:

$$\begin{aligned} \Phi(X_T) &= F(t,x) \\ &+ \int_t^T \sigma(s, X_s) \frac{\partial F}{\partial x}(s, X_s) dW_s. \end{aligned}$$

Take expectations.

$$F(t,x) = E_{t,x}\left[\Phi\left(X_T\right)\right],$$

Feynman-Kac

The solution $F(t,x)$ to the PDE

$$\begin{aligned}\frac{\partial F}{\partial t}+\mu(t,x)\frac{\partial F}{\partial x}+\frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F}{\partial x^2}-rF &= 0,\\ F(T,x) &= \Phi(x).\end{aligned}$$

is given by

$$F(t,x)=e^{-r(T-t)}E_{t,x}\left[\Phi\left(X_T\right)\right],$$

where X satisfies the SDE

$$\begin{aligned}dX_s &= \mu(s,X_s)dt+\sigma(s,X_s)dW_s,\\ X_t &= x.\end{aligned}$$

Chapters 6-7

Black-Scholes

Tomas Björk

Contents

1.

Introduction

European Call Option

The holder of this paper has the right

to buy

1 ACME INC

on the date

July 30, 2018

at the price

\$100

Financial Derivative

- A financial asset which is defined **in terms of** some **underlying** asset.
- Future stochastic claim.

Examples

- European calls and puts
- American options
- Forward rate agreements
- Convertibles
- Futures
- Bond options
- Caps & Floors
- Interest rate swaps
- CDO:s
- CDS:s

Main problems

- What is a "reasonable" price for a derivative?
- How do you hedge yourself against a derivative.

Natural Answers

Consider a random cash payment $\mathcal{Z}$ at time T.

What is a reasonable price $\Pi_0[\mathcal{Z}]$ at time 0?

Natural answers:

1. Price = Discounted present value of future payouts.

$$\Pi_0[\mathcal{Z}] = e^{-rT} E[\mathcal{Z}]$$

2. The question is meaningless.

Both answers are incorrect!

- Given some assumptions we **can** really talk about “the correct price” of an option.

- The correct pricing formula is **not** the one on the previous slide.

Philosophy

- The derivative is **defined in terms of** underlying.
- The derivative can be **priced in terms of** underlying price.
- **Consistent** pricing.
- **Relative** pricing.

Before we can go on further we need some simple portfolio theory

2.

Portfolio Theory

Portfolios

We consider a market with N assets.

$$S_t^i = \text{price at } t\text{, of asset No } i.$$

A **portfolio** strategy is an adapted vector process

$$h_t = (h_t^1, \cdots, h_t^N)$$

where

$$\begin{aligned} h_t^i &= \text{number of units of asset } i, \\ V_t &= \text{market value of the portfolio} \end{aligned}$$

$$V_t = \sum_{i=1}^{N} h_t^i S_t^i$$

The portfolio is typically of the form

$$h_t = h(t, S_t)$$

i.e. today's portfolio is based on today's prices.

Self financing portfolios

We want to study self financing portfolio strategies, i.e. portfolios where purchase of a "new" asset must be financed through sale of an "old" asset.

How is this formalized?

Definition:
The strategy h is **self financing** if

$$dV_t = \sum_{i=1}^{N} h_t^i dS_t^i$$

Interpret!

See Appendix B for details.

Relative weights

Definition:

$\omega_t^i =$ relative portfolio weight on asset No i.

We have

$$\omega_t^i = \frac{h_t^i S_t^i}{V_t}$$

Insert this into the self financing condition

$$dV_t = \sum_{i=1}^{N} h_t^i dS_t^i$$

We obtain

Portfolio dynamics:

$$dV_t = V_t \sum_{i=1}^{N} \omega_t^i \frac{dS_t^i}{S_t^i}$$

Interpret!

3.

Deriving the Black-Scholes PDE

Back to Financial Derivatives

Consider the Black-Scholes model

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

We want to price a European call with strike price K and exercise time T. This is a stochastic claim on the future. The future pay-out (at T) is a stochastic variable, $\mathcal{Z}$, given by

$$\mathcal{Z} = \max[S_T - K, 0]$$

More general:

$$\mathcal{Z} = \Phi(S_T)$$

for some contract function Φ.

Main problem: What is a "reasonable" price, $\Pi_t[\mathcal{Z}]$, for $\mathcal{Z}$ at t?

Main Idea

- We demand **consistent** pricing between derivative and underlying.
- No **mispricing** between derivative and underlying.
- No **arbitrage possibilities** on the market (B, S, Π)

Arbitrage

The portfolio ω is an **arbitrage** portfolio if

- The portfolio strategy is self financing.
- $V_0 = 0$.
- $V_T > 0$ with probability one.

Moral:

- **Arbitrage = Free Lunch**
- **No arbitrage possibilities in an efficient market.**

Arbitrage test

Suppose that a portfolio ω is self financing whith dynamics

$$dV_t = kV_t dt$$

- No driving Wiener process
- Risk free rate of return.
- "Synthetic bank" with rate of return k.

If the market is free of arbitrage we must have:

$$k = r$$

Main Idea of Black-Scholes

- Since the derivative is defined in terms of the underlying, the derivative price should be highly correlated with the underlying price.

- We should be able to balance dervative against underlying in our portfolio, so as to cancel the randomness.

- Thus we will obtain a riskless rate of return k on our portfolio.

- Absence of arbitrage must imply

$$k = r$$

Two Approaches

The program above can be formally carried out in two slightly different ways:

- The way Black-Scholes did it in the original paper. This leads to some logical problems.

- A more conceptually satisfying way, first presented by Merton.

Here we use the Merton method. You will find the original BS method in Appendix C at the end of this lecture.

Quiz: What is the problem with the original B-S argument?

Formalized program a la Merton

- Assume that the derivative price is of the form

$$\Pi_t\left[\mathcal{Z}\right] = f(t, S_t).$$

- Form a portfolio based on the underlying S and the derivative f, with portfolio dynamics

$$dV_t = V_t \left\{ \omega_t^S \cdot \frac{dS_t}{S_t} + \omega_t^f \cdot \frac{df}{f} \right\}$$

- Choose ω^S and ω^f such that the dW-term is wiped out. This gives us

$$dV_t = V_t \cdot k dt$$

- Absence of arbitrage implies

$$k = r$$

- This relation will say something about f.

Back to Black-Scholes

$$
\begin{aligned}
dS_t &= \mu S_t dt + \sigma S_t dW_t, \\
\Pi_t [\mathcal{Z}] &= f(t, S_t)
\end{aligned}
$$

Itô's formula gives us the f dynamics as

$$
\begin{aligned}
df &= \left\{ \frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial s} + \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial s^2} \right\} dt \\
&+ \sigma S \frac{\partial f}{\partial s} dW
\end{aligned}
$$

Write this as

$$
df = \mu_f \cdot f dt + \sigma_f \cdot f dW
$$

where

$$
\begin{aligned}
\mu_f &= \frac{\frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial s} + \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial s^2}}{f} \\
\sigma_f &= \frac{\sigma S \frac{\partial f}{\partial s}}{f}
\end{aligned}
$$

$$df = \mu_f \cdot f dt + \sigma_f \cdot f dW$$

$$\begin{aligned} dV &= V \left\{ \omega^S \cdot \frac{dS}{S} + \omega^f \cdot \frac{df}{f} \right\} \\ &= V \left\{ \omega^S (\mu dt + \sigma dW) + \omega^f (\mu_f dt + \sigma_f dW) \right\} \end{aligned}$$

$$dV = V \left\{ \omega^S \mu + \omega^f \mu_f \right\} dt + V \left\{ \omega^S \sigma + \omega^f \sigma_f \right\} dW$$

Now we kill the dW-term!

Choose (ω^S, ω^f) such that

$$\begin{aligned} \omega^S \sigma + \omega^f \sigma_f &= 0 \\ \omega^S + \omega^f &= 1 \end{aligned}$$

Linear system with solution

$$\omega^S = \frac{\sigma_f}{\sigma_f - \sigma}, \quad \omega^f = \frac{-\sigma}{\sigma_f - \sigma}$$

Plug into dV!

We obtain

$$dV = V \left\{ \omega^S \mu + \omega^f \mu_f \right\} dt$$

This is a risk free "synthetic bank" with short rate

$$\left\{ \omega^S \mu + \omega^f \mu_F \right\}$$

.

Absence of arbitrage implies

$$\left\{ \omega^S \mu + \omega^f \mu_f \right\} = r$$

Plug in the expressions for ω^S, ω^f, μ_f and simplify. This will give us the following result.

Black-Schole's PDE

The price is given by

$$\Pi_t [\mathcal{Z}] = f(t, S_t)$$

where the pricing function f satisfies the PDE (partial differential equation)

$$\begin{cases} \dfrac{\partial f}{\partial t}(t,s) + rs\dfrac{\partial f}{\partial s}(t,s) + \dfrac{1}{2}\sigma^2 s^2 \dfrac{\partial^2 f}{\partial s^2}(t,s) - rf(t,s) & = & 0 \\ f(T,s) & = & \Phi(s) \end{cases}$$

There is a unique solution to the PDE so there is a unique arbitrage free price process for the contract.

Black-Scholes' PDE ct'd

$$\begin{cases} \dfrac{\partial f}{\partial t} + rs\dfrac{\partial f}{\partial s} + \dfrac{1}{2}\sigma^2 s^2 \dfrac{\partial^2 f}{\partial s^2} - rf & = & 0 \\ f(T,s) & = & \Phi(s) \end{cases}$$

- The price of **all** derivative contracts have to satisfy the **same** PDE

$$\frac{\partial f}{\partial t} + rs\frac{\partial f}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 f}{\partial s^2} - rf = 0$$

otherwise there will be an arbitrage opportunity.

- The only difference between different contracts is in the boundary value condition

$$f(T,s) = \Phi(s)$$

Data needed

- The contract function Φ.
- Today's date t.
- Today's stock price S.
- Short rate r.
- Volatility σ.

Note: The pricing formula does **not** involve the mean rate of return μ!

??

Black-Scholes Basic Assumptions

Assumptions:

- The stock price is Geometric Brownian Motion
- Continuous trading.
- Frictionless efficient market.
- Short positions are allowed.
- Constant volatility σ.
- Constant short rate r.
- Flat yield curve.

Black-Scholes' Formula
European Call

T=date of expiration,
t=today's date,
K=strike price,
r=short rate,
s=today's stock price,
σ=volatility.

$$f(t,s) = sN[d_1] - e^{-r(T-t)}KN[d_2].$$

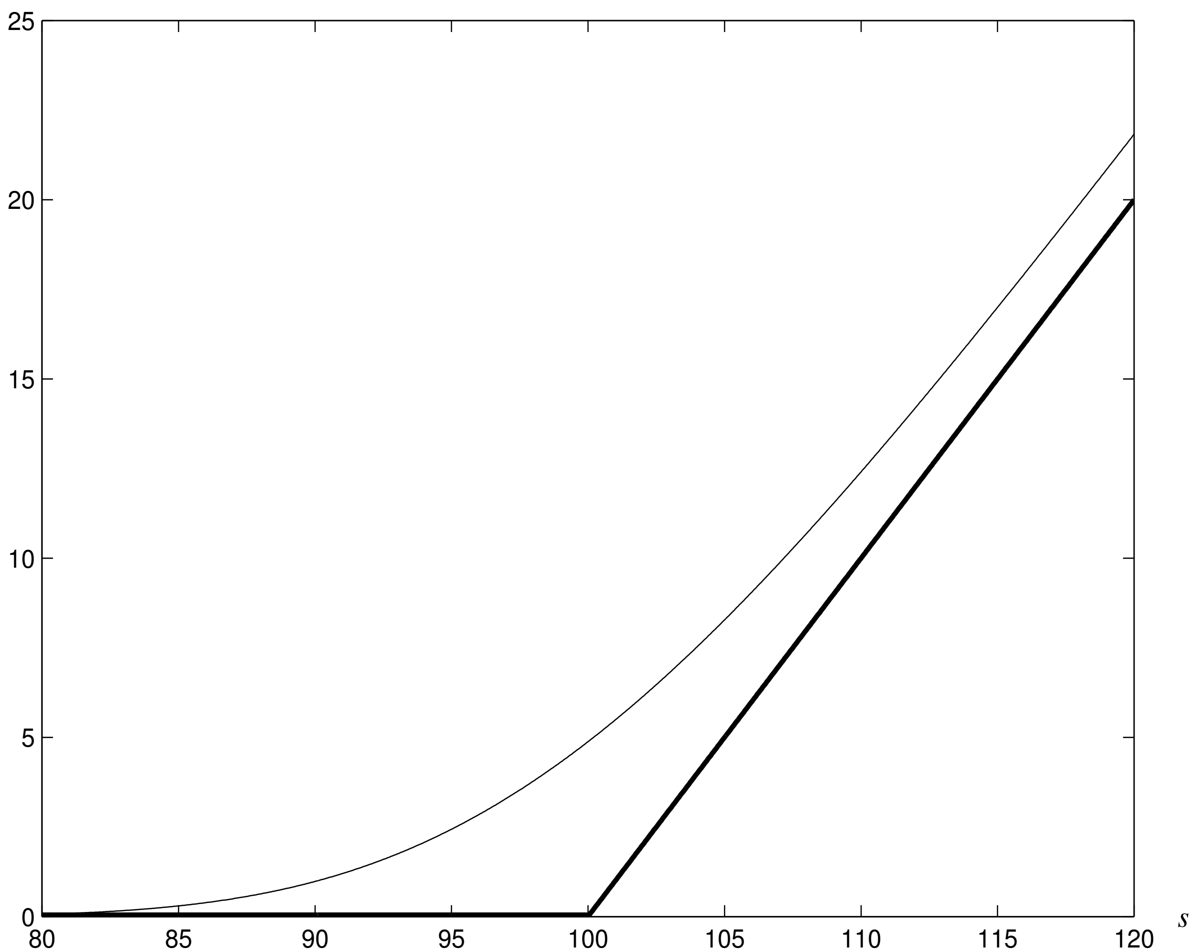
$N[\cdot]$=cdf for $N(0,1)$-distribution.

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T-t}}\left\{\ln\left(\frac{s}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)\right\}, \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

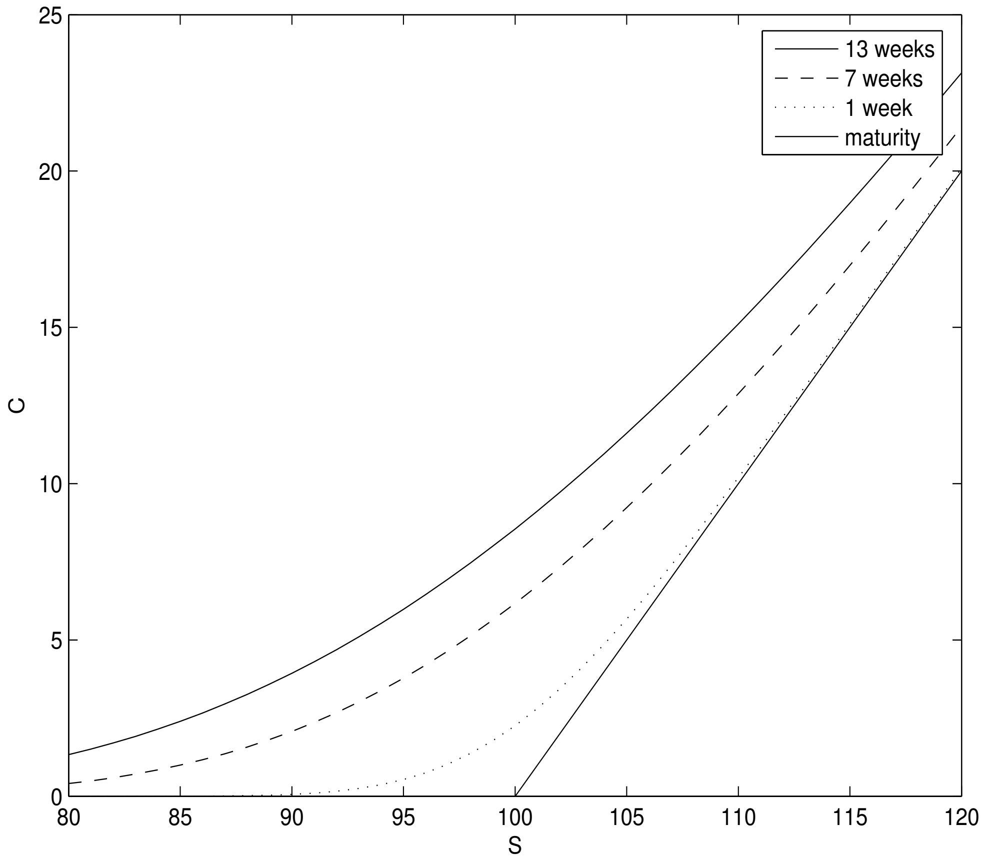
Black-Scholes

European Call,

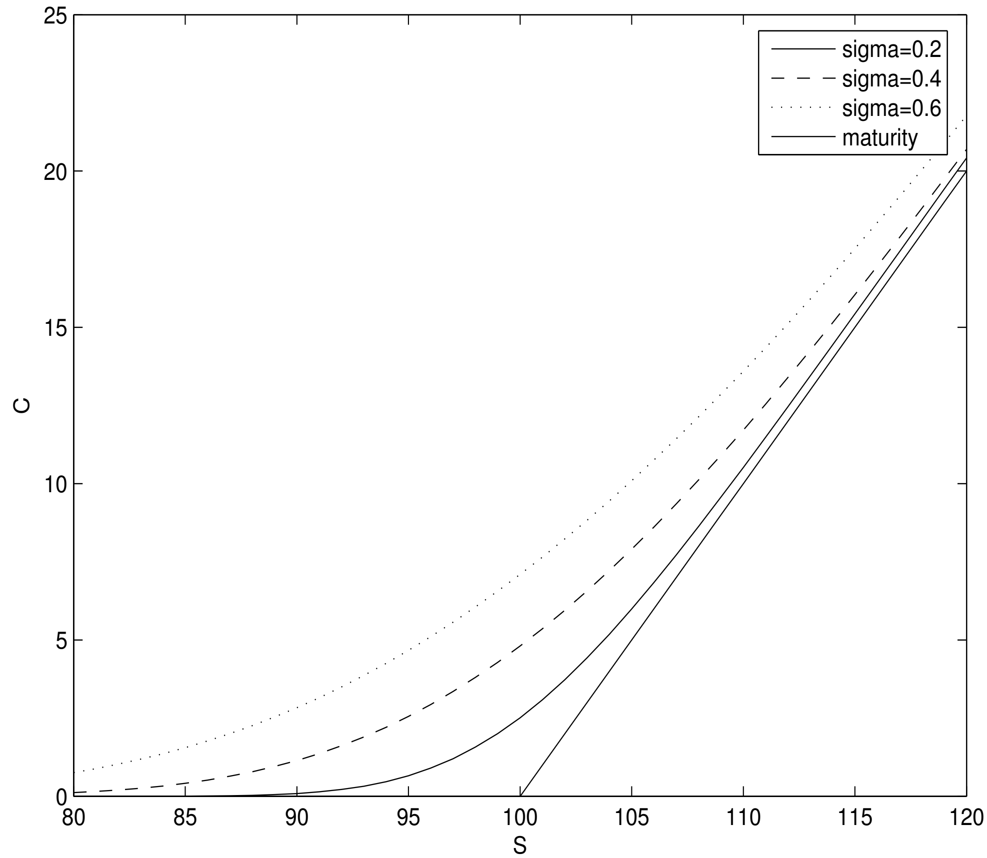
$$K = 100, \quad \sigma = 20\%, \quad r = 7\%, \quad T - t = 1/4$$



Dependence on Time to Maturity



Dependence on Volatility



4.

Risk Neutral Valuation

Risk neutral valuation

Appplying Feynman-Kac to the Black-Scholes PDE we obtain

$$\Pi_t[X] = e^{-r(T-t)} E^Q_{t,s}[X]$$

Q-**dynamics:**

$$\begin{cases} dS_t &= rS_t dt + \sigma S_t dW_t^Q, \\ dB_t &= rB_t dt. \end{cases}$$

- Price $=$ Expected discounted value of future payments.
- The expectation shall **not** be taken under the "objective" probability measure P, but under the "risk adjusted" measure ("martingale measure") Q.

Note: $P \sim Q$

Interpretation of the risk adjusted measure

- **Assume** a risk neutral world.
- Then the following must hold

$$s = S_0 = e^{-rt} E\left[S_t\right]$$

- In our model this means that

$$dS_t = rS_t dt + \sigma S_t dW_t^Q$$

- The risk adjusted probabilities can be intrepreted as probabilities in a fictuous risk neutral economy.

Moral

- When we compute prices, we can compute **as if** we live in a risk neutral world.
- This does **not** mean that we live (or think that we live) in a risk neutral world.
- The formulas above hold regardless of the investor's attitude to risk, as long as he/she prefers more to less.
- The valuation formulas are therefore called "preference free valuation formulas".

Properties of Q

The probability measure Q is fundamental for the theory. One can easily prove the following.

- The process

$$\frac{S_t}{B_t}$$

  is a Q-**martingale**.

- If $F(t, s)$ is the arbitrage free pricing function for a contract Φ then the process

$$\frac{f(t, S_t)}{B_t}$$

  is a Q-**martingale**.

- The volatility of the process $F(t, S_t)$ is the same under P and Q.

A small lemma

Consider an Ito process X. The following statements are then equivalent.

- $$\frac{X_t}{B_t}$$
  is a martingale under Q.

- X has Q dynamics of the form
  $$dX_t = rX_t dt + \sigma_t dW_t^Q.$$
  where W^Q is Q-Wiener. The point is that the **local rate of return of X under Q equals the risk free rate r.**

Properties of Q ct'd

- $P \sim Q$ (this will be explained later)

- For the price pricess π of any traded asset, derivative or underlying, the process

$$Z_t = \frac{\pi_t}{B_t}$$

  is a Q-martingale.

- Under Q, the price pricess π of any traded asset, derivative or underlying, has r as its local rate of return:

$$d\pi_t = r\pi_t dt + \sigma_\pi \pi_t dW_t^Q$$

- The volatility of π is the same under Q as under P.

A Preview of Martingale Measures

Consider a market, under an objective probability measure P, with underlying assets

$$B, S^1, \ldots, S^N$$

Definition: A probability measure Q is called a **martingale measure** if

- $P \sim Q$ (this will be explained later)
- For every i, the process

$$Z_t^i = \frac{S_t^i}{B_t}$$

  is a Q-martingale.

Theorem: The market is arbitrage free **iff** there exists a martingale measure.

5.

Appendices

Appendix A: Black-Scholes vs Binomial

Consider a binomial model for an option with a fixed time to maturity T and a fixed strike price X.

- Build a binomial model with n periods for each $n = 1, 2, ...$
- Use the standard formulas for scaling the jumps:

$$u = e^{\sigma\sqrt{\Delta t}} \quad d = e^{-\sigma\sqrt{\Delta t}} \quad \Delta t = T/n$$

- For a large n, the stock **price** at time T will then be a **product** of a large number of i.i.d. random variables.
- More precisely

$$S_T = S_0 Z_1 Z_2 \cdots Z_n,$$

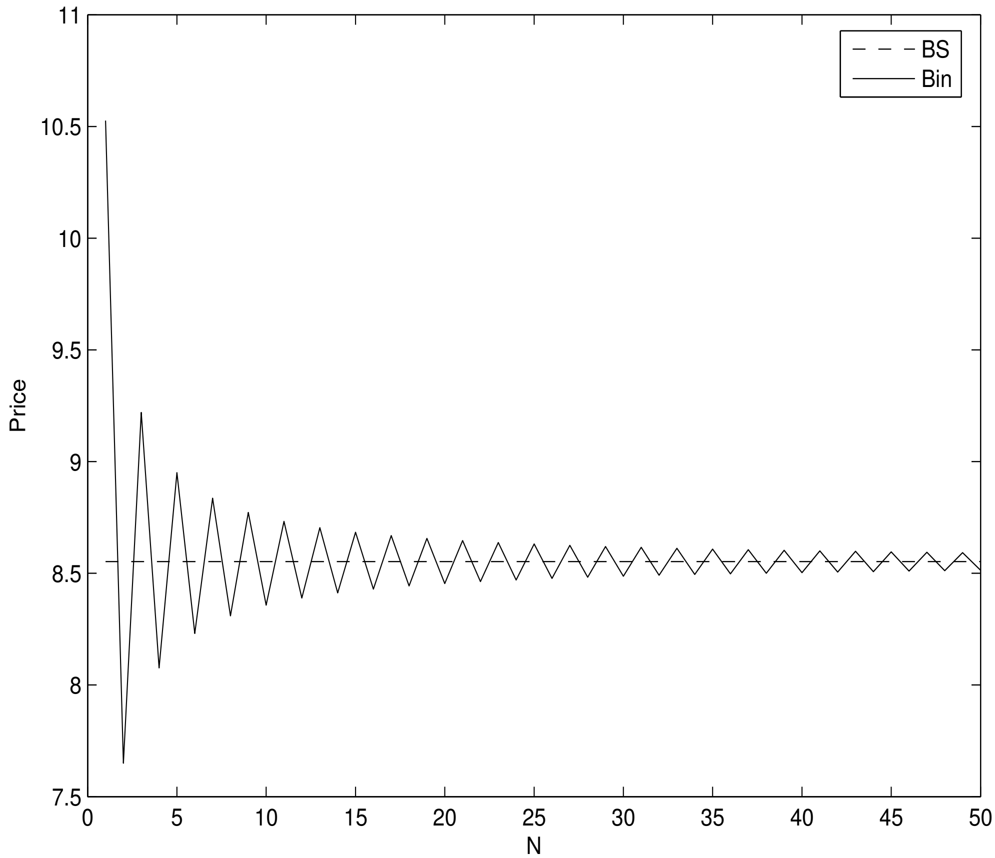
where n is the number of periods in the binomial model and $Z_i = u, d$

Recall

$$S_T = S_0 Z_1 Z_2 \cdots Z_n,$$

- The stock **price** at time T will be a **product** of a large number of i.i.d. random variables.
- The **return** will be a large **sum** of i.i.d. variables.
- The Central Limit Theorem will kick in.
- In the limit, **returns** will be **normally** distributed.
- Stock **prices** will be **lognormally** distributed.
- We are in the Black-Scholes model.
- The binomial price will converge to the Black-Scholes price.

Binomial convergence to Black-Scholes



Binomial $\sim$ Black-Scholes

The intuition from the Binomial model carries over to Black-Scholes.

- The B-S model is "just" a binomial model where we rebalance the portfolio infinitely often.
- The B-S model is thus complete.
- Completeness explains the unique prices for options in the B-S model.
- The B-S price for a derivative is the limit of the binomial price when the number of periods is very large.

Appendix B: Portfolio theory

We consider a market with N assets.

$$S_t^i = \text{price at } t\text{, of asset No } i.$$

A **portfolio** strategy is an adapted vector process

$$h_t = (h_t^1, \cdots, h_t^N)$$

where

$$\begin{aligned} h_t^i &= \text{number of units of asset } i, \\ V_t &= \text{market value of the portfolio} \end{aligned}$$

$$V_t = \sum_{i=1}^{N} h_t^i S_t^i$$

The portfolio is typically of the form

$$h_t = h(t, S_t)$$

i.e. today's portfolio is based on today's prices.

Self financing portfolios

We want to study **self financing** portfolio strategies, i.e. portfolios where

- There is now external infusion and/or withdrawal of money to/from the portfolio.
- Purchase of a “new” asset must be financed through sale of an “old” asset.

How is this formalized?

Problem: Derive an expression for dV_t for a self financing portfolio.

We analyze in discrete time, and then go to the continuous time limit.

Discrete time portfolios

We trade at discrete points in time $t = 0, 1, 2, \ldots$.

Price vector process:

$$S_n = (S_n^1, \cdots, S_n^N), \quad n = 0, 1, 2, \ldots$$

Portfolio process:

$$h_n = (h_n^1, \cdots, h_n^N), \quad n = 0, 1, 2, \ldots$$

Interpretation: At time n we buy the portfolio h_n at the price S_n, and keep it until time $n + 1$.

Value process:

$$V_n = \sum_{i=1}^{N} h_n^i S_n^i = h_n S_n$$

The self financing condition

- At time $n-1$ we buy the portfolio h_{n-1} at the price S_{n-1}.
- At time n this portfolio is worth $h_{n-1}S_n$.
- At time n we buy the new portfolio h_n at the price S_n.
- The cost of this new portfolio is h_nS_n.
- The self financing condition is the **budget constraint**
$$h_{n-1}S_n = h_nS_n$$

The self financing condition

Recall:

$$V_n = h_n S_n$$

Definition: For any sequence $x_1, x_2, \ldots$ we define the sequence Δx_n by

$$\Delta x_n = x_n - x_{n-1}$$

Problem: Derive an expression for ΔV_n for a self financing portfolio.

Lemma: For any pair of sequences $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ we have the relation

$$\Delta (xy)_n = x_{n-1} \Delta y_n + y_n \Delta x_n$$

Proof: Do it yourself.

Recall

$$V_n = h_n S_n$$

From the Lemma we have

$$\Delta V_n = \Delta(hS)_n = h_{n-1}\Delta S_n + S_n \Delta h_n$$

Recall the self financing condition

$$h_{n-1}S_n = h_n S_n$$

which we can write as

$$S_n \Delta h_n = 0$$

Inserting this into the expression for ΔV_n gives us.

Proposition: The dynamics of a self financing portfolio are given by

$$\Delta V_n = h_{n-1}\Delta S_n$$

Note the forward increments!

Portfolios in continuous time

Price process:

$$S_t^i = \text{price at } t \text{, of asset No } i.$$

Portfolio:

$$h_t = (h_t^1, \cdots, h_t^N)$$

Value process

$$V_t = \sum_{i=1}^{N} h_t^i S_t^i$$

From the self financing condition in discrete time

$$\Delta V_n = h_{n-1} \Delta S_n$$

we are led to the following definition.

Definition: The portfolio h is self financing if and only if

$$dV_t = \sum_{i=1}^{N} h_t^i dS_t^i$$

Relative weights

Definition:

$\omega_t^i =$ relative portfolio weight on asset No i.

We have

$$\omega_t^i = \frac{h_t^i S_t^i}{V_t}$$

Insert this into the self financing condition

$$dV_t = \sum_{i=1}^{N} h_t^i dS_t^i$$

We obtain

Portfolio dynamics:

$$dV_t = V_t \sum_{i=1}^{N} \omega_t^i \frac{dS_t^i}{S_t^i}$$

Interpret!

Appendix C: The original Black-Scholes PDE argument

Consider the following portfolio.

- Short one unit of the derivative, with pricing function $f(t,s)$.
- Hold x units of the underlying S.

The portfolio value is given by

$$V = -f(t, S_T) + xS_t$$

The object is to choose x such that the portfolio is risk free for an infinitesimal interval of length dt.

We have $dV = -df + xdS$ and from Itô we obtain

$$\begin{aligned} dV &= -\left\{ \frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial s} + \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial s^2} \right\} dt \\ &- \sigma S \frac{\partial f}{\partial s} dW + x\mu S dt + x\sigma S dW \end{aligned}$$

$$\begin{aligned} dV &= \left\{ x\mu S - \frac{\partial f}{\partial t} - \mu S \frac{\partial f}{\partial s} - \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial s^2} \right\} dt \\ &+ \sigma S \left\{ x - \frac{\partial f}{\partial s} \right\} dW \end{aligned}$$

We obtain a risk free portfolio if we choose x as

$$x = \frac{\partial f}{\partial s}$$

and then we have, after simplification,

$$dV = \left\{ -\frac{\partial f}{\partial t} - \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial s^2} \right\} dt$$

Using $V = -f + xS$ and x as above, the return dV/V is thus given by

$$\frac{dV}{V} = \frac{-\frac{\partial f}{\partial t} - \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial s^2}}{-f + S \frac{\partial f}{\partial s}} dt$$

We had

$$\frac{dV}{V} = \frac{-\frac{\partial f}{\partial t} - \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial s^2}}{-f + S\frac{\partial f}{\partial s}}dt$$

This portfolio is risk free, so absence of arbitrage implies that

$$\frac{-\frac{\partial f}{\partial t} - \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial s^2}}{-f + S\frac{\partial f}{\partial s}} = r$$

Simplifying this expression gives us the Black-Scholes PDE.

$$\begin{aligned} \frac{\partial f}{\partial t} + rs\frac{\partial f}{\partial s} + \frac{1}{2}\sigma^2 s^2\frac{\partial^2 f}{\partial s^2} - rf &= 0, \\ f(T, s) &= \Phi(s). \end{aligned}$$

Ch 8-9

Completeness and Hedging

Tomas Björk

Problems around Standard Black-Scholes

- We **assumed** that the derivative was traded. How do we price OTC products?

- Why is the option price independent of the expected rate of return α of the underlying stock?

- Suppose that we have sold a call option. Then we face financial risk, so how do we hedge against that risk?

All this has to do with **completeness**.

Definition:
We say that a T-claim X can be **replicated**, alternatively that it is **reachable** or **hedgeable**, if there exists a self financing portfolio h such that

$$V_T^h = X, \quad P - a.s.$$

In this case we say that h is a **hedge** against X. Alternatively, h is called a **replicating** or **hedging** portfolio. If every contingent claim is reachable we say that the market is **complete**

Basic Idea: If X can be replicated by a portfolio h then the arbitrage free price for X is given by

$$\Pi_t [X] = V_t^h.$$

Trading Strategy

Consider a replicable claim X which we want to sell at $t = 0$..

- Compute the price $\Pi_0[X]$ and sell X at a slightly (well) higher price.
- Buy the hedging portfolio and invest the surplus in the bank.
- Wait until expiration date T.
- The liabilities stemming from X is exactly matched by V_T^h, and we have our surplus in the bank.

Completeness of Black-Scholes

Theorem: The Black-Scholes model is complete.

Proof. Fix a claim $X = \Phi(S_T)$. We want to find processes V, ω^B and ω^S such that

$$dV = V\left\{\omega^B \frac{dB}{B} + \omega^S \frac{dS}{S}\right\}$$

$$V_T = \Phi(S_T).$$

i.e.

$$dV = V\left\{\omega^B r + \omega^S \alpha\right\} dt + V\omega^S \sigma dW,$$

$$V_T = \Phi(S_T).$$

Heuristics:
Let us **assume** that X is replicated by $\omega = (\omega^B, \omega^S)$ with value process V.

Ansatz:

$$V_t = F(t, S_t)$$

Ito gives us

$$dV = \left\{ F_t + \alpha S F_s + \frac{1}{2}\sigma^2 S^2 F_{ss} \right\} dt + \sigma S F_s dW,$$

Write this as

$$dV = V \left\{ \frac{F_t + \alpha S F_s + \frac{1}{2}\sigma^2 S^2 F_{ss}}{V} \right\} dt + V \frac{S F_s}{V} \sigma dW.$$

Compare with

$$dV = V \left\{ \omega^B r + \omega^S \alpha \right\} dt + V \omega^S \sigma dW$$

Define ω^S by

$$u_t^S = \frac{S_t F_s(t, S_t)}{F(t, S_t)},$$

This gives us the eqn

$$dV = V \left\{ \frac{F_t + \frac{1}{2}\sigma^2 S^2 F_{ss}}{rF} r + \omega^S \alpha \right\} dt + V \omega^S \sigma dW.$$

Compare with

$$dV = V \left\{ \omega^B r + \omega^S \alpha \right\} dt + V \omega^S \sigma dW$$

Natural choice for ω^B is given by

$$\omega^B = \frac{F_t + \frac{1}{2}\sigma^2 S^2 F_{ss}}{rF},$$

The relation $\omega^B + \omega^S = 1$ gives us the Black-Scholes PDE

$$F_t + rSF_s + \frac{1}{2}\sigma^2 S^2 F_{ss} - rF = 0.$$

The condition

$$V_T = \Phi\left(S_T\right)$$

gives us the boundary condition

$$F(T, s) = \Phi(s)$$

Moral: The model is complete and we have explicit formulas for the replicating portfolio.

Main Result

Theorem: Define F as the solution to the boundary value problem

$$\begin{cases} F_t + rsF_s + \frac{1}{2}\sigma^2 s^2 F_{ss} - rF = 0, \\ F(T,s) = \Phi(s). \end{cases}$$

Then X can be replicated by the relative portfolio

$$\begin{aligned} u_t^S &= \frac{F(t,S_t) - S_t F_s(t,S_t)}{F(t,S_t)}, \\ u_t^S &= \frac{S_t F_s(t,S_t)}{F(t,S_t)}. \end{aligned}$$

The corresponding absolute portfolio is given by

$$\begin{aligned} h_t^B &= \frac{F(t,S_t) - S_t F_s(t,S_t)}{B_t}, \\ h_t^S &= F_s(t,S_t), \end{aligned}$$

and the value process V^h is given by

$$V_t^h = F(t,S_t).$$

Notes

- Completeness explains unique price - the claim is superfluous!
- Replicating the claim $P - a.s. \Longleftrightarrow$ Replicating the claim $Q - a.s.$ for any $Q \sim P$. Thus the price only depends on the support of P.
- Thus (Girsanov) it will not depend on the drift α of the state equation.
- The completeness theorem is a nice theoretical result, but the replicating portfolio is **continuously rebalanced**. Thus we are facing very high transaction costs.

Completeness vs No Arbitrage

Question:
When is a model arbitrage free and/or complete?

Answer:
Count the number of risky assets, and the number of random sources.

$$R = \text{number of random sources}$$
$$N = \text{number of risky assets}$$

Intuition:
If N is large, compared to R, you have lots of possibilities of forming clever portfolios. Thus lots of chances of making arbitrage profits. Also many chances of replicating a given claim.

Meta-Theorem

Generically, the following hold.

- The market is arbitrage free if and only if

$$N \leq R$$

- The market is complete if and only if

$$N \geq R$$

Example:
The Black-Scholes model. R=N=1. Arbitrage free and complete.

Parity Relations

Let Φ and Ψ be contract functions for the T-claims $\mathcal{X} = \Phi(S(T))$ and $Y = \Psi(S(T))$. Then for any real numbers α and β we have the following price relation.

$$\Pi_t\left[\alpha\Phi + \beta\Psi\right] = \alpha\Pi_t\left[\Phi\right] + \beta\Pi_t\left[\Psi\right].$$

Proof. Linearity of mathematical expectation.

Consider the following "basic" contract functions.

$$\begin{aligned}
\Phi_S(x) &= x, \\
\Phi_B(x) &\equiv 1, \\
\Phi_{C,K}(x) &= \max\left[x - K, 0\right].
\end{aligned}$$

Prices:

$$\begin{aligned}
\Pi_t\left[\Phi_S\right] &= S_t, \\
\Pi_t\left[\Phi_B\right] &= e^{-r(T-t)}, \\
\Pi_t\left[\Phi_{C,K}\right] &= c(t, S_t; K, T).
\end{aligned}$$

If we have

$$\Phi = \alpha\Phi_S + \beta\Phi_B + \sum_{i=1}^{n} \gamma_i \Phi_{C,K_i},$$

then

$$\Pi_t\left[\Phi\right] = \alpha\Pi_t\left[\Phi_S\right] + \beta\Pi_t\left[\Phi_B\right] + \sum_{i=1}^{n} \gamma_i \Pi_t\left[\Phi_{C,K_i}\right]$$

We may replicate the claim Φ using a portfolio consisting of basic contracts that is **constant** over time, i.e. a **buy-and hold** portfolio:

- α shares of the underlying stock,
- β zero coupon T-bonds with face value $\$1$,
- γ_i European call options with strike price K_i, all maturing at T.

Put-Call Parity

Consider a European put contract

$$\Phi_{P,K}(s) = \max\left[K - s, 0\right]$$

It is easy to see (draw a figure) that

$$\begin{aligned}
\Phi_{P,K}(x) &= \Phi_{C,K}(x) - s + K \\
&= \Phi_{P,K}(x) - \Phi_S(x) + \Phi_B(x)
\end{aligned}$$

We immediately get

Put-call parity:

$$p(t, s; K) = c(t, s; K) - s + Ke^{r(T-t)}$$

Thus you can construct a synthetic put option, using a buy-and-hold portfolio.

Delta Hedging

Consider a fixed claim

$$X = \Phi(S_T)$$

with pricing function

$$F(t, s).$$

Setup:
We are at time t, and have a short (interpret!) position in the contract.

Goal:
Offset the risk in the derivative by buying (or selling) the (highly correlated) underlying.

Definition:
A position in the underlying is a **delta hedge** against the derivative if the portfolio (underlying + derivative) is immune against small changes in the underlying price.

Formal Analysis

$$\begin{aligned} -1 &= \text{number of units of the derivative product} \\ x &= \text{number of units of the underlying} \\ s &= \text{today's stock price} \\ t &= \text{today's date} \end{aligned}$$

Value of the portfolio:

$$V = -1 \cdot F(t, s) + x \cdot s$$

A delta hedge is characterized by the property that

$$\frac{\partial V}{\partial s} = 0.$$

We obtain

$$-\frac{\partial F}{\partial s} + x = 0$$

Solve for x!

Result:
We should have

$$\hat{x} = \frac{\partial F}{\partial s}$$

shares of the underlying in the delta hedged portfolio.

Definition:
For any contract, its "delta" is defined by

$$\Delta = \frac{\partial F}{\partial s}.$$

Result:
We should have

$$\hat{x} = \Delta$$

shares of the underlying in the delta hedged portfolio.

Warning:
The delta hedge must be rebalanced over time. (why?)

Black Scholes

For a European Call in the Black-Scholes model we have

$$\Delta = N[d_1]$$

NB This is **not** a trivial result!

From put call parity it follows (how?) that Δ for a European Put is given by

$$\Delta = N[d_1] - 1$$

Check signs and interpret!

Rebalanced Delta Hedge

- Sell one call option a time $t = 0$ at the B-S price F.

- Compute Δ and by Δ shares. (Use the income from the sale of the option, and borrow money if necessary.)

- Wait one day (week, minute, second..). The stock price has now changed.

- Compute the new value of Δ, and borrow money in order to adjust your stock holdings.

- Repeat this procedure until $t = T$. Then the value of your portfolio (B+S) will match the value of the option almost exactly.

- Lack of perfection comes from discrete, instead of continuous, trading.

- You have created a “synthetic” option. (Replicating portfolio).

Formal result:
The relative weights in the replicating portfolio are

$$
\begin{aligned}
u_S &= \frac{S \cdot \Delta}{F}, \\
u_B &= \frac{F - S \cdot \Delta}{F}
\end{aligned}
$$

Portfolio Delta

Assume that you have a portfolio consisting of derivatives

$$\Phi_i(S_{T_i}), \quad i = 1, \cdots, n$$

all **written on the same underlying** stock S.

$$\begin{aligned} F_i(t,s) &= \text{pricing function for i:th derivative} \\ \Delta_i &= \frac{\partial F_i}{\partial s} \\ h_i &= \text{units of i:th derivative} \end{aligned}$$

Portfolio value:

$$\Pi = \sum_{i=1}^{n} h_i F_i$$

Portfolio delta:

$$\Delta_\Pi = \sum_{i=1}^{n} h_i \Delta_i$$

Gamma

A problem with discrete delta-hedging is.

- As time goes by S will change.
- This will cause $\Delta = \frac{\partial F}{\partial s}$ to change.
- Thus you are sitting with the wrong value of delta.

Moral:

- If delta is sensitive to changes in S, then you have to rebalance often.
- If delta is insensitive to changes in S you do not need to rebalance so often.

Definition:
Let Π be the value of a derivative (or portfolio). **Gamma** (Γ) is defined as

$$\Gamma = \frac{\partial \Delta}{\partial s}$$

i.e.

$$\Gamma = \frac{\partial^2 \Pi}{\partial s^2}$$

Gamma is a measure of the sensitivity of Δ to changes in S.

Result: For a European Call in a Black-Scholes model, Γ can be calculated as

$$\Gamma = \frac{N'[d_1]}{S\sigma\sqrt{T-t}}$$

Important fact:
For a position in the underlying stock itself we have

$$\Gamma = 0$$

Gamma Neutrality

A portfolio Π is said to be **gamma neutral** if its gamma equals zero, i.e.

$$\Gamma_{\Pi} = 0$$

- Since $\Gamma = 0$ for a stock you can not gamma-hedge using only stocks. item Typically you use some derivative to obtain gamma neutrality.

General procedure

Given a portfolio Π with underlying S. Consider two derivatives with pricing functions F and G.

$$\begin{aligned} x_F &= \text{number of units of } F \\ x_G &= \text{number of units of } G \end{aligned}$$

Problem:
Choose x_F and x_G such that the entire portfolio is delta- and gamma-neutral.

Value of hedged portfolio:

$$V = \Pi + x_F \cdot F + x_G \cdot G$$

Value of hedged portfolio:

$$V = \Pi + x_F \cdot F + x_G \cdot G$$

We get the equations

$$\begin{aligned} \frac{\partial V}{\partial s} &= 0, \\ \frac{\partial^2 V}{\partial s^2} &= 0. \end{aligned}$$

i.e.

$$\begin{aligned} \Delta_\Pi + x_F \Delta_F + x_G \Delta_G &= 0, \\ \Gamma_\Pi + x_F \Gamma_F + x_G \Gamma_G &= 0 \end{aligned}$$

Solve for x_F and x_G!

Further Greeks

$$\Theta = \frac{\partial \Pi}{\partial t},$$

$$V = \frac{\partial \Pi}{\partial \sigma},$$

$$\rho = \frac{\partial \Pi}{\partial r}$$

V is pronounced "Vega".

NB!

- A delta hedge is a hedge against the movements in the underlying stock, given a **fixed model**.
- A Vega-hedge is not a hedge against movements of the underlying asset. It is a hedge against a **change of the model itself**.

Chapter 11

The Martingale Approach

I: Mathematics

Tomas Björk

Introduction

In order to understand and to apply the martingale approach to derivative pricing and hedging we will need to some basic concepts and results from measure theory. These will be introduced below in an informal manner - for full details see the textbook.

Many propositions below will be proved but we will also present a couple of central results without proofs, and these must then be considered as dogmatic truths. You are of course not expected to know the proofs of such results (this is outside the scope of this course) but you are supposed to be able to **use** the results in an operational manner.

Contents

1.

Events and sigma-algebras

Events and sigma-algebras

Consider a probability measure P on a sample space Ω. An **event** is simply a subset $A \subseteq \Omega$ and $P(A)$ is the probability that the event A occurs.

For technical reasons, a probability measure can only be defined for a certain "nice" class $\mathcal{F}$ of events, so for $A \in \mathcal{F}$ we are allowed to write $P(A)$ as the probability for the event A.

For technical reasons the class $\mathcal{F}$ must be a **sigma-algebra**, which means that $\mathcal{F}$ is closed under the usual set theoretic operations like complements, countable intersections and countable unions.

Interpretation: We can view a σ-algebra $\mathcal{F}$ as formalizing the idea of information. More precisely: A σ-algebra $\mathcal{F}$ is a collection of events, and if we assume that we have access to the information contained in $\mathcal{F}$, this means that for every $A \in \mathcal{F}$ we know exactly if A has occured or not.

Borel sets

Definition: The **Borel algebra** $\mathcal{B}$ is the smallest sigma-algebra on R which contains all intervals. A set B in $\mathcal{B}$ is called a **Borel set**.

Remark: There is no constructive definition of $\mathcal{B}$, but almost all subsets of R that you will ever see will in fact be Borel sets, so the reader can without danger think about a Borel set as "an arbitrary subset of R".

Random variables

An $\mathcal{F}$-**measurable random variable** X is a a mapping

$$X : \Omega \to R$$

such that $\{X \in B\} = \{\omega \in \Omega : \ X(\omega) \in B\}$ belongs to $\mathcal{F}$ for all Borel sets B. This guarantees that we are allowed to write $P(X \in B)$. Instad of writing "X is $\mathcal{F}$-measurable" we will often write $X \in \mathcal{F}$.

This means that if $X \in \mathcal{F}$ then the value of X is completely determined by the information contained in $\mathcal{F}$.

If we have another σ-algebra $\mathcal{G}$ with $\mathcal{G} \subseteq \mathcal{F}$ then we interpret this as "$\mathcal{G}$ contains less information than $\mathcal{F}$".

2.

Conditional Expectation

Conditional Expectation

If $X \in \mathcal{F}$ and if $\mathcal{G} \subseteq \mathcal{F}$ then we write $E[X|\mathcal{G}]$ for the conditional expectation of X given the information contained in $\mathcal{G}$. Sometimes we use the notation $E_{\mathcal{G}}[X]$.

The following proposition contains everything that we will need to know about conditional expectations within this course.

Main Results

Proposition 1: Assume that $X \in \mathcal{F}$, and that $\mathcal{G} \subseteq \mathcal{F}$. Then the following hold.

- The random variable $E[X|\mathcal{G}]$ is completely determined by the information in $\mathcal{G}$ so we have

$$E[X|\mathcal{G}] \in \mathcal{G}$$

- If we have $Y \in \mathcal{G}$ then Y is completely determined by $\mathcal{G}$ so we have

$$E[XY|\mathcal{G}] = YE[X|\mathcal{G}]$$

  In particular we have

$$E[Y|\mathcal{G}] = Y$$

- If $\mathcal{H} \subseteq \mathcal{G}$ then we have the "law of iterated expectations"

$$E[E[X|\mathcal{G}]|\mathcal{H}] = E[X|\mathcal{H}]$$

- In particular we have

$$E[X] = E[E[X|\mathcal{G}]]$$

3.

Changing Measures

Changing Measures

Consider a probability measure P on $(\Omega, \mathcal{F})$, and assume that $L \in \mathcal{F}$ is a random variable with the properties that

$$L \geq 0$$

and

$$E^P[L] = 1.$$

For every event $A \in \mathcal{F}$ we now define the real number $Q(A)$ by the prescription

$$Q(A) = E^P[L \cdot I_A]$$

where the random variable I_A is the indicator for A, i.e.

$$I_A = \begin{cases} 1 & \text{if} \quad A \text{ occurs} \\ 0 & \text{if} \quad A^c \text{ occurs} \end{cases}$$

Recall that

$$Q(A) = E^P [L \cdot I_A]$$

We now see that $Q(A) \geq 0$ for all A, and that

$$Q(\Omega) = E^P [L \cdot I_\Omega] = E^P [L \cdot 1] = 1$$

We also see that if $A \cap B = \emptyset$ then

$$\begin{aligned} Q(A \cup B) &= E^P [L \cdot I_{A \cup B}] = E^P [L \cdot (I_A + I_B)] \\ &= E^P [L \cdot I_A] + E^P [L \cdot I_B] \\ &= Q(A) + Q(B) \end{aligned}$$

Furthermore we see that

$$P(A) = 0 \quad \Rightarrow \quad Q(A) = 0$$

We have thus more or less proved the following

Proposition 2: If $L \in \mathcal{F}$ is a nonnegative random variable with $E^P[L] = 1$ and Q is defined by

$$Q(A) = E^P[L \cdot I_A]$$

then Q will be a probability measure on $\mathcal{F}$ with the property that

$$P(A) = 0 \quad \Rightarrow \quad Q(A) = 0.$$

I turns out that the property above is a very important one, so we give it a name.

Absolute Continuity

Definition: Given two probability measures P and Q on $\mathcal{F}$ we say that Q is **absolutely continuous** w.r.t. P on $\mathcal{F}$ if, for all $A \in \mathcal{F}$, we have

$$P(A) = 0 \quad \Rightarrow \quad Q(A) = 0$$

We write this as

$$Q << P.$$

If $Q << P$ and $P << Q$ then we say that P and Q are **equivalent** and write

$$Q \sim P$$

Equivalent measures

It is easy to see that P and Q are equivalent if and only if

$$P(A) = 0 \quad \Leftrightarrow \quad Q(A) = 0$$

or, equivalently,

$$P(A) = 1 \quad \Leftrightarrow \quad Q(A) = 1$$

Two equivalent measures thus agree on all certain events and on all impossible events, but can disagree on all other events.

Simple examples:

- All non degenerate Gaussian distributions on R are equivalent.

- If P is Gaussian on R and Q is exponential then $Q << P$ but not the other way around.

Absolute Continuity ct'd

We have seen that if we are given P and **define** Q by

$$Q(A) = E^P [L \cdot I_A]$$

for $L \geq 0$ with $E^P [L] = 1$, then Q is a probability measure and $Q << P$. .

A natural question is now if **all** measures $Q << P$ are obtained in this way. The answer is yes, and the precise (quite deep) result is as follows. The proof is difficult and therefore omitted.

The Radon Nikodym Theorem

Consider two probability measures P and Q on $(\Omega, \mathcal{F})$, and assume that $Q << P$ on $\mathcal{F}$. Then there exists a unique random variable L with the following properties

1. $Q(A) = E^P[L \cdot I_A], \quad \forall A \in \mathcal{F}$
2. $L \geq 0, \quad P - a.s.$
3. $E^P[L] = 1,$
4. $L \in \mathcal{F}$

The random variable L is denoted as

$$L = \frac{dQ}{dP}, \quad \text{on } \mathcal{F}$$

and it is called the **Radon-Nikodym derivative** of Q w.r.t. P on $\mathcal{F}$, or the **likelihood ratio** between Q and P on $\mathcal{F}$.

A simple example

The Radon-Nikodym derivative L is intuitively the local scale factor between P and Q. If the sample space Ω is finite so $\Omega = \{\omega_1, \dots, \omega_n\}$ then P is determined by the probabilities $p_1, \dots, p_n$ where

$$p_i = P(\omega_i) \quad i = 1, \dots, n$$

Now consider a measure Q with probabilities

$$q_i = Q(\omega_i) \quad i = 1, \dots, n$$

If $Q << P$ this simply says that

$$p_i = 0 \quad \Rightarrow \quad q_i = 0$$

and it is easy to see that the Radon-Nikodym derivative $L = dQ/dP$ is given by

$$L(\omega_i) = \frac{q_i}{p_i} \quad i = 1, \dots, n$$

If $p_i = 0$ then we also have $q_i = 0$ and we can define the ratio q_i/p_i arbitrarily.

If $p_1, \ldots, p_n$ as well as $q_1, \ldots, q_n$ are all positive, then we see that $Q \sim P$ and in fact

$$\frac{dP}{dQ} = \frac{1}{L} = \left(\frac{dQ}{dP} \right)^{-1}$$

as could be expected.

Computing expected values

A main use of Radon-Nikodym derivatives is for the computation of expected values.

Suppose therefore that $Q << P$ on $\mathcal{F}$ and that X is a random variable with $X \in \mathcal{F}$. With $L = dQ/dP$ on $\mathcal{F}$ then have the following result.

Proposition 3: With notation as above we have

$$E^Q[X] = E^P[L \cdot X]$$

Proof: We only give a proof for the simple example above where $\Omega = \{\omega_1, \ldots, \omega_n\}$. We then have

$$\begin{aligned} E^Q[X] &= \sum_{i=1}^{n} X(\omega_i) q_i = \sum_{i=1}^{n} X(\omega_i) \frac{q_i}{p_i} p_i \\ &= \sum_{i=1}^{n} X(\omega_i) L(\omega_i) p_i = E^P[X \cdot L] \end{aligned}$$

The Abstract Bayes' Formula

We can also use Radon-Nikodym derivatives in order to compute conditional expectations. The result, known as the abstract **Bayes' Formula**, is as follows.

Theorem 4: Consider two measures P and Q with $Q << P$ on $\mathcal{F}$ and with

$$L^{\mathcal{F}} = \frac{dQ}{dP} \quad \text{on } \mathcal{F}$$

Assume that $\mathcal{G} \subseteq \mathcal{F}$ and let X be a random variable with $X \in \mathcal{F}$. Then the following holds

$$E^{Q}\left[X \middle| \mathcal{G}\right] = \frac{E^{P}\left[L^{\mathcal{F}} X \middle| \mathcal{G}\right]}{E^{P}\left[L^{\mathcal{F}} \middle| \mathcal{G}\right]}$$

Dependence of the σ-algebra

Suppose that we have $Q << P$ on $\mathcal{F}$ with

$$L^{\mathcal{F}} = \frac{dQ}{dP} \quad \text{on } \mathcal{F}$$

Now consider smaller σ-algebra $\mathcal{G} \subseteq \mathcal{F}$. Our problem is to find the R-N derivative

$$L^{\mathcal{G}} = \frac{dQ}{dP} \quad \text{on } \mathcal{G}$$

We recall that $L^{\mathcal{G}}$ is characterized by the following properties

1. $Q(A) = E^P\left[L^{\mathcal{G}} \cdot I_A\right] \quad \forall A \in \mathcal{G}$

2. $L^{\mathcal{G}} \geq 0$

3. $E^P\left[L^{\mathcal{G}}\right] = 1$

4. $L^{\mathcal{G}} \in \mathcal{G}$

A natural guess would perhaps be that $L^{\mathcal{G}} = L^{\mathcal{F}}$, so let us check if $L^{\mathcal{F}}$ satisfies points 1-4 above.

By assumption we have

$$Q(A) = E^{P}\left[L^{\mathcal{F}} \cdot I_A\right] \quad \forall A \in \mathcal{F}$$

Since $\mathcal{G} \subseteq \mathcal{F}$ we then have

$$Q(A) = E^{P}\left[L^{\mathcal{F}} \cdot I_A\right] \quad \forall A \in \mathcal{G}$$

so point 1 above is certainly satisfied by $L^{\mathcal{F}}$. It is also clear that $L^{\mathcal{F}}$ satisfies points 2 and 3. It thus seems that $L^{\mathcal{F}}$ is also a natural candidate for the R-N derivative $L^{\mathcal{G}}$, but the problem is that we do not in general have $L^{\mathcal{F}} \in \mathcal{G}$.

This problem can, however, be fixed. By iterated expectations we have, for all $A \in \mathcal{G}$,

$$E^{P}\left[L^{\mathcal{F}} \cdot I_A\right] = E^{P}\left[E^{P}\left[\left. L^{\mathcal{F}} \cdot I_A \right| \mathcal{G}\right]\right]$$

Since $A \in \mathcal{G}$ we have

$$E^P\left[L^{\mathcal{F}} \cdot I_A \middle| \mathcal{G}\right] = E^P\left[L^{\mathcal{F}} \middle| \mathcal{G}\right] I_A$$

Let us now define $L^{\mathcal{G}}$ by

$$L^{\mathcal{G}} = E^P\left[L^{\mathcal{F}} \middle| \mathcal{G}\right]$$

We then obviously have $L^{\mathcal{G}} \in \mathcal{G}$ and

$$Q(A) = E^P\left[L^{\mathcal{G}} \cdot I_A\right] \quad \forall A \in \mathcal{G}$$

It is easy to see that also points 2-3 are satisfied so we have proved the following result.

A formula for $L^{\mathcal{G}}$

Proposition 5: If $Q << P$ on $\mathcal{F}$ and $\mathcal{G} \subseteq \mathcal{F}$ then, with notation as above, we have

$$L^{\mathcal{G}} = E^{P}\left[L^{\mathcal{F}} \middle| \mathcal{G} \right]$$

The likelihood process on a filtered space

We now consider the case when we have a probability measure P on some space Ω and that instead of just one σ-algebra $\mathcal{F}$ we have a **filtration**, i.e. an increasing family of σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$.

The interpretation is as usual that $\mathcal{F}_t$ is the information available to us at time t, and that we have $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$.

Now assume that we also have another measure Q, and that for some fixed T, we have $Q << P$ on $\mathcal{F}_T$. We define the random variable L_T by

$$L_T = \frac{dQ}{dP} \quad \text{on } \mathcal{F}_T$$

Since $Q << P$ on $\mathcal{F}_T$ we also have $Q << P$ on $\mathcal{F}_t$ for all $t \leq T$ and we define

$$L_t = \frac{dQ}{dP} \quad \text{on } \mathcal{F}_t \quad 0 \leq t \leq T$$

For every t we have $L_t \in \mathcal{F}_t$, so L is an adapted process, known as the **likelihood process**.

The L process is a P martingale

We recall that

$$L_t = \frac{dQ}{dP} \quad \text{on } \mathcal{F}_t \quad 0 \le t \le T$$

Since $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \le t$ we can use Proposition 5 and deduce that

$$L_s = E^P \left[L_t | \mathcal{F}_s \right] \quad s \le t \le T$$

and we have thus proved the following result.

Proposition: Given the assumptions above, the likelihood process L is a P-martingale.

Where are we heading?

We are now going to perform measure transformations on Wiener spaces, where P will correspond to the objective measure and Q will be the risk neutral measure.

For this we need define the proper likelihood process L and, since L is a P-martingale, we have the following natural questions.

- What does a martingale look like in a Wiener driven framework?
- Suppose that we have a P-Wiener process W and then change measure from P to Q. What are the properties of W under the new measure Q?

These questions are handled by the Martingale Representation Theorem, and the Girsanov Theorem respectively.

4.

The Martingale Representation Theorem

Intuition

Suppose that we have a Wiener process W under the measure P. We recall that if h is adapted (and integrable enough) and if the process X is defined by

$$X_t = x_0 + \int_0^t h_s dW_s$$

then X is a a martingale. We now have the following natural question:

Question: Assume that X is an arbitrary martingale. Does it then follow that X has the form

$$X_t = x_0 + \int_0^t h_s dW_s$$

for some adapted process h?

In other words: Are **all** martingales stochastic integrals w.r.t. W?

Answer

It is immediately clear that all martingales can **not** be written as stochastic integrals w.r.t. W. Consider for example the process X defined by

$$X_t = \begin{cases} 0 & \text{for} \quad 0 \le t < 1 \\ Z & \text{for} \quad t \ge 1 \end{cases}$$

where Z is an random variable, independent of W, with $E[Z] = 0$.

X is then a martingale (why?) but it is clear (how?) that it cannot be written as

$$X_t = x_0 + \int_0^t h_s dW_s$$

for any process h.

Intuition

The intuitive reason why we cannot write

$$X_t = x_0 + \int_0^t h_s dW_s$$

in the example above is of course that the random variable Z "has nothing to do with" the Wiener process W. In order to exclude examples like this, we thus need an assumption which guarantees that our probability space only contains the Wiener process W and nothing else.

This idea is formalized by assuming that the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ **is the one generated by the Wiener process** W.

The Martingale Representation Theorem

Theorem. Let W be a P-Wiener process and assume that the filtation is the **internal** one i.e.

$$\mathcal{F}_t = \mathcal{F}_t^W = \sigma\left\{W_s;\ 0 \leq s \leq t\right\}$$

Then, for every $(P, \mathcal{F}_t)$-martingale X, there exists a real number x and an adapted process h such that

$$X_t = x + \int_0^t h_s dW_s,$$

i.e.

$$dX_t = h_t dW_t.$$

Proof: Hard. This is very deep result.

Note

For a given martingale X, the Representation Theorem above guarantees the existence of a process h such that

$$X_t = x + \int_0^t h_s dW_s,$$

The Theorem does **not**, however, tell us how to find or construct the process h.

5.

The Girsanov Theorem

Setup

Let W be a P-Wiener process and fix a time horizon T. Suppose that we want to change measure from P to Q on $\mathcal{F}_T$. For this we need a P-martingale L with $L_0 = 1$ to use as a likelihood process, and a natural way of constructing this is to choose a process g and then define L by

$$\begin{cases} dL_t &= g_t dW_t \\ L_0 &= 1 \end{cases}$$

This definition does not guarantee that $L \geq 0$, so we make a small adjustment. We choose a process φ and define L by

$$\begin{cases} dL_t &= L_t \varphi_t dW_t \\ L_0 &= 1 \end{cases}$$

The process L will again be a martingale and we easily obtain

$$L_t = e^{\int_0^t \varphi_s dW_s - \frac{1}{2}\int_0^t \varphi_s^2 ds}$$

Thus we are guaranteed that $L \geq 0$. We now change measure form P to Q by setting

$$dQ = L_t dP, \quad \text{on } \mathcal{F}_t,\ 0 \leq t \leq T$$

The main problem is to find out what the properties of W are, under the new measure Q. This problem is resolved by the **Girsanov Theorem**.

The Girsanov Theorem

Let W be a P-Wiener process. Fix a time horizon T.

Theorem: Choose an adapted process φ, and define the process L by

$$\begin{cases} dL_t = L_t \varphi_t dW_t \\ L_0 = 1 \end{cases}$$

Assume that $E^P[L_T] = 1$, and define a new mesure Q on $\mathcal{F}_T$ by

$$dQ = L_t dP, \quad \text{on } \mathcal{F}_t,\ 0 \le t \le T$$

Then $Q << P$ and the process W^Q, defined by

$$W_t^Q = W_t - \int_0^t \varphi_s ds$$

is Q-Wiener. We can also write this as

$$dW_t = \varphi_t dt + dW_t^Q$$

Changing the drift in an SDE

The single most common use of the Girsanov Theorem is as follows.

Suppose that we have a process X with P dynamics

$$dX_t = \mu_t dt + \sigma_t dW_t$$

where μ and σ are adapted and W is P-Wiener.

We now do a Girsanov Transformation as above, and the question is what the Q-dynamics look like.

From the Girsanov Theorem we have

$$dW_t = \varphi_t dt + dW_t^Q$$

and substituting this into the P-dynamics we obtain the Q dynamics as

$$dX_t = \{\mu_t + \sigma_t \varphi_t\} dt + \sigma_t dW_t^Q$$

Moral: The drift changes but the diffusion is unaffected.

The Converse Girsanov Theorem

Let W be a P-Wiener process. Fix a time horizon T.

Theorem. Assume that:

- $Q << P$ on $\mathcal{F}_T$, with likelihood process

$$L_t = \frac{dQ}{dP}, \quad \text{on } \mathcal{F}_t \; 0, \leq t \leq T$$

- The filtation is the **internal** one .i.e.

$$\mathcal{F}_t = \sigma \left\{ W_s; \; 0 \leq s \leq t \right\}$$

Then there exists a process φ such that

$$\begin{cases} dL_t &= L_t \varphi_t dW_t \\ L_0 &= 1 \end{cases}$$

Chapter 10

The Martingale Approach

II. Pricing and Hedging

Tomas Björk

Financial Markets

Price Processes:

$$S_t = \left[S_t^0, ..., S_t^N \right]$$

Example: (Black-Scholes, $S^0 := B,\ S^1 := S$)

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= r B_t dt. \end{aligned}$$

Portfolio:

$$h_t = \left[h_t^0, ..., h_t^N \right]$$

$h_t^i =$ number of units of asset i at time t.

Value Process:

$$V_t^h = \sum_{i=0}^{N} h_t^i S_t^i = h_t S_t$$

Self Financing Portfolios

Definition: (intuitive)
A portfolio is **self-financing** if there is no exogenous infusion or withdrawal of money. “The purchase of a new asset must be financed by the sale of an old one.”

Definition: (mathematical)
A portfolio is **self-financing** if the value process satisfies

$$dV_t = \sum_{i=0}^{N} h_t^i dS_t^i$$

Major insight:
If the price process S is a **martingale**, and if h is **self-financing**, then V is a **martingale**.

NB! This simple observation is in fact the basis of the following theory.

Arbitrage

We now give the full technical definition of arbitrage.

Definition: The portfolio u is an **arbitrage** if

- The portfolio strategy is self financing.
- $V_0 = 0$.
- $V_T \geq 0, \ P - a.s.$
- $P\left(V_T > 0\right) > 0$

Main Question: When is the market free of arbitrage?

First Attempt

Proposition: If $S_t^0, \cdots, S_t^N$ are P-martingales, then the market is free of arbitrage.

Proof:
Assume that V is an arbitrage strategy. Since

$$dV_t = \sum_{i=0}^{N} h_t^i dS_t^i,$$

V is a P-martingale, so

$$V_0 = E^P \left[V_T \right] > 0.$$

This contradicts $V_0 = 0$.

True, but useless.

Example: (Black-Scholes)

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= rB_t dt. \end{aligned}$$

(We would have to assume that $\alpha = r = 0$)

We now try to improve on this result.

Choose S_0 as numeraire

Definition:
The **normalized price vector** Z is given by

$$Z_t = \frac{S_t}{S_t^0} = \left[1, Z_t^1, ..., Z_t^N\right]$$

The **normalized value process** V^Z is given by

$$V_t^Z = \sum_0^N h_t^i Z_t^i.$$

Idea:
The arbitrage and self financing concepts should be independent of the accounting unit.

Invariance of numeraire

Proposition: One can show (see the book) that

- S-arbitrage $\iff$ Z-arbitrage.

- S-self-financing $\iff$ Z-self-financing.

Insight:

- If h self-financing then

$$dV_t^Z = \sum_1^N h_t^i dZ_t^i$$

- Thus, if the **normalized** price process Z is a P-martingale, then V^Z is a martingale.

Second Attempt

Proposition: If $Z_t^0, \cdots, Z_t^N$ are P-martingales, then the market is free of arbitrage.

True, but still fairly useless.

Example: (Black-Scholes)

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= rB_t dt. \end{aligned}$$

$$\begin{aligned} dZ_t^1 &= (\alpha - r) Z_t^1 dt + \sigma Z_t^1 dW_t, \\ dZ_t^0 &= 0dt. \end{aligned}$$

We would have to assume "risk-neutrality", i.e. that $\alpha = r$.

Arbitrage

Recall that h is an arbitrage if

- h is self financing
- $V_0 = 0$.
- $V_T \geq 0, \; P - a.s.$
- $P\left(V_T > 0\right) > 0$

Major insight

This concept is invariant under an **equivalent change of measure!**

Martingale Measures

Definition: A probability measure Q is called an **equivalent martingale measure** (EMM) if and only if it has the following properties.

- Q and P are equivalent, i.e.

$$Q \sim P$$

- The normalized price processes

$$Z_t^i = \frac{S_t^i}{S_t^0}, \quad i = 0, \ldots, N$$

  are **Q-martingales**.

Wan now state the main result of arbitrage theory.

First Fundamental Theorem

Theorem: The market is arbitrage free

iff

there exists an equivalent martingale measure.

Comments

- It is very easy to prove that existence of EMM imples no arbitrage (see below).
- The other implication is technically very hard.
- For discrete time and finite sample space Ω the hard part follows easily from the separation theorem for convex sets.
- For discrete time and more general sample space we need the Hahn-Banach Theorem.
- For continuous time the proof becomes technically very hard, mainly due to topological problems. See the textbook.

Proof that EMM implies no arbitrage

Assume that there exists an EMM denoted by Q. Assume that $P(V_T \geq 0) = 1$ and $P(V_T > 0) > 0$. Then, since $P \sim Q$ we also have $Q(V_T \geq 0) = 1$ and $Q(V_T > 0) > 0$.

Recall:

$$dV_t^Z = \sum_1^N h_t^i dZ_t^i$$

Q is a martingale measure

$$\Downarrow$$

V^Z is a Q-martingale

$$\Downarrow$$

$$V_0 = V_0^Z = E^Q\left[V_T^Z\right] > 0$$

$$\Downarrow$$

No arbitrage

Choice of Numeraire

The **numeraire** price S_t^0 can be chosen arbitrarily. The most common choice is however that we choose S^0 as the **bank account**, i.e.

$$S_t^0 = B_t$$

where

$$dB_t = r_t B_t dt$$

Here r is the (possibly stochastic) short rate and we have

$$B_t = e^{\int_0^t r_s ds}$$

Example: The Black-Scholes Model

$$
\begin{aligned}
dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\
dB_t &= r B_t dt.
\end{aligned}
$$

Look for martingale measure. We set $Z = S/B$.

$$dZ_t = Z_t(\alpha - r)dt + Z_t \sigma dW_t,$$

Girsanov transformation on $[0, T]$:

$$
\begin{cases}
dL_t &= L_t \varphi_t dW_t, \\
L_0 &= 1.
\end{cases}
$$

$$dQ = L_T dP, \quad \text{on } \mathcal{F}_T$$

Girsanov:

$$dW_t = \varphi_t dt + dW_t^Q,$$

where W^Q is a Q-Wiener process.

The Q-dynamics for Z are given by

$$dZ_t = Z_t \left[\alpha - r + \sigma\varphi_t\right] dt + Z_t \sigma dW_t^Q.$$

Unique martingale measure Q, with Girsanov kernel given by

$$\varphi_t = \frac{r - \alpha}{\sigma}.$$

Q-dynamics of S:

$$dS_t = rS_t dt + \sigma S_t dW_t^Q.$$

Conclusion: The Black-Scholes model is free of arbitrage.

Pricing

We consider a market $B_t, S_t^1, \ldots, S_t^N$.

Definition:
A **contingent claim** with **delivery time** T, is a random variable

$$X \in \mathcal{F}_T.$$

"At $t = T$ the amount X is paid to the holder of the claim".

Example: (European Call Option)

$$X = \max\left[S_T - K, 0\right]$$

Let X be a contingent T-claim.

Problem: How do we find an arbitrage free price process $\Pi_t\left[X\right]$ for X?

Solution

The extended market

$$B_t, S_t^1, \ldots, S_t^N, \Pi_t[X]$$

must be arbitrage free, so there must exist a martingale measure Q for $(S_t, \Pi_t[X])$. In particular

$$\frac{\Pi_t[X]}{B_t}$$

must be a Q-martingale, i.e.

$$\frac{\Pi_t[X]}{B_t} = E^Q\left[\left.\frac{\Pi_T[X]}{B_T}\right| \mathcal{F}_t\right]$$

Since we obviously (why?) have

$$\Pi_T[X] = X$$

we have proved the main pricing formula.

Risk Neutral Valuation

Theorem: For a T-claim X, the arbitrage free price is given by the formula

$$\Pi_t[X] = E^Q\left[e^{-\int_t^T r_s ds} \times X \middle| \mathcal{F}_t \right]$$

Example: The Black-Scholes Model

Q-dynamics:

$$dS_t = rS_t dt + \sigma S_t dW_t^Q.$$

Simple claim:

$$X = \Phi(S_T),$$

$$\Pi_t[X] = e^{-r(T-t)} E^Q[\Phi(S_T) | \mathcal{F}_t]$$

Kolmogorov $\Rightarrow$

$$\Pi_t[X] = F(t, S_t)$$

where $F(t, s)$ solves the Black-Scholes equation:

$$\begin{cases} \frac{\partial F}{\partial t} + rs\frac{\partial F}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 F}{\partial s^2} - rF = 0, \\ F(T, s) = \Phi(s). \end{cases}$$

Problem

Recall the valuation formula

$$\Pi_t[X] = E^Q\left[e^{-\int_t^T r_s ds} \times X \middle| \mathcal{F}_t\right]$$

What if there are several different martingale measures Q?

This is connected with the **completeness** of the market.

Hedging

Def: A portfolio is a **hedge** against X ("replicates X") if

- h is self financing
- $V_T = X, \quad P - a.s.$

Def: The market is **complete** if every X can be hedged.

Pricing Formula:
If h replicates X, then a natural way of pricing X is

$$\Pi_t[X] = V_t^h$$

When can we hedge?

Second Fundamental Theorem

The second most important result in arbitrage theory is the following.

Theorem:

The market is complete

iff

the martingale measure Q is unique.

Proof: It is obvious (why?) that if the market is complete, then Q must be unique. The other implication is very hard to prove. It basically relies on duality arguments from functional analysis.

Main Results

- The market is arbitrage free $\Leftrightarrow$ There exists a martingale measure Q
- The market is complete $\Leftrightarrow$ Q is unique.
- Every X must be priced by the formula

$$\Pi_t[X] = E^Q\left[e^{-\int_t^T r_s ds} \times X \middle| \mathcal{F}_t \right]$$

  for some choice of Q.

- In a non-complete market, different choices of Q will produce different prices for X.
- For a hedgeable claim X, all choices of Q will produce the same price for X:

$$\Pi_t[X] = V_t = E^Q\left[e^{-\int_t^T r_s ds} \times X \middle| \mathcal{F}_t \right]$$

Stochastic Discount Factors

Given a model under P. For every EMM Q we define the corresponding **Stochastic Discount Factor**, or **SDF**, by

$$D_t = e^{-\int_0^t r_s ds} L_t,$$

where

$$L_t = \frac{dQ}{dP}, \quad \text{on } \mathcal{F}_t$$

There is thus a one-to-one correspondence between EMMs and SDFs.

The risk neutral valuation formula for a T-claim X can now be expressed under P instead of under Q.

Proposition: With notation as above we have

$$\Pi_t[X] = \frac{1}{D_t} E^P\left[D_T X \middle| \mathcal{F}_t \right]$$

Proof: Bayes' formula.

Martingale Property of $S \cdot D$

Proposition: If S is an arbitrary price process, then the process

$$S_t D_t$$

is a P-martingale.

Proof: Bayes' formula.

Technical appendix on completeness

The main tool for completeness is the following fact.

Existence of hedge

$$\Updownarrow$$

Existence of stochastic integral representation

A small but usefull observation

Fix T-claim X.

If h is a hedge for X then

- $V_T^Z = \frac{X}{B_T}$
- h is self financing, i.e.

$$dV_t^Z = \sum_1^K h_t^i dZ_t^i$$

Thus V^Z is a Q-**martingale**.

$$V_t^Z = E^Q \left[\left. \frac{X}{B_T} \right| \mathcal{F}_t \right]$$

Main technical result

Proposition:
Fix T-claim X. Define martingale M by

$$M_t = E^Q \left[\frac{X}{B_T} \middle| \mathcal{F}_t \right]$$

Suppose that there exist predictable processes $h^1, \cdots, h^N$ such that

$$M_t = x + \sum_{i=1}^{N} \int_0^t h_s^i dZ_s^i,$$

Then X can be replicated by the portfolio $h^B, h^1, \ldots, h^N$, where $h^1, \ldots, h^N$ are as above and h^B is given by

$$h_t^B = M_t - \sum_{i=1}^{N} h_t^i Z_t^i.$$

Proof

We guess that

$$M_t = V_t^Z = h_t^B \cdot 1 + \sum_{i=1}^{N} h_t^i Z_t^i$$

Define: h^B by

$$h_t^B = M_t - \sum_{i=1}^{N} h_t^i Z_t^i.$$

We have $M_t = V_t^Z$, and we get

$$dV_t^Z = dM_t = \sum_{i=1}^{N} h_t^i dZt^i,$$

so the portfolio is self financing. Furthermore:

$$V_T^Z = M_T = E^Q \left[\left. \frac{X}{B_T} \right| \mathcal{F}_T \right] = \frac{X}{B_T}.$$

Black-Scholes Model

Q-dynamics

$$\begin{aligned} dS_t &= rS_t dt + \sigma S_t dW_t^Q, \\ dZ_t &= Z_t \sigma dW_t^Q \end{aligned}$$

$$M_t = E^Q\left[e^{-rT} X \middle| \mathcal{F}_t \right],$$

Representation theorem for Wiener processes

$\Downarrow$

there exists g such that

$$M_t = M(0) + \int_0^t g_s dW_s^Q.$$

Thus

$$M_t = M_0 + \int_0^t h_s^1 dZ_s,$$

with $h_t^1 = \frac{g_t}{\sigma Z_t}$.

Result:
X can be replicated using the portfolio defined by

$$\begin{aligned} h_t^1 &= g_t/\sigma Z_t, \\ h_t^B &= M_t - h_t^1 Z_t. \end{aligned}$$

Moral: The Black Scholes model is complete.

Special Case: Simple Claims

Assume X is of the form $X = \Phi(S_T)$

$$M_t = E^Q\left[e^{-rT}\Phi(S_T) \middle| \mathcal{F}_t \right],$$

Kolmogorov backward equation $\Rightarrow M_t = f(t, S_t)$

$$\begin{cases} \frac{\partial f}{\partial t} + rs\frac{\partial f}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 f}{\partial s^2} & = & 0, \\ f(T, s) & = & e^{-rT}\Phi(s). \end{cases}$$

Itô $\Rightarrow$

$$dM_t = \sigma S_t \frac{\partial f}{\partial s} dW_t^Q,$$

so

$$g_t = \sigma S_t \cdot \frac{\partial f}{\partial s},$$

Replicating portfolio h:

$$\begin{aligned} h_t^B &= f - S_t \frac{\partial f}{\partial s}, \\ h_t^1 &= B_t \frac{\partial f}{\partial s}. \end{aligned}$$

Interpretation: $f(t, S_t) = V_t^Z$.

Define $F(t,s)$ by

$$F(t,s) = e^{rt} f(t,s)$$

so $F(t, S_t) = V_t$. Then

$$\begin{cases} h_t^B = \frac{F(t,S_t) - S_t \frac{\partial F}{\partial s}(t,S_t)}{B_t}, \\ h_t^1 = \frac{\partial F}{\partial s}(t, S_t) \end{cases}$$

where F solves the **Black-Scholes equation**

$$\begin{cases} \frac{\partial F}{\partial t} + rs\frac{\partial F}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 F}{\partial s^2} - rF = 0, \\ F(T,s) = \Phi(s). \end{cases}$$

Chapter 15

Incomplete Markets

Tomas Björk

Derivatives on Non Financial Underlying

Recall: The Black-Scholes theory assumes that the market for the underlying asset has (among other things) the following properties.

- The underlying is a liquidly traded asset.
- Shortselling allowed.
- Portfolios can be carried forward in time.

There exists a large market for derivatives, where the underlying does not satisfy these assumptions.

Examples:

- Weather derivatives.
- Derivatives on electric energy.
- CAT-bonds.

Typical Contracts

Weather derivatives:

"Heating degree days". Payoff at maturity T is given by

$$\mathcal{Z} = \max\{X_T - 30, 0\}$$

where X_T is the (mean) temperature at some place.

Electricity option:

The right (but not the obligation) to buy, at time T, at a predetermined price K, a constant flow of energy over a predetermined time interval.

CAT bond:

A bond for which the payment of coupons and nominal value is contingent on some (well specified) natural disaster to take place.

Problems

Weather derivatives:
The temperature is not the price of a traded asset.

Electricity derivatives:
Electric energy cannot easily be stored.

CAT-bonds:
Natural disasters are not traded assets.

We will treat all these problems within a **factor model**.

Typical Factor Model Setup

Given:

- An underlying factor process X, which is **not** the price process of a traded asset, with dynamics under the objective probability measure P as

$$dX_t = \mu\left(t, X_t\right)dt + \sigma\left(t, X_t\right)dW_t.$$

- A risk free asset with dynamics

$$dB_t = rB_t dt,$$

Problem:

Find arbitrage free price $\Pi_t\left[\mathcal{Z}\right]$ of a derivative of the form

$$\mathcal{Z} = \Phi(X_T)$$

Concrete Examples

Assume that X_t is the temperature at time t at the village of Peniche (Portugal).

Heating degree days:

$$\Phi(X_T) = 100 \cdot \max\{X_T - 30, 0\}$$

Holiday Insurance:

$$\Phi(X_T) = \begin{cases} 1000, & \text{if } X_T < 20 \\ 0, & \text{if } X_T \geq 20 \end{cases}$$

Question

Is the price $\Pi_t[\Phi]$ uniquely determined by the P-dynamics of X, and the requirement of an arbitrage free derivatives market?

NO!!

WHY?

Stock Price Model $\sim$ Factor Model

Black-Scholes:

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

Factor Model:

$$\begin{aligned} dX_t &= \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \\ dB_t &= r B_t dt. \end{aligned}$$

What is the difference?

Answer

- X is **not the price of a traded asset!**
- We can not form a portfolio based on X.

1. **Rule of thumb:**

$$\begin{aligned} N &= 0, \quad \text{(no risky asset)} \\ R &= 1, \quad \text{(one source of randomness, } W) \end{aligned}$$

   We have $N < R$. The exogenously given market, consisting only of B, is incomplete.

2. **Replicating portfolios:**
   We can only invest money in the bank, and then sit down passively and wait.

We do **not** have **enough underlying assets** in order to price X-derivatives.

- There is **not** a unique price for a **particular** derivative.

- In order to avoid arbitrage, **different** derivatives have to satisfy **internal consistency** relations.

- If we take **one** "benchmark" derivative as given, then all other derivatives can be priced **in terms of** the market price of the benchmark.

We consider two given claims $\Phi(X_T)$ and $\Gamma(X_T)$. We assume they are traded with prices

$$\begin{aligned}\Pi_t[\Phi] &= f(t, X_t) \\ \Pi_t[\Gamma] &= g(t, X_t)\end{aligned}$$

Program:

- Form portfolio based on Φ and Γ. Use Itô on f and g to get portfolio dynamics.

$$dV = V\left\{\omega^f \frac{df}{f} + \omega^g \frac{dg}{g}\right\}$$

- Choose portfolio weights such that the $dW-$ term vanishes. Then we have

$$dV = V \cdot kdt,$$

("synthetic bank" with k as the short rate)

- Absence of arbitrage implies

$$k = r$$

- Read off the relation $k = r$!

From Itô:

$$df = f\mu_f dt + f\sigma_f dW,$$

where

$$\begin{cases} \mu_f & = & \frac{f_t + \mu f_x + \frac{1}{2}\sigma^2 f_{xx}}{f}, \\ \sigma_f & = & \frac{\sigma f_x}{f}. \end{cases}$$

Portfolio dynamics

$$dV = V\left\{\omega^f \frac{df}{f} + \omega^g \frac{dg}{g}\right\}.$$

Reshuffling terms gives us

$$dV = V \cdot \left\{\omega^f \mu_f + \omega^g \mu_g\right\} dt + V \cdot \left\{\omega^f \sigma_f + \omega^g \sigma_g\right\} dW.$$

Let the portfolio weights solve the system

$$\begin{cases} \omega^f + \omega^g & = & 1, \\ \omega^f \sigma_f + \omega^g \sigma_g & = & 0. \end{cases}$$

$$
\begin{aligned}
\omega^f &= -\frac{\sigma_g}{\sigma_f - \sigma_g}, \\
\omega^g &= \frac{\sigma_f}{\sigma_f - \sigma_g},
\end{aligned}
$$

Portfolio dynamics

$$dV = V \cdot \left\{ \omega^f \mu_f + \omega^g \mu_g \right\} dt.$$

i.e.

$$dV = V \cdot \left\{ \frac{\mu_g \sigma_f - \mu_f \sigma_g}{\sigma_f - \sigma_g} \right\} dt.$$

Absence of arbitrage requires

$$\frac{\mu_g \sigma_f - \mu_f \sigma_g}{\sigma_f - \sigma_g} = r$$

which can be written as

$$\frac{\mu_g - r}{\sigma_g} = \frac{\mu_f - r}{\sigma_f}.$$

$$\frac{\mu_g - r}{\sigma_g} = \frac{\mu_f - r}{\sigma_f}.$$

Note!
The quotient does **not** depend upon the particular choice of contract.

Result

Assume that the market for X-derivatives is free of arbitrage. Then there exists a universal process λ, such that

$$\frac{\mu_f(t) - r}{\sigma_f(t)} = \lambda(t, X_t),$$

holds for all t and for every choice of contract f.

NB: The same λ for all choices of f.

$$\begin{aligned} \lambda &= \text{Risk premium per unit of volatility} \\ &= \text{“Market Price of Risk” (cf. CAPM).} \\ &= \text{Sharpe Ratio} \end{aligned}$$

Slogan:
"On an arbitrage free market all X-derivatives have the same market price of risk."

The relation

$$\frac{\mu_f - r}{\sigma_f} = \lambda$$

is actually a PDE!

Pricing Equation

$$\begin{cases} f_t + \{\mu - \lambda\sigma\} f_x + \frac{1}{2}\sigma^2 f_{xx} - rf &= 0 \\ f(T,x) &= \Phi(x), \end{cases}$$

P-**dynamics:**

$$dX = \mu(t,X)dt + \sigma(t,X)dW.$$

Can we solve the PDE?

No!!

Why??

Answer

Recall the PDE

$$
\begin{cases}
f_t + \{\mu - \lambda\sigma\} f_x + \frac{1}{2}\sigma^2 f_{xx} - rf &= 0 \\
f(T,x) &= \Phi(x),
\end{cases}
$$

- In order to solve the PDE **we need to know λ.**
- λ is not given exogenously.
- λ is not determined endogenously.

Question:

Who determines λ?

Answer:

THE MARKET!

Interpreting λ

Recall that the f dynamics are

$$df = f\mu_f dt + f\sigma_f dW_t$$

and λ is defined as

$$\frac{\mu_f(t) - r}{\sigma_f(t)} = \lambda(t, X_t),$$

- λ **measures the aggregate risk aversion** in the market.
- If λ is big then the market is highly risk averse.
- If λ is zero then the market is **risk netural**.
- If you make an assumption about λ, then you implicitly make an assumption about the aggregate risk aversion of the market.

Moral

- Since the market is incomplete the requirement of an arbitrage free market will **not** lead to unique prices for X-derivatives.

- Prices on derivatives are determined by two main factors.

  1. **Partly** by the requirement of an arbitrage free derivative market. **All** pricing functions satisfies the **same** PDE.
  2. **Partly** by supply and demand on the market. These are in turn determined by attitude towards risk, liquidity consideration and other factors. All these are aggregated into the particular λ used (implicitly) by the market.

Risk Neutral Valuation

We recall the PDE

$$\begin{cases} f_t + \{\mu - \lambda\sigma\} f_x + \frac{1}{2}\sigma^2 f_{xx} - rf & = & 0 \\ f(T,x) & = & \Phi(x), \end{cases}$$

Using Feynman-Kac we obtain a risk neutral valuation formula.

Risk Neutral Valuation

$$f(t,x) = e^{-r(T-t)} E^Q_{t,x}\left[\Phi(X_T)\right]$$

Q-**dynamics:**

$$dX_t = \{\mu - \lambda\sigma\}\, dt + \sigma dW^Q_t$$

- Price = expected value of future payments
- The expectation should **not** be taken under the "objective" probabilities P, but under the "risk adjusted" probabilities Q.

Interpretation of the risk adjusted probabilities

- The risk adjusted probabilities can be interpreted as probabilities in a (fictuous) risk neutral world.
- When we **compute prices**, we can calculate **as if** we live in a risk neutral world.
- This does **not** mean that we live in, or think that we live in, a risk neutral world.
- The formulas above hold regardless of the attitude towards risk of the investor, as long as he/she prefers more to less.

Diversification argument about λ

- If the risk factor is **idiosyncratic** and **diversifiable**, then one can argue that the factor should not be priced by the market. Compare with APT.

- Mathematically this means that $\lambda = 0$, i.e. $P = Q$, i.e. **the risk neutral distribution coincides with the objective distribution**.

- We thus have the **"actuarial pricing formula"**

$$f(t,x) = e^{-r(T-t)} E^P_{t,x}\left[\Phi(X_T)\right]$$

where we use the objective probabiliy measure P.

Modeling Issues

Temperature:

A standard model is given by

$$dX_t = \{m(t) - bX_t\} \, dt + \sigma dW_t,$$

where m is the mean temperature capturing seasonal variations. This often works reasonably well.

Electricity:

A (naive) model for the spot electricity price is

$$dS_t = S_t \{m(t) - a \ln S_t\} \, dt + \sigma S_t dW_t$$

This implies lognormal prices (why?). Electricty prices are however very far from lognormal, because of "spikes" in the prices. Complicated.

CAT bonds:

Here we have to use the theory of point processes and the theory of extremal statistics to model natural disasters. Complicated.

Martingale Analysis

Model: Under P we have

$$\begin{aligned} dX_t &= \mu(t, X_t)\, dt + \sigma(t, X_t)\, dW_t, \\ dB_t &= rB_t dt, \end{aligned}$$

We look for martingale measures. Since B is the only traded asset we need to find $Q \sim P$ such that

$$\frac{B_t}{B_t} = 1$$

is a Q martingale.

Result: In this model, **every** $Q \sim P$ is a martingale measure.

Girsanov

$$dL_t = L_t \varphi_t dW_t$$

P-dynamics

$$dX_t = \mu\left(t, X_t\right) dt + \sigma\left(t, X_t\right) dW_t,$$

$$dL_t = L_t \varphi_t dW_t$$

$dQ = L_t dP$ on $\mathcal{F}_t$

Girsanov:

$$dW_t = \varphi_t dt + dW_t^Q$$

Martingale pricing:

$$F(t, x) = e^{-r(T-t)} E^Q \left[Z \middle| \mathcal{F}_t \right]$$

Q-dynamics of X:

$$dX_t = \left\{ \mu\left(t, X_t\right) + \sigma\left(t, X_t\right) \varphi_t \right\} dt + \sigma\left(t, X_t\right) dW_t^Q,$$

Result: We have $\lambda_t = -\varphi_t$, i.e,. the Girsanov kernel φ equals minus the market price of risk.

Several Risk Factors

We recall the dynamics of the f-derivative

$$df = f\mu_f dt + f\sigma_f dW_t$$

and the Market Price of Risk

$$\frac{\mu_f - r}{\sigma_f} = \lambda, \quad \text{i.e.} \quad \mu_f - r = \lambda\sigma_f.$$

In a multifactor model of the type

$$dX_t = \mu\left(t, X_t\right) dt + \sum_{i=1}^{n} \sigma_i\left(t, X_t\right) dW_t^i,$$

it follows from Girsanov that for every risk factor W^i there will exist a market price of risk $\lambda_i = -\varphi_i$ such that

$$\mu_f - r = \sum_{i=1}^{n} \lambda_i \sigma_i$$

Compare with CAPM.

Chapters 16 & 26

Forwards, Futures, and Futures Options

Tomas Björk

Contents

1. Dividends

Dividends

Black-Scholes model:

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= r B_t dt. \end{aligned}$$

New feature:
The underlying stock pays **dividends**.

$$D_t = \text{The cumulative dividends over the interval } [0, t]$$

Interpretation:
Over the interval $[t, t+dt]$ you obtain the amount dD_t

Two cases

- Discrete dividends (realistic but messy).
- Continuous dividends (unrealistic but easy to handle).

Portfolios and Dividends

Consider a market with N assets.

$$
\begin{array}{rcl}
S_t^i & = & \text{price at } t\text{, of asset No } i \\
D_t^i & = & \text{cumulative dividends for } S^i \text{ over} \\
 & & \text{the interval } [0,t] \\
h_t^i & = & \text{number of units of asset } i \\
V_t & = & \text{market value of the portfolio } h \text{ at } t
\end{array}
$$

Assumption: We assume that D has continuous trajctories.

Definition: The **value process** V is defined by

$$V_t = \sum_{i=1}^{N} h_t^i S_t^i$$

Interpretation of D

Consider a price dividend pair (S, D). recall that

$$D_t = \text{cumulative dividends for } S \text{ over the interval } [0, t]$$

Thus $D_t =$ the sum of all dividends during $[0, t]$.

The intuitive interpretation of the cumulative dividend process D is:

$$dD_t = D_{t+dt} - D_t = \text{dividends obtained during } (t, t + dt]$$

Self financing portfolios

Recall:

$$V_t = \sum_{i=1}^{N} h_t^i S_t^i$$

Definition: The strategy h is **self financing** if

$$dV_t = \sum_{i=1}^{N} h_t^i dG_t^i$$

where the **gain** process G^i is defined by

$$dG_t^i = dS_t^i + dD_t^i$$

Interpret!

Note: The definitions above rely on the assumption that D is continuous. In the case of a discontinuous D, the definitions are more complicated.

Relative weights

$\omega_t^i =$ the relative share of the portfolio value, which is invested in asset No i.

$$\omega_t^i = \frac{h_t^i S_t^i}{V_t}$$

$$dV_t = \sum_{i=1}^{N} h_t^i dG_t^i$$

Substitute!

$$dV_t = V_t \sum_{i=1}^{N} \omega_t^i \frac{dG_t^i}{S_t^i}$$

Quiz

Problem 2:
Suppose the stock pays 5 dollars at time t. What happens to the stock price S at time t

Problem 2:
How does a dividend affect the price of a European Call? (compared to a non-dividend paying stock).

Continuous Dividend Yield

Definition: The stock S pays a **continuous dividend yield** of q, if D has the form

$$dD_t = qS_t dt$$

Black-Scholes with Cont. Dividend Yield

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dD_t &= q S_t dt \end{aligned}$$

Gain process:

$$dG_t = (\alpha + q) S_t dt + \sigma S_t dW_t$$

Consider a fixed claim

$$X = \Phi(S_T)$$

and assume that

$$\Pi_t [X] = F(t, S_t)$$

Standard Procedure

- Assume that the derivative price is of the form

$$\Pi_t[X] = F(t, S_t).$$

- Form a portfolio based on underlying S and derivative F, with portfolio dynamics

$$dV_t = V_t \left\{ \omega_t^S \cdot \frac{dG_t}{S_t} + \omega_t^F \cdot \frac{dF}{F} \right\}$$

- Choose ω^S and ω^F such that the dW-term is wiped out. This gives us

$$dV_t = V_t \cdot k_t dt$$

- Absence of arbitrage implies

$$k_t = r$$

- This relation will say something about F.

Value dynamics:

$$dV = V \cdot \left\{ \omega^S \frac{dG}{S} + \omega^F \frac{dF}{F} \right\},$$

$$dG = S(\alpha + q)dt + \sigma S dW.$$

From Itô we obtain

$$dF = \alpha_F F dt + \sigma_F F dW,$$

where

$$\begin{aligned} \alpha_F &= \frac{1}{F} \left\{ \frac{\partial F}{\partial t} + \alpha S \frac{\partial F}{\partial s} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial s^2} \right\}, \\ \sigma_F &= \frac{1}{F} \cdot \sigma S \frac{\partial F}{\partial s}. \end{aligned}$$

Collecting terms gives us

$$\begin{aligned} dV &= V \cdot \left\{ \omega^S (\alpha + q) + \omega^F \alpha_F \right\} dt \\ &+ V \cdot \left\{ \omega^S \sigma + \omega^F \sigma_F \right\} dW, \end{aligned}$$

Define ω^S and ω^F by the system

$$\begin{aligned} \omega^S \sigma + \omega^F \sigma_F &= 0, \\ \omega^S + \omega^F &= 1. \end{aligned}$$

Solution

$$\begin{aligned} \omega^S &= \frac{\sigma_F}{\sigma_F - \sigma}, \\ \omega^F &= \frac{-\sigma}{\sigma_F - \sigma}, \end{aligned}$$

Value dynamics

$$dV = V \cdot \left\{ \omega^S (\alpha + q) + \omega^F \alpha_F \right\} dt.$$

Absence of arbitrage implies

$$\omega^S (\alpha + q) + \omega^F \alpha_F = r,$$

We get

$$\frac{\partial F}{\partial t} + (r - q) S \frac{\partial F}{\partial s} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 F}{\partial s^2} - rF = 0.$$

Pricing PDE

Proposition: The pricing function F is given as the solution to the PDE

$$\begin{cases} \dfrac{\partial F}{\partial t} + (r-q)s\dfrac{\partial F}{\partial s} + \dfrac{1}{2}\sigma^2 s^2 \dfrac{\partial^2 F}{\partial s^2} - rF & = & 0, \\ F(T,s) & = & \Phi(s). \end{cases}$$

We can now apply Feynman-Kac to the PDE in order to obtain a risk neutral valuation formula.

Risk Neutral Valuation

The pricing function has the representation

$$F(t,s) = e^{-r(T-t)} E^{Q}_{t,s}\left[\Phi(S_T)\right],$$

where the Q-dynamics of S are given by

$$dS_t = (r-q)S_t dt + \sigma S_t dW^Q_t.$$

Question: Which object is a martingale under the meausre Q?

Martingale Property

Proposition: Under the martingale measure Q the **normalized gain process**

$$G_t^Z = e^{-rt}S_t + \int_0^t e^{-ru} dD_u$$

is a Q-martingale.

Proof: Exercise.

Note: The result above holds in great generality.

Interpretation:
In a risk neutral world, today's stock price should be the expected value of all future discounted earnings which arise from holding the stock.

$$S_0 = E^Q \left[\int_0^t e^{-ru} dD_u + e^{-rt} S_t \right],$$

Pricing formula

Pricing formula for claims of the type

$$\mathcal{Z} = \Phi(S_T)$$

We are standing at time t, with dividend yield q. Today's stock price is s.

- Suppose that you have the pricing function

$$F^0(t, s)$$

  for a non dividend stock.

- Denote the pricing function for the dividend paying stock by

$$F^q(t, s)$$

Proposition: With notation as above we have

$$F^q(t, s) = F^0\left(t, se^{-q(T-t)}\right)$$

Moral

Use your old formulas, but replace today's stock price s with $se^{-q(T-t)}$.

European Call on Dividend-Paying-Stock

$$F^q(t,s) = se^{-q(T-t)}N\left[d_1\right] - e^{-r(T-t)}KN\left[d_2\right].$$

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T-t}}\left\{\ln\left(\frac{s}{K}\right) + \left(r - q + \frac{1}{2}\sigma^2\right)(T-t)\right\} \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

Martingale Analysis

Basic task: We have a general model for stock price S and cumulative dividends D, under P. How do we find a martingale measure Q, and exactly which objects will be martingales under Q?

Main Idea: We attack this situation by reducing it to the well known case of a market without dividends. Then we apply standard techniques.

The Reduction Technique

- Consider the self financing portfolio where you keep 1 unit of the stock and invest all dividends in the bank. Denote the portfolio value by V.

- This portfolio can be viewed as a traded asset **without dividends**.

- Now apply the First Fundamental Theorem to the market (B, V) instead of the original market (B, S).

- Thus there exists a martingale measure Q such that $\frac{\Pi_t}{B_t}$ is a Q martingale for all traded assets (underlying and derivatives) without dividends.

- In particular the process

$$\frac{V_t}{B_t}$$

is a Q martingale.

Problems for discussion

- Suppose you get x dollars at time s. You put it into the bank and keep it until time t. How much money do you then have at the bank at time t?

- Derive an expression for the value V_t of the portfolio above using economic arguments. Recall the setup.
  - A general stock price process S
  - A general cumulative dividend process D.
  - A bank account with constant short rate r.
  - We hold one unit of the stock and invest all dividends in the bank account.

The V Process

Let h_t denote the number of units in the bank account, where $h_0 = 0$. V is then characterized by

$$V_t = 1 \cdot S_t + h_t B_t \tag{1}$$

$$dV_t = dS_t + dD_t + h_t dB_t \tag{2}$$

From (**??**) we obtain

$$dV_t = dS_t + h_t dB_t + B_t dh_t$$

Comparing this with (2) gives us

$$B_t dh_t = dD_t$$

Integrating this gives us

$$h_t = \int_0^t \frac{1}{B_s} dD_s$$

We thus have

$$V_t = S_t + B_t \int_0^t \frac{1}{B_s} dD_s \tag{3}$$

and the first fundamental theorem gives us the following result.

Proposition: For a market with dividends, the martingale measure Q is characterized by the fact that the **normalized gain process**

$$G_t^Z = \frac{S_t}{B_t} + \int_0^t \frac{1}{B_s} dD_s$$

is a Q martingale.

Quiz: Could you have guessed the formula (3) for V?

Continuous Dividend Yield

Model under P

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dD_t &= q S_t dt \end{aligned}$$

We recall

$$G_t^Z = \frac{S_t}{B_t} + \int_0^t \frac{1}{B_s} dD_s$$

Easy calculation gives us

$$dG_t^Z = Z_t \left(\alpha - r + q\right) dt + Z_t \sigma dW_t$$

where $Z = S/B$.

Girsanov transformation $dQ = LdP$, where

$$dL_t = L_t \varphi_t dW_t$$

We have

$$dW_t = \varphi_t dt + dW_t^Q$$

Insert this into dG^Z

The Q dynamics for G^Z are

$$dG_t^Z = Z_t \left(\alpha - r + q + \sigma\varphi_t\right) dt + Z_t \sigma dW_t^Q$$

Martingale condition

$$\alpha - r + q + \sigma\varphi_t = 0$$

Q-dynamics of S

$$dS_t = S_t \left(\alpha + \sigma\varphi\right) dt + S_t \sigma dW_t^Q$$

Using the martingale condition this gives us the Q-dynamics of S as

$$dS_t = S_t \left(r - q\right) dt + S_t \sigma dW_t^Q$$

Risk Neutral Valuation

Theorem: For a T-claim X, the price process $\Pi_t[X]$ is given by

$$\Pi_t[X] = e^{-r(T-t)} E^Q[X|\mathcal{F}_t],$$

where the Q-dynamics of S are given by

$$dS_t = (r-q)S_t dt + \sigma S_t dW_t^Q.$$

2. Forward and Futures Contracts

Forward Contracts

A **forward contract** on the T-claim X, **contracted at** t, is defined by the following payment scheme.

- The holder of the forward contract receives, at time T, the stochastic amount X from the underwriter.
- The holder of the contract pays, at time T, the **forward price** $f(t;T,X)$ to the underwriter.
- The forward price $f(t;T,X)$ is determined at time t.
- The forward price $f(t;T,X)$ is determined in such a way that the price of the forward contract equals zero, at the time t when the contract is made.

General Risk Neutral Formula

Suppose we have a bank account B with dynamics

$$dB_t = r_t B_t dt, \quad B_0 = 1$$

with a (possibly stochastic) short rate r_t. Then

$$B_t = e^{\int_0^t r_s ds}$$

and we have the following risk neutral valuation for a T-claim X

$$\Pi_t[X] = E^Q\left[e^{-\int_t^T r_s ds} \cdot X \,\middle|\, \mathcal{F}_t \right]$$

The price of a zero coupon bond

Setting $X = 1$ we have the price, at time t, of a zero coupon bond maturing at T as

$$p(t,T) = E^Q \left[e^{-\int_t^T r_s ds} \middle| \mathcal{F}_t \right]$$

Forward Price Formula

Theorem: The forward price of the claim X is given by

$$f(t,T) = \frac{1}{p(t,T)} E^Q \left[e^{-\int_t^T r_s ds} \cdot X \middle| \mathcal{F}_t \right]$$

where $p(t,T)$ denotes the price at time t of a zero coupon bond maturing at time T.

In particular, if the short rate r is deterministic we have

$$f(t,T) = E^Q \left[X \middle| \mathcal{F}_t \right]$$

Proof

The net cash flow at maturity is $X - f(t,T)$. If the value of this at time t equals zero we obtain

$$\Pi_t\left[X\right] = \Pi_t\left[f(t,T)\right]$$

We have

$$\Pi_t\left[X\right] = E^Q\left[e^{-\int_t^T r_s ds} \cdot X \middle| \mathcal{F}_t\right]$$

and, since $f(t,T)$ is known at t, we obviously (why?) have

$$\Pi_t\left[f(t,T)\right] = p(t,T)f(t,T).$$

This proves the main result. If r is deterministic then $p(t,T) = e^{-r(T-t)}$ which gives us the second formula.

Futures Contracts

A **futures contract** on the T-claim X, is a financial asset with the following properties.

(i) At every point of time t with $0 \leq t \leq T$, there exists in the market a quoted object $F(t;T,X)$, known as the **futures price** for X at t, for delivery at T.

(ii) At the time T of delivery, the holder of the contract pays $F(T;T,X)$ and receives the claim X.

(iii) During an arbitrary time interval $(s,t]$ the holder of the contract receives the amount $F(t;T,X) - F(s;T,X)$.

(iv) The spot price, at any time t prior to delivery, for buying or selling the futures contract, is by definition equal to zero.

Futures Price Formula

From the definition it is clear that a futures contract is a **price-dividend pair** (S, D) with

$$S \equiv 0, \quad dD_t = dF(t, T)$$

From general theory, the normalized gains process

$$G_t^Z = \frac{S_t}{B_t} + \int_0^t \frac{1}{B_s} dD_s$$

is a Q-martingale.

Since $S \equiv 0$ and $dD_t = dF(t, T)$ this implies that

$$\frac{1}{B_t} dF(t, T)$$

is a martingale increment, which implies (why?) that $dF(t, T)$ is a martingale increment. Thus F is a Q-martingale and we have

$$F(t, T) = E^Q \left[F(T, T) \middle| \mathcal{F}_t \right] = E^Q \left[X \middle| \mathcal{F}_t \right]$$

Theorem: The futures price process is given by

$$F(t,T) = E^Q\left[X\,|\,\mathcal{F}_t\right].$$

Corollary. If the short rate is deterministic, then the futures and forward prices coincide.

3. Futures Options

Futures Options

We denote the futures price process, at time t with delivery time at T by

$$F(t,T).$$

When T is fixed we sometimes suppress it and write F_t, i.e. $F_t = F(t,T)$

Definition:
A European futures call option, with strike price K and exercise date T, on a futures contract with delivery date T_1 will, if exercised at T, pay to the holder:

- The amount $F(T,T_1) - K$ in **cash**.
- A long postition in the underlying futures contract.

NB! The long position above can immediately be closed at no cost.

Institutional fact:
The exercise date T of the futures option is typcally very close to the date of delivery of the underlying T_1 futures contract.

Why do Futures Options exist?

- On many markets (such as commodity markets) the futures market is much more liquid than the underlying market.
- Futures options are typically settled in **cash**. This relieves you from handling the underlying (tons of copper, hundreds of pigs, etc.).
- The market place for futures and futures options is often the same. This facilitates hedging etc.

Pricing Futures Options – Black-76

We consider a futures contract with delivery date T_1 and use the notation $F_t = F(t, T_1)$. We assume the following dynamics for F.

$$dF_t = \mu F_t dt + \sigma F_t dW_t$$

Now suppose we want to price a derivative with exercise date T with the T_1-futures price F as underlying, i.e. a claim of the form

$$\Phi(F_T)$$

This turns out to be quite easy.

From risk neutral valuation we know that the price process $\Pi_t[\Phi]$ is of the form

$$\Pi_t[\Phi] = f(t, F_t)$$

where f is given by

$$f(t, F) = e^{-r(T-t)} E^Q_{t,F}[\Phi(F_T)]$$

so it only remains to find the Q-dynamics for F.

We now recall

Proposition: The futures price process F_t is a Q-martingale.

Thus the Q-dynamics of F are given by

$$dF_t = \sigma F_t dW_t^Q$$

We thus have

$$f(t, F) = e^{-r(T-t)} E^Q_{t,F} [\Phi(F_T)]$$

with Q-dynamics

$$dF_t = \sigma F_t dW^Q_t$$

Now recall the formula for a stock with continuous dividend yield q.

$$f(t, s) = e^{-r(T-t)} E^Q_{t,s} [\Phi(S_T)]$$

with Q-dynamics

$$dS_t = (r - q) S_t + \sigma S_t dW^Q_t$$

Note: If we set $q = r$ the formulas are **identical**!

Pricing Formulas

Let $f^0(t, s)$ be the pricing function for the contract $\Phi(S_T)$ for the case when S is a stock without dividends. Let $f(t, F)$ be the pricing formula for the claim $\Phi(F_T)$.

Proposition: With notation as above we have

$$f(t, F) = f^0(t, Fe^{-r(T-t)})$$

Moral: Reset today's futures price F to $Fe^{-r(T-t)}$ and use your formulas for stock options.

Black-76 Formula

The price of a futures option with exercise date T and exercise price K is given by

$$c = e^{-r(T-t)} \left\{ FN\left[d_1\right] - KN\left[d_2\right] \right\}.$$

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T-t}} \left\{ \ln\left(\frac{F}{K}\right) + \frac{1}{2}\sigma^2(T-t) \right\}, \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

Chapter 17

Currency Derivatives

Tomas Björk

Pure Currency Contracts

Consider two markets, domestic (England) and foreign (USA).

$$
\begin{aligned}
r^d &= \text{domestic short rate} \\
r^f &= \text{foreign short rate} \\
X &= \text{exchange rate}
\end{aligned}
$$

NB! The exchange rate X is quoted as

$$
\frac{\text{units of the domestic currency}}{\text{unit of the foreign currency}}
$$

Simple Model (Garman-Kohlhagen)

The P-dynamics are given as:

$$\begin{aligned} dX_t &= X_t \alpha dt + X_t \sigma dW_t, \\ dB_t^d &= r^d B_t^d dt, \\ dB_t^f &= r^f B_t^f dt, \end{aligned}$$

Main Problem:
Find arbitrage free price for currency derivative, Z, of the form

$$Z = \Phi(X_T)$$

Typical example: European Call on X.

$$Z = \max\left[X_T - K, 0\right]$$

Naive idea

For the European Call, use the standard Black-Scholes formula, with S replaced by X and r replaced by r^d.

Is this OK?

NO!

WHY?

Main Idea

- When you buy stock you just keep the asset until you sell it.
- When you buy dollars, these are put into a bank account, giving the interest r^f.

Moral:
Buying a currency is like buying a dividend-paying stock with dividend yield $q = r^f$.

Technique

- Transform all objects into **domestically traded** asset prices.
- Use standard techniques on the transformed model.

Transformed Market

1. Investing foreign currency in the foreign bank gives value dynamics **in foreign currency** according to

$$dB_t^f = r^f B_t^f dt.$$

2. B_f units of the foreign currency is worth $X \cdot B_f$ in the domestic currency.

3. Trading in the foreign currency is equivalent to trading in a domestic market with the domestic price process

$$S_t^f = B_t^f \cdot X_t$$

4. Study the domestic market consisting of

$$S^f, \quad B^d$$

Market dynamics

$$
\begin{aligned}
dX_t &= X_t \alpha dt + X_t \sigma dW \\
S_t^f &= B_t^f \cdot X_t
\end{aligned}
$$

Using Itô we have domestic market dynamics

$$
\begin{aligned}
S_t^f &= S_t^f \left(\alpha + r^f\right) dt + S_t^f \sigma dW_t \\
dB_t^d &= r^d B_t^d dt
\end{aligned}
$$

Standard results gives us Q-dynamics for domestically traded asset prices:

$$
\begin{aligned}
S_t^f &= S_t^f r^d dt + S_t^f \sigma dW_t^Q \\
dB_t^d &= r^d B_t^d dt
\end{aligned}
$$

Itô gives us Q-dynamics for $X_t = S_t^f / B_t^f$:

$$
dX_t = X_t(r^d - r^f)dt + X_t \sigma dW_t^Q
$$

Risk neutral Valuation

Theorem: The arbitrage free price $\Pi_t[\Phi]$ is given by $\Pi_t[\Phi] = F(t, X_t)$ where

$$F(t,x) = e^{-r^d(T-t)} E^Q_{t,x}[\Phi(X_T)]$$

The Q-dynamics of X are given by

$$dX_t = X_t(r^d - r^f)dt + X_t \sigma dW^Q_t$$

Pricing PDE

Theorem: The pricing function F solves the boundary value problem

$$\begin{aligned}\frac{\partial F}{\partial t}+x(r^d-r^f)\frac{\partial F}{\partial x}+\frac{1}{2}x^2\sigma_X^2\frac{\partial^2 F}{\partial x^2}-r^dF &= 0,\\ F(T,x) &= \Phi(x)\end{aligned}$$

Currency vs Equity Derivatives

Proposition: Introduce the notation:

- $F^0(t,x)$ = the pricing function for the claim $\mathcal{Z} = \Phi(X_T)$, where we interpret X as the price of an ordinary stock without dividends.
- $F(t,x)$ = the pricing function of the same claim when X is interpreted as an exchange rate.

Then the following holds

$$F(t,x) = F_0\left(t, xe^{-r^f(T-t)}\right).$$

Currency Option Formula

The price of a European currency call is given by

$$F(t,x) = xe^{-r^f(T-t)}N\left[d_1\right] - e^{-r^d(T-t)}KN\left[d_2\right],$$

where

$$d_1 = \frac{1}{\sigma_X\sqrt{T-t}}\left\{\ln\left(\frac{x}{K}\right) + \left(r^d - r^f + \frac{1}{2}\sigma_X^2\right)(T-t)\right\}$$

$$d_2 = d_1(t,x) - \sigma_X\sqrt{T-t}$$

Siegel's Paradox

Assume that the domestic and the foreign markets are risk neutral and assume constant short rates. We now have the following surprising (?) argument.

A: Let us consider a T claim of 1 dollar. The arbitrage free dollar value at $t = 0$ is of course

$$e^{-r^f T}$$

so the Euro value at at $t = 0$ is given by

$$X_0 e^{-r^f T}.$$

The 1-dollar claim is, however, identical to a T-claim of X_T euros. Given domestic risk neutrality, the Euro value at $t = 0$ is then

$$e^{-r^d T} E^P \left[X_T \right].$$

We thus have

$$X_0 e^{-r^f T} = e^{-r^d T} E^P \left[X_T \right]$$

Siegel's Paradox ct'd

B: We now consider a T-claim of one Euro and compute the dollar value of this claim. The Euro value at $t = 0$ is of course

$$e^{-r^d T}$$

so the dollar value is

$$\frac{1}{X_0} e^{-r^d T}.$$

The 1-Euro claim is identical to a T-claim of X_T^{-1} Euros so, by foreign risk neutrality, we obtain the dollar price as

$$e^{-r^f T} E^P \left[\frac{1}{X_T} \right]$$

which gives us

$$\frac{1}{X_0} e^{-r^d T} = e^{-r^f T} E^P \left[\frac{1}{X_T} \right]$$

Siegel's Paradox ct'd

Recall our earlier results

$$
\begin{aligned}
X_0 e^{-r^f T} &= e^{-r^d T} E^P \left[X_T \right] \\
\frac{1}{X_0} e^{-r^d T} &= e^{-r^f T} E^P \left[\frac{1}{X_T} \right]
\end{aligned}
$$

Combining these gives us

$$
E^P \left[\frac{1}{X_T} \right] = \frac{1}{E^P \left[X_T \right]}
$$

which, by Jensen's inequality, is impossible unless X_T is deterministic. This is sometimes referred to as (one formulation of) "Siegel's paradox."

It thus seems that Americans cannot be risk neutral at the same time as Europeans.

What is going on?

Martingale Analysis

$$\begin{aligned} Q^d &= \text{domestic martingale measure} \\ Q^f &= \text{foreign martingale measure} \end{aligned}$$

$$L_t = \frac{dQ^f}{dQ^d}, \quad L_t^d = \frac{dQ^d}{dP}, \quad L_t^f = \frac{dQ^f}{dP}$$

P-dynamics of X

$$dX_t = X_t \alpha_t dt + X_t \sigma_t dW_t$$

where α and σ are arbitrary adapted processes and W is P-Wiener.

Problem: How are Q^d and Q^f related?

Main Idea

Fix an arbitrary foreign T-claim Z.

- Compute foreign price and change to domestic currency. The price at $t = 0$ will be

$$\Pi_0[Z] = X_0 E^{Q^f}\left[e^{-\int_0^T r_s^f ds} Z\right]$$

This can be written as

$$\Pi_0[Z] = X_0 E^{Q^d}\left[L_T e^{-\int_0^T r_s^f ds} Z\right]$$

- Change into domestic currency at T and then compute arbitrage free price. This gives us

$$\Pi_0[Z] = E^{Q^d}\left[e^{-\int_0^T r_s^d ds} X_T \cdot Z\right]$$

- These expressions must be equal for all choices of $Z \in \mathcal{F}_T$.

We thus obtain

$$E^{Q^d}\left[e^{-\int_0^T r_s^d ds} X_T \cdot Z\right] = X_0 E^{Q^d}\left[L_T e^{-\int_0^T r_s^f ds} Z\right]$$

for all T-claims Z. This implies the following result.

Theorem: The exchange rate X is given by

$$X_t = X_0 e^{\int_0^t (r_s^d - r_s^f) ds} L_t$$

alternatively by

$$X_t = X_0 \frac{D_t^f}{D_t^d}$$

where D_t^d is the domestic stochastic discount factor etc.

Proof: The last part follows from

$$L = \frac{dQ^f}{dQ^d} = \frac{dQ^f}{dP} \Bigg/ \frac{dQ^d}{dP}$$

Q^d-Dynamics of X

In particular, since L is a Q^d-martingale the Q^d dynamics of L are of the form

$$dL_t = L_t \varphi_t dW_t^d$$

where W^d is Q^d-Wiener. From

$$X_t = X_0 e^{\int_0^t (r_s^d - r_s^f) ds} L_t$$

the Q^d-dynamics of X follows as

$$dX_t = (r_t^d - r_t^f) X_t dt + X_t \varphi_t dW_t^d$$

so the Girsanov kernel φ equals the exchange rate volatility σ and we have the general Q^d dynamics.

Theorem: The Q^d dynamics of X are of the form

$$dX_t = (r_t^d - r_t^f) X_t dt + X_t \sigma_t dW_t^d$$

Market Prices of Risk

Recall

$$D_t^d = e^{-\int_0^t r_s^d ds} L_t^d$$

We also have

$$dL_t^d = L_t^d \varphi_t^d dW_t$$

where $-\varphi_t^d = \lambda^d$ is the domestic market price of risk and similar for φ^f etc. From

$$X_t = X_0 \frac{D_t^f}{D_t^d}$$

we now easily obtain

$$dX_t = X_t \alpha_t dt + X_t \left(\lambda_t^d - \lambda_t^f \right) dW_t,$$

where we do not care about the exact shape of α. We thus have

Theorem: The exchange rate volatility is given by

$$\sigma_t = \lambda_t^d - \lambda_t^f$$

Siegel's Paradox

Sometimes it is assumed that (for computational simplicity) the market is risk neutral.

Question: Can we assume that both the domestic and the foreign markets are risk neutral?

Answer: Generally no.

Proof: The assumption would be equivalent to assuming the $P = Q^d = Q^f$ i.e.

$$\lambda_t^d = \lambda_t^f = 0$$

However, we know that

$$\sigma_t = \lambda_t^d - \lambda_t^f$$

so we would need to have $\sigma_t = 0$ i.e. a non-stochastic exchange rate.

Chapters 22-23

Bonds and Short Rate Models

Tomas Björk

Definitions

A **zero coupon bond** with maturity T (a "T-bond") is a contract paying $\$1$ at the date of maturity T.

$$\begin{aligned} p(t,T) &= \text{price, at } t \text{, of a } T\text{-bond.} \\ p(T,T) &= 1. \end{aligned}$$

Main Problem

- Investigate the **term structure**, i.e. how prices of bonds with different dates of maturity are related to each other.
- Compute arbitrage free prices of interest rate derivatives (bond options, swaps, caps, floors etc.)

Risk Free Interest Rates

At time t:

- Sell one S-bond
- Buy exactly $p(t,S)/p(t,T)$ $T-$bonds
- Net investment at t: ± 0.

At time S:

- Pay $\$1$

At time T:

- Collect $\$$ $p(t,S)/p(t,T) \cdot 1$

Net Effect

- The contract is made at t.
- An investment of 1 at time S has yielded $p(t,S)/p(t,T)$ at time T.
- The implied interest rate can be quoted in two different ways: As a **continuous** interest rate R, or as a **simple** interest rate L.

Continuous rate:

$$e^{R\cdot(T-S)} \cdot 1 = \frac{p(t,S)}{p(t,T)}$$

Simple rate:

$$[1 + L \cdot (T-S)] \cdot 1 = \frac{p(t,S)}{p(t,T)}$$

Continuous Interest Rates

1. The **continuously compounded forward rate** for the period $[S, T]$, contracted at t is defined by

$$R(t; S, T) = -\frac{\ln p(t,T) - \ln p(t,S)}{T - S}.$$

2. The **spot rate**, $R(t,T)$, for the period $[t,T]$ is defined by

$$R(t,T) = R(t; t, T).$$

3. The **instantaneous forward rate at T, conracted at** t is defined by

$$f(t,T) = -\frac{\partial \ln p(t,T)}{\partial T} = \lim_{S \to T} R(t; S, T).$$

4. The **instantaneous short rate at** t is defined by

$$r(t) = f(t,t).$$

Simple Rates (LIBOR)

1. The **simple forward rate** L(t,S,T) for the period $[S, T]$, contracted at t is defined by

$$L(t, S, T) = \frac{1}{T - S} \cdot \frac{p(t, S) - p(t, T)}{p(t, T)}$$

2. The **simple spot rate,** $L(t, T)$, for the period $[t, T]$ is defined by

$$L(t, T) = \frac{1}{T - t} \cdot \frac{1 - p(t, T)}{p(t, T)}$$

Forward vs Spot Rates

- The spot rate $L(t, T)$ is the risk free rate for the time interval $[t, T]$, contracted at t.
- The forward rate $L(t, S, T)$ is the risk free rate for the time interval $[S, T]$, contracted at t.
- The spot rate $L(S, T)$ is the risk free rate for the time interval $[S, T]$, contracted at S.

Forward vs Spot Rates ct'd

- The short rate r_t is the risk free rate for the time interval $[t, t + dt]$, contracted at t.
- The forward rate $f(t, T)$ is the risk free rate for the time interval $[T, T + dT]$, contracted at t.
- The short rate r_T is the risk free rate for the time interval $[T, T + dT]$, contracted at T.

An Expectation Hypothesis

- Both the forward rate $f(t,T)$ and the short rate r_T are risk free interest rates for $[T, T+dT]$.
- The forward rate $f(t,T)$ is known today.
- The future short rate r_T is known only at time T.

Conjecture: Is it the case that the forward rate $f(t,T)$ is an unbiased estimator for the future spot rate r_T?

Question: How do we formalize this? There are at least two possibilities:

$$f(t,T) = E^P\left[r_T \middle| \mathcal{F}_t \right]$$

and

$$f(t,T) = E^Q\left[r_T \middle| \mathcal{F}_t \right]$$

Which of these is true, if indeed any?

Bond prices $\sim$ forward rates

$$p(t,T) = p(t,s) \cdot e^{-\int_s^T f(t,u)du},$$

In particular we have

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

The Bank Account: Discrete Time

Definitions:

- The price at time n of a bond maturing at k is denoted by
$$p(n, k)$$

- The (possible stochastic) discrete **short rate** r_n, for the period $[n, n+1]$, is defined as
$$p(n, n+1) = \frac{1}{1 + r_n}$$

Roll-over strategy:

- At time n we invest the entire portfolio value in bonds maturing at time $n+1$.
- At time $n+1$ the bonds mature, and we invest everything in bonds maturing at $n+2$ etc. etc.
- The value process B_n is the **bank account**.

Dynamics of the Bank Account

By h_n we denote the number of bonds, bought at time n, and maturing at $n + 1$.

- We have (why?)

$$h_n = \frac{B_n}{p(n, n+1)}$$

- At $n + 1$ we have (why?)

$$B_{n+1} = h_n \cdot 1$$

- Thus

$$B_{n+1} = \frac{B_n}{p(n, n+1)} = B_n(1 + r_n)$$

- Thus

$$B_{n+1} - B_n = B_n r_n$$

Properties of the Bank Account

- We have
$$r_n = \frac{1 - p(n, n+1)}{p(n, n+1)}$$

- The (possibly stochastic) short rate r_n for the interval $[n, n+1]$ is thus known at time n.

- Recall
$$B_{n+1} - B_n = B_n r_n$$

- Thus the bank account is **locally riskless**, i.e. the return over the interval $[n, n+1]$ is risk free.

- Note that the return of B over a **longer** interval such as $[n, n+2]$ is **stochastic**.

The Bank Account: Continuous Time

The intuitive definition is as follows.

Definition:

- The bank account B is defined as a roll-over portfolio where at time t you invest the entire portfolio value in "just maturing" bonds.
- At time t you thus invest the entire portfolio in zero coupon bonds maturing at time $t + dt$.
- The value process for this roll-over portfolio is denoted B_t
- The dynamics of B are given by

$$dB_t = r_t B_t dt$$

- Compare with

$$B_{n+1} - B_n = B_n r_n$$

Caps

Basic idea: Buy an insurance against high interest rates in the future.

1. The contract is written at $t = 0$. At that time also the **principal**, K, and the fixed **cap rate**, R are determined.

2. The **resetttlement dates**

$$T_0 < T_1 < \ldots < T_n$$

are specified, with **tenor**

$$\alpha = T_{i+1} - T_i, \quad i = 0, \ldots, n-1.$$

Typically $\alpha = 1/4$, i.e. quarterly resettlement.

3. A cap is a sum of elementary cash flows, $X_1, \ldots, X_n$, paid at $T_1, \ldots, T_n$, called **caplets**.

Denote by $L_i(t)$ the LIBOR forward rate for $[T_{i-1}, T_i]$.

In particular, $L_i(T_{i-1})$ is the spot rate at time T_{i-1}.

Definition:
A **cap** with **cap rate** R, **nominal value** K, and **resettlement dates** $T_0, \ldots, T_n$ is a contract which at each T_i give the holder the amount

$$X_i = K \cdot \alpha \cdot \max\left[L_i(T_{i-1}) - R, 0\right], \quad i = 1, \ldots, N$$

The cap is thus a portfolio of **caplets** $X_1, \ldots, X_n$.

Recap: The Black-76 Formula

The price of a futures option with exercise date T and exercise price X is given by

$$c = e^{-r(T-t)} \left\{ F_t N\left[d_1\right] - XN\left[d_2\right] \right\}.$$

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T-t}} \left\{ \ln\left(\frac{F_t}{X}\right) + \frac{1}{2}\sigma^2(T-t) \right\}, \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

Idea: For a caplet with maturity T_i we...

- Replace the futures price F_t by the forward rate $L_i(t)$.
- Replace the strike X by the cap rate R.
- Replace $e^{-r(T-t)}$ by the market discount factor $p_i(t) = p(t, T_i)$.

Black-76:

The **Black-76** formula, at time t, for the caplet

$$X_i = \alpha_i \cdot \max\left[L(T_{i-1}, T_i) - R, 0\right], \qquad (4)$$

is obtained by simply applying the standard Black-76 formula for futures options, to forward rates:

$$\mathbf{Capl_i^B}(t) = \alpha \cdot p_i(t)\left\{L_i(t)N[d_1] - RN[d_2]\right\}$$

where

$$\begin{aligned} d_1 &= \frac{1}{\sigma_i\sqrt{T_i - t}}\left[\ln\left(\frac{L_i(t)}{R}\right) + \frac{1}{2}\sigma_i^2(T_{i-1} - t)\right], \\ d_2 &= d_1 - \sigma_i\sqrt{T_{i-1} - t}. \end{aligned}$$

- The constants $\sigma_1, \ldots, \sigma_N$ are known as the **Black volatilities**

Problems with Black-76

In the original Black-76 formula we assume the following.

- The underlying is lognormal.
- The underlying is a futures, or forward, **price**.
- The short rate is **constant**.

Using Black-76 for caplets means that:

- We need to assume that the LIBOR rates are lognormal. (Possible)
- We apply a formula for underlying forward **prices** to a contract with underlying forward **rates**. (Problematic)
- We use a formula for **constant** interest rates to compute the price of a contract which is relevent only for **random** interest rates. (Ouch!)

Deeply felt need

A consistent **arbitrage free** model for the bond market

Stochastic interest rates

We assume that the short rate r is a stochastic process.

Money in the bank will then grow according to:

$$\begin{cases} dB_t &= r_t B_t dt, \\ B_0 &= 1. \end{cases}$$

i.e.

$$B_t = e^{\int_0^t r_s ds}$$

We need a model for the short rate r.

Models for the short rate

P-dynamics

$$\begin{aligned} dr_t &= \mu(t, r_t)dt + \sigma(t, r_t)dW_t, \\ dB_t &= r_t B_t dt. \end{aligned}$$

Question: Are bond prices uniquely determined by the P-dynamics of r, and the requirement of an arbitrage free bond market?

NO!!
WHY?

Stock Models $\sim$ Interest Rates

Black-Scholes:

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= r B_t dt. \end{aligned}$$

Interest Rates:

$$\begin{aligned} dr_t &= \mu(t, r_t) dt + \sigma(t, r_t) dW_t, \\ dB_t &= r_t B_t dt. \end{aligned}$$

Question: What is the difference?

Answer: The short rate r is **not the price of a traded asset!**

1. **Rule of Thumb:**

$$\begin{aligned} N &= 0, \quad \text{(no risky asset)} \\ R &= 1, \quad \text{(one source of randomness, } W\text{)} \end{aligned}$$

   We have $N < R$. The exogenously given market, consisting only of B, is incomplete.

2. **Replicating portfolios:**
   We can only invest money in the bank, and then sit down passively and wait.

We do **not** have **enough underlying assets** in order to price bonds.

- There is **not** a unique price for a **particular** $T-$bond.

- In order to avoid arbitrage, bonds of **different maturities** have to satisfy internal **consistency** relations.

- If we take **one** "benchmark" T_0-bond as given, then all other bonds can be priced **in terms of** the market price of the benchmark bond.

Assumption:

$$
\begin{aligned}
p(t,T) &= F(t, r_t, T) \\
p(t,T) &= F^T(t, r_t), \\
F^T(T,T) &= 1.
\end{aligned}
$$

Program:

- Form portfolio based on T and S bonds. Use Itô on $F^T(t, r_t)$ to get bond- and portfolio dynamics.

$$dV = V \left\{ u^T \frac{dF^T}{F^T} + u^S \frac{dF^S}{F^S} \right\}$$

- Choose portfolio weights such that the $dW-$ term vanishes. Then we have

$$dV = V \cdot k dt,$$

("synthetic bank" with k as the short rate)

- Absence of arbitrage $\Rightarrow k = r$.

- Read off the relation $k = r$!

Notation:

$$F_t = \frac{\partial F}{\partial t}, \quad F_r = \frac{\partial F}{\partial r}, \qquad F_{rr} = \frac{\partial^2 F}{\partial r^2}$$

From Itô:

$$dF^T = F^T \alpha_T dt + F^T \sigma_T dW,$$

where

$$\begin{cases} \alpha_T & = & \frac{F_t^T + \mu F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T}{F^T}, \\ \sigma_T & = & \frac{\sigma F_r^T}{F^T}. \end{cases}$$

Portfolio dynamics

$$dV = V \left\{ u^T \frac{dF^T}{F^T} + u^S \frac{dF^S}{F^S} \right\}.$$

Reshuffling terms gives us

$$dV = V \cdot \left\{ u^T \alpha_T + u^S \alpha_S \right\} dt + V \cdot \left\{ u^T \sigma_T + u^S \sigma_S \right\} dW.$$

Let the portfolio weights solve the system

$$\begin{cases} u^T + u^S & = & 1, \\ u^T \sigma_T + u^S \sigma_S & = & 0. \end{cases}$$

$$\begin{cases} u^T & = & -\frac{\sigma_S}{\sigma_T - \sigma_S}, \\ u^S & = & \frac{\sigma_T}{\sigma_T - \sigma_S}, \end{cases}$$

Portfolio dynamics

$$dV = V \cdot \left\{ u^T \alpha_T + u^S \alpha_S \right\} dt.$$

i.e.

$$dV = V \cdot \left\{ \frac{\alpha_S \sigma_T - \alpha_T \sigma_S}{\sigma_T - \sigma_S} \right\} dt.$$

Absence of arbitrage requires

$$\frac{\alpha_S \sigma_T - \alpha_T \sigma_S}{\sigma_T - \sigma_S} = r$$

which can be written as

$$\frac{\alpha_S - r}{\sigma_S} = \frac{\alpha_T - r}{\sigma_T}.$$

$$\frac{\alpha_S(t, r_t) - r_t}{\sigma_S(t, r_t)} = \frac{\alpha_T(t, r_t) - r_t}{\sigma_T(t, r_t)}.$$

Note!

The quotient does **not** depend upon the particular choice of maturity date.

Result

Assume that the bond market is free of arbitrage. Then there exists a universal process λ, such that

$$\frac{\alpha_T(t, r_t) - r_t}{\sigma_T(t, r_t)} = \lambda(t, r_t),$$

holds for all t and for every choice of maturity T.

NB: The same λ for all choices of T.

$$\begin{aligned} \lambda &= \text{Risk premium per unit of volatility} \\ &= \text{“Market Price of Risk” (cf. CAPM).} \end{aligned}$$

Slogan:
“On an arbitrage free market all bonds have the same market price of risk.”

The relation

$$\frac{\alpha_T - r}{\sigma_T} = \lambda$$

is actually a PDE!

The Term Structure Equation

$$\begin{cases} F_t^T + \{\mu - \lambda\sigma\} F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T - rF^T & = & 0, \\ F^T(T,r) & = & 1. \end{cases}$$

P**-dynamics:**

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

$$\lambda = \frac{\alpha_T - r}{\sigma_T}, \quad \text{for alla } T$$

In order to solve the TSE we need to know λ.

General Term Structure Equation

Contingent claim:

$$\mathcal{Z} = \Phi(r_T)$$

Result:
The price is given by

$$\Pi_t\left[\mathcal{Z}\right] = F(t, r_t)$$

where F solves

$$\begin{cases} F_t + \{\mu - \lambda\sigma\} F_r + \frac{1}{2}\sigma^2 F_{rr} - rF & = & 0, \\ F(T, r) & = & \Phi(r). \end{cases}$$

In order to solve the TSE we need to know λ.

Risk Neutral Valuation

We have the pricing PDE

$$\begin{cases} F_t + \{\mu - \lambda\sigma\} F_r + \frac{1}{2}\sigma^2 F_{rr} - rF & = & 0, \\ F(T, r) & = & \Phi(r). \end{cases}$$

This can of course be attacked by Kolmogorov-Feynman-Kac.

PDE $\sim$ SDE

With a slight extension of the standard Feynman-Kac, one can show that the following are equivalent.

- The function $f(t,x)$ solves the PDE

$$\left\{ \begin{array}{rcl} \dfrac{\partial F}{\partial t} + a(t,x)\dfrac{\partial F}{\partial x} + \dfrac{1}{2}b^2(t,x)\dfrac{\partial^2 F}{\partial x^2} - k(x)F & = & 0, \\ F(T,x) & = & \Phi(x). \end{array} \right.$$

- The function $F(t,x)$ is given by the relation

$$F(t,x) = E^Q_{t,x}\left[e^{-\int_t^T k(X_s)ds}\Phi(X_T)\right],$$

where the Q dynamics of X are given by

$$dX_t = a(t,X_t)dt + b(t,X_t)dW_t^Q$$

Risk Neutral Valuation

In our case

$$\begin{cases} F_t + \{\mu - \lambda\sigma\} F_r + \frac{1}{2}\sigma^2 F_{rr} - rF &= 0 \\ F(T, r) &= \Phi(r), \end{cases}$$

We obtain

$$F(t, r) = E_{t,r}^Q \left[e^{-\int_t^T r_s ds} \times \Phi(r_T) \right].$$

with Q-dynamics

$$dr_t = \{\mu - \lambda\sigma\} dt + \sigma dW_t^Q$$

Bond prices are given by

$$p(t, T) = F^T(t, r_t)$$

where

$$F^T(t, r) = E_{t,r}^Q \left[e^{-\int_t^T r_s ds} \right].$$

General Risk Neutral Valuation

Derivative prices are given by

$$\Pi_t\left[\mathcal{Z}\right] = E^Q\left[e^{-\int_t^T r_s ds} \times \mathcal{Z}\middle|\, \mathcal{F}_t\right].$$

with Q-dynamics

$$dr_t = \{\mu(t, r_t) - \lambda(t, r_t)\sigma(t, r_t)\}dt + \sigma(t, r_t)dW_t^Q$$

Moral:

- Price = expected discounted value of future payments
- The expectation should **not** be taken under the "objective" probabilities P, but under the "risk adjusted" probabilities Q.

To compute the expected value (or solve the previous PDE) we need to know λ.

Question:

Who determines λ?

Answer:

THE MARKET!

Moral

- Since the market is incomplete the requirement of an arbitrage free bond market will not lead to unique bond prices.

- Prices on bonds and other interest rate derivatives are determined by two main factors.

  1. **Partly** by the requirement of an arbitrage free bond market (the pricing functions satisfies the TSE).
  2. **Partly** by supply and demand on the market. These are in turn determined by attitude towards risk, liquidity consideration and other factors. All these are aggregated into the particular λ used (implicitly) by the market.

Martingale Analysis

Model: Under P we have

$$\begin{aligned} dr_t &= \mu(t, r_t)\, dt + \sigma(t, r_t)\, dW_t, \\ dB_t &= rB_t dt, \end{aligned}$$

We look for martingale measures. Since B is the only traded asset we need to find $Q \sim P$ such that

$$\frac{B_t}{B_t} = 1$$

is a Q martingale.

Result: In a short rate model, **every** $Q \sim P$ is a martingale measure.

Girsanov

$$dL_t = L_t \varphi_t dW_t$$

P-dynamics

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

$$dL_t = L_t \varphi_t dW_t$$

$dQ = L_t dP$ on $\mathcal{F}_t$

Girsanov:

$$dW_t = \varphi_t dt + dW_t^Q$$

Martingale pricing:

$$\Pi_t\left[Z\right] = E^Q\left[e^{-\int_t^T r_s ds} Z \middle| \mathcal{F}_t \right]$$

Q-dynamics of r:

$$dr_t = \left\{\mu\left(t, r_t\right) + \sigma\left(t, r_t\right)\varphi_t\right\} dt + \sigma\left(t, r_t\right) dW_t^Q,$$

Result: We have $\lambda_t = -\varphi_t$, i.e,. the Girsanov kernel φ equals minus the market price of risk.

Chapter 24

Martingale Models for the Short Rate

Tomas Björk

Contents

I. Recap

Recap

P-dynamics for the short rate.

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

The price of a T-claim $\Phi(r_T)$ is given by

$$\Pi_t[\Phi] = F(t, r_t)$$

where F solves

$$\begin{cases} F_t + \{\mu - \lambda\sigma\} F_r + \frac{1}{2}\sigma^2 F_{rr} - rF & = & 0, \\ F(T, r) & = & \Phi(r). \end{cases}$$

Risk neutral valuation

P-dynamics for the short rate.

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

Risk neutral valuation:

$$\Pi_t [\mathcal{Z}] = E^Q \left[e^{-\int_t^T r_s ds} \times \mathcal{Z} \right| \left. \mathcal{F}_t \right]$$

Q**-dynamics:**

$$dr_t = \{\mu(t, r_t) + \varphi_t \sigma(t, r_t)\} dt + \sigma(t, r_t)dW_t^Q$$

II. Martingale Modeling

Martingale Modeling

Basic Idea:

- All prices are determined by Q-dynamics of r.
- Model dr directly under Q!

Standard procedure: In the literature (theoretical and applied) it is common to **model the relevant interest rates directly under a risk neutral measure** Q. We will now follow this approach.

Problem: Parameter estimation!

Note: Observe that, for simplicity of notation, μ, σ and W will, from now on, denote the drift, diffusion and Wiener process **under the risk neutral measure** Q. Also recall that σ is the same under P and Q.

Pricing under Q

We model the short rate r directly under Q.

Q-dynamics:

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

The price of a T-claim $\mathcal{Z}$ is given by

$$\begin{aligned} \Pi_t [\mathcal{Z}] &= E_t^Q \left[e^{-\int_t^T r_s ds} \times \mathcal{Z} \right] \\ p(t, T) &= E_t^Q \left[e^{-\int_t^T r_s ds} \times 1 \right] \end{aligned}$$

The case $\mathcal{Z} = \Phi(r_T)$:

$$\begin{aligned} F_t + \mu F_r + \frac{1}{2}\sigma^2 F_{rr} - rF &= 0, \\ F(T, r) &= \Phi(r). \end{aligned}$$

Models for the Short Rate

1. Vasicek

$$dr_t = (b - ar_t)\,dt + \sigma dW_t,$$

2. Cox-Ingersoll-Ross

$$dr_t = (b - ar_t)\,dt + \sigma\sqrt{r_t}dW_t,$$

3. Dothan

$$dr_t = ar_t dt + \sigma r_t dW_t,$$

4. Black-Derman-Toy

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

5. Ho-Lee

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

6. Hull-White (extended Vasicek)

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

Properties of the models

1. Models with linear dynamics
2. Models with mean reversion
3. Lognormal models
4. Models with positive interest rates
5. Affine Term Structure Models

1. Models with linear dynamics

- Vasicek

$$dr_t = (b - ar_t)\,dt + \sigma dW_t,$$

- Ho-Lee

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

- Hull-White extended Vasicek

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

These models all lead to a **normally distributed short rate**.

- This is **good** from a computational point of view.

- It is **bad** from an economic point of view, since we may then have negative nominal interest rates.

2. Models with mean reversion

- Vasicek

$$dr_t = (b - ar_t)\,dt + \sigma dW_t,$$

- Hull-White extended Vasicek

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

- Hull-White extended Cox-Ingersoll-Ross

$$dr_t = \{\Phi(t) - ar_t\}\,dt + \sigma\sqrt{r_t}dW_t,$$

All these models exhibit **mean reversion**, i.e. they tend to revert to a (possibly time dependent) mean value. This is reasonable from an economic point of view (why?).

Mean Reversion ct'd

As an example we consider the Vasicek model

$$dr_t = (b - ar_t)\, dt + \sigma dW_t,$$

where all parameters are assumed to be positive.

Write the model as

$$dr_t = a\left(\frac{b}{a} - r_t\right) dt + \sigma dW_t,$$

- If $r_t < \frac{b}{a}$ the drift is positive, and r has a tendency to increase.
- If $r_t > \frac{b}{a}$ the drift is negative, and r has a tendency to decrease.
- The short rate r will thus have a tendency to revert to the value b/a.
- One can in fact show that r has a limiting Gaussian distribution with mean b/a.

3. Lognormal models

- Dothan

$$dr_r = ar_t dt + \sigma r_t dW_t,$$

- Black-Derman-Toy

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

For these models the short rate is **lognormal** (why?).

- Nice, since the short rate is then always positive.
- Not nice, since these models are terrible from a computational point of view.
- Not nice, since $E^Q[B(t)] = +\infty$ for every $t > 0$ which leads to nonsensical pricing rules.

NB: These properties refer to the continuous time versions of the models. There is a tree version of BDT which is used quite a lot in practice.

5. Models with positive interest rates

- CIR

$$dr_t = \{b - ar_t\}\, dt + \sigma\sqrt{r_t} dW_t,$$

- Hull-White extended CIR

$$dr_t = \{\Phi(t) - ar_t\}\, dt + \sigma\sqrt{r_t} dW_t,$$

- Dothan

$$dr_t = ar_t dt + \sigma r_t dW_t,$$

- Black-Derman-Toy

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

Dothan and BDT are already discusssed.

CIR and positive interest rates

Model:

$$dr_t = \{b - ar_t\}\, dt + \sigma\sqrt{r_t}dW_t,$$

Assume a, b and σ are positive.

Intuitive argument for positivity

- Suppose that $r_t = 0$.
- Then the diffusion part $\sigma\sqrt{r_t}$ vanishes.
- Thus we have
$$dr_t = bdt$$
so the r process increases from $r = 0$ (to a positive value).
- Thus the r process never becomes negative.

This is just the basic intuition, but it can be shown that if $2b > \sigma^2$ then the r process stays strictly positive.

4 Affine Term Structure Models

Definition: A short rate models has an **affine term structure** if the bond prices are of the form

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

where A and B are deterministic.

Moral: The ATS models are the **only ones** who are **computationally tractable**.

Precise Result for ATS

Theorem: Assume that μ and σ are of the form

$$\begin{aligned} \mu(t,r) &= \alpha(t)r + \beta(t), \\ \sigma^2(t,r) &= \gamma(t)r + \delta(t). \end{aligned}$$

Then the model admits an affine term structure

$$F(t,r;T) = e^{A(t,T) - B(t,T)r},$$

where A and B satisfy the system

$$\begin{cases} B_t(t,T) = -\alpha(t)B(t,T) + \frac{1}{2}\gamma(t)B^2(t,T) - 1, \\ B(T;T) = 0. \end{cases}$$

$$\begin{cases} A_t(t,T) = \beta(t)B(t,T) - \frac{1}{2}\delta(t)B^2(t,T), \\ A(T;T) = 0. \end{cases}$$

Proof: It is easy to see that F defined as above satisfies the relevant PDE.

Affine Term Structure Models

- Vasicek

$$dr_t = (b - ar_t)\,dt + \sigma dW_t,$$

- Cox-Ingersoll-Ross

$$dr_t = (b - ar_t)\,dt + \sigma\sqrt{r_t}dW_t,$$

- Ho-Lee

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

- Hull-White (extended Vasicek)

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

- Hull-White extended CIR

$$dr_t = \{\Phi(t) - ar_t\}\,dt + \sigma\sqrt{r_t}dW_t,$$

III. Inverting the Yield Curve

Parameter Estimation

Suppose that we have chosen a specific model, e.g. Vasicek:

$$dr_t = (b - ar_t)\, dt + \sigma dW_t,$$

How do we estimate the parameters $a,\ b,\ \sigma$?

Naive answer:
Use standard methods from statistical theory.

NONSENSE!!

Why?

- The model parameters are Q-parameters.
- Our observations are **not** under Q, but under P.
- Standard statistical techniques can **not** be used.
- We need to know the market price of risk λ.
- Who determines λ?
- **The Market!**
- We must get **price information from the market** in order to estimate parameters.

Inversion of the Yield Curve

We consider a model having the following Q-dynamics with parameter list α. For Vasicek we would for example have $\alpha = (a, b, \sigma)$

$$dr = \mu(t, r; \alpha)dt + \sigma(t, r; \alpha)dW$$

- By solving the PDE we obtain the **theoretical** term structure (bond price curve)

$$p(0, T; \alpha); \quad T \geq 0$$

- By going to the market we obtain the **observed** term structure (bond price curve)

$$p^{\star}(0, T); \quad T \geq 0.$$

Basic Idea

We would like to have a model such that the **theoretical** prices of today coincide with the **observed** prices of today. We thus want to choose the parameter vector α such that

$$p(0,T;\alpha) = p^{\star}(0,T), \quad \text{for all } T \geq 0$$

- This is a system of equations where α is the unknown.
- Number of equations $= \infty$ (one for each T).
- Number of unknowns $=$ number of parameters.

Need:
Infinite parameter list.

The time dependent function Φ in Hull-White and Ho-Lee is precisely such an infinite parameter list (one parameter value $\Phi(t)$ for every t).

Result

- The Hull-White extensions of Vasicek and CIR, as well as the Ho-Lee models can be calibrated exactly to any initial term structure.

- Example: For the Ho-Lee model, the calibrated model has the form

$$dr_t = \left\{ f_T^\star(0,t) + \sigma^2 t \right\} dt + \sigma dW_t$$

where the observed forward rates are defined by

$$f^\star(0,T) = -\frac{\partial}{\partial T} \ln p^\star(0,T)$$

and

$$f_T^\star(0,T) = \frac{\partial}{\partial T} f^\star(0,T)$$

- There are analytical formulas for interest rate options.

Short rate models

Pro:

- Easy to model r.
- Analytical formulas for bond prices and bond options.

Con:

- Inverting the yield curve can be hard work.
- Hard to model a flexible volatility structure for forward rates.
- With a one factor model, all points on the yield curve are perfectly correlated.

Chapter 25

Foprward Rate Models

Tomas Björk

Recap

The instantaneous forward rate with maturity T, contracted at t is defined as

$$f(t,T) = \frac{\partial}{\partial T} \ln p(t,T)$$

Bond prices are then given by

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

Heath-Jarrow-Morton

Idea: Model the dynamics of the **entire forward rate curve**.

The forward rate curve itself (rather than the short rate r) is the explanatory variable.

Model forward rates. Use the observed forward rate curve as initial data.

Q-dynamics:

$$\begin{aligned} df(t,T) &= \alpha(t,T)dt + \sigma(t,T)dW_t, \\ f(0,T) &= f^{\star}(0,T). \end{aligned}$$

One SDE for each maturity date T.

Q-dynamics

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t$$

$$f(t,T) = \frac{\partial \log p(t,T)}{\partial T},$$

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

- Specifying forward rate dynamics is equivalent to specifying bond dynamics.
- For bond prices we know that the **mean rate of return under Q equals the short rate r_t.**
- Thus, modeling forward rates under Q implies restrictions on α and σ.
- Which are these restrictions?

Practical Toolbox

Theorem: Assume that the forward rate dynamics are given by

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW$$

Then the bond price dynamics are given by

$$\begin{aligned} dp(t,T) &= p(t,T)\left\{r(t) + A(t,T) + \frac{1}{2}\|S(t,T)\|^2\right\}dt \\ &+ p(t,T)S(t,T)dW \end{aligned}$$

where

$$\begin{cases} A(t,T) = -\int_t^T \alpha(t,s)ds, \\ S(t,T) = -\int_t^T \sigma(t,s)ds \end{cases}$$

HJM Drift Condition

Theorem: Under the **risk neutral** measure Q, the following must hold.

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds.$$

Proof: Follows from toolbox. ∎

Moral: The volatility can be specified freely. The forward rate drift is then uniquely specified.

Example

We consider the simplest possible forward rate model where $\sigma(t,T)$ is constant for all t and T.

From the drift condition we have

$$\alpha(t,T) = \sigma \int_t^T \sigma ds = \sigma^2(T-t)$$

Forward rate dynamics:

$$\begin{aligned} df(t,T) &= \sigma^2(T-t)dt + \sigma dW_t, \\ f(0,T) &= f^\star(0,T). \end{aligned}$$

Fix T and integrate over the interval $[0,t]$

$$f(t,T) = f^\star(0,T) + \int_0^t \sigma^2(T-s)ds + \int_0^t \sigma dW_s$$

$$f(t,T) = f^\star(0,T) + \sigma^2 Tt - \frac{\sigma^2}{2}t^2 + \sigma W_t$$

Example ct'd

Recall:

$$f(t,T) = f^{\star}(0,T) + \sigma^2 Tt - \frac{\sigma^2}{2}t^2 + \sigma W_t$$

We see that the forward rate curve has random horizontal shifts.

The short rate $r_t = f(t,t)$ is given by

$$r_t = f^{\star}(0,t) + \frac{\sigma^2}{2}t^2 + \sigma W_t$$

and thus

$$dr_t = \left\{ f_T^{\star}(0,t) + \sigma^2 t \right\} dt + \sigma dW_t$$

This is a well known short rate model. Which?

HJM in Practice

$$df(t,T) = \alpha(t,T)dt + \sum_{i=1}^{n} \sigma_i(t,T)dW_t^i$$

- Very often the volatilities $\sigma_1, \ldots, \sigma_n$ are chosen as **deterministic** functions of t and T.

- A constant σ will lead to forward rate curves $f(t,T)$ such that
$$\lim_{T \to \infty} f(t,T) = +\infty$$

- An exponential volatility term of the form

$$\sigma(t,T) = p(T-t)e^{-a(T-t)}$$

where p is a polynomial, roughly corresponds to mean reversion.

- In applications it is common to do PCA on the historical forward rate curves. Typically one find three main factors, implying that $n = 3$.

- The factors coming out of the PCA are then chosen as the volatility functions in the HJM model

Forward rate models

Pro:

- Easy to model a flexible volatility structure for forward rates.
- Easy to include multiple factors.

Con:

- The short rate will generically not be a Markov process.
- Hard computational problems.
- Numerical procedures.

Chapter 26

Change of Numeraire

Tomas Björk

Recap of General Theory

Consider a market with asset prices

$$S_t^0, S_t^1, \ldots, S_t^N$$

Theorem: The market is arbitrage free

iff

there exists an EMM, i.e. a measure Q such that

- Q and P are equivalent, i.e.

$$Q \sim P$$

- The normalized price processes

$$\frac{S_t^0}{S_t^0}, \frac{S_t^1}{S_t^0}, \ldots, \frac{S_t^N}{S_t^0}$$

  are Q-martingales.

Recap continued

Recall the normalized market

$$\left(Z_t^0, Z_t^1, \dots Z_t^N\right) = \left(\frac{S_t^0}{S_t^0}, \frac{S_t^1}{S_t^0} \dots, \frac{S_t^N}{S_t^0}\right)$$

- We obviously have

$$Z_t^0 \equiv 1$$

- Thus Z^0 is a risk free asset in the normalized economy.

- Z^0 is a bank account in the normalized economy.

- In the normalized economy **the short rate is zero**.

Dependence on numeraire

- The EMM Q will obviously depend on the choice of numeraire, so we should really write Q^0 to emphasize that we are using S^0 as numeraire.

- So far we have only considered the case when the numeraire asset is the bank account, i.e. when $S_t^0 = B_t$. In this case, the martingale measure Q^B is referred to as "the risk neutral martingale measure".

- Henceforth the notation Q (without upper case index) will only be used for the risk neutral martingale measure, i.e. $Q = Q^B$.

- We will now consider the case of a general numeraire.

General change of numeraire.

- Consider a financial market, including a bank account B.
- Assume that the market is using a fixed risk neutral measure Q as pricing measure.
- Choose a fixed asset S as numeraire, and denote the corresponding martingale measure by Q^S.

Problems:

- Determine Q^S, i.e. determine

$$L_t = \frac{dQ^S}{dQ}, \quad \text{on } \mathcal{F}_t$$

- Develop pricing formulas for contingent claims using Q^S instead of Q.

Constructing $\mathbf{Q}^S$

Fix a T-claim X. From general theory we know that

$$\Pi_0[X] = E^Q\left[\frac{X}{B_T}\right]$$

Since Q^S is a martingale measure for the numeraire S, the normalized process

$$\frac{\Pi_t[X]}{S_t}$$

is a Q^S-martingale. We thus have

$$\frac{\Pi_0[X]}{S_0} = E^S\left[\frac{\Pi_T[X]}{S_T}\right] = E^S\left[\frac{X}{S_T}\right] = E^Q\left[L_T\frac{X}{S_T}\right]$$

From this we obtain

$$\Pi_0[X] = E^Q\left[L_T\frac{X \cdot S_0}{S_T}\right],$$

For all $X \in \mathcal{F}_T$ we thus have

$$E^Q \left[\frac{X}{B_T} \right] = E^Q \left[L_T \frac{X \cdot S_0}{S_T} \right]$$

Recall the following basic result from probability theory.

Proposition: Consider a probability space $(\Omega, \mathcal{F}, P)$ and assume that

$$E\left[Y \cdot X\right] = E\left[Z \cdot X\right], \quad \text{for all } Z \in \mathcal{F}.$$

Then we have

$$Y = Z, \quad P - a.s.$$

From this result we conclude that

$$\frac{1}{B_T} = L_T \frac{S_0}{S_T}$$

Main result

Proposition: The likelihood process

$$L_t = \frac{dQ^S}{dQ}, \quad \text{on } \mathcal{F}_t$$

is given by

$$L_t = \frac{S_t}{B_t} \cdot \frac{1}{S_0}$$

Easy exercises

1. Convince yourself that L is a Q-martingale.

2. Assume that a process A_t has the property that A_t/B_t is a Q martingale. Show that this implies that A_t/S_t is a Q^S-martingale. Interpret the result.

Pricing

Theorem: For every T-claim X we have the pricing formula

$$\Pi_t[X] = S_t E^S \left[\left. \frac{X}{S_T} \right| \mathcal{F}_t \right]$$

Proof: Follows directly from the Q^S-martingale property of $\Pi_t[X]/S_t$. ∎

Note 1: We observe S_t directly on the market.

Note 2: The pricing formula above is particularly useful when X is of the form

$$X = S_T \cdot Y$$

In this case we obtain

$$\Pi_t[X] = S_t E^S [Y | \mathcal{F}_t]$$

Important example

Consider a claim of the form

$$X = \Phi\left[S_T^0, S_T^1\right]$$

We assume that Φ is **linearly homogeneous**, i.e.

$$\Phi(\lambda x, \lambda y) = \lambda \Phi(x, y), \quad \text{for all } \lambda > 0$$

Using Q^0 we obtain

$$\Pi_t\left[X\right] = S_t^0 E^0\left[\left.\frac{\Phi\left[S_T^0, S_T^1\right]}{S_T^0}\right| \mathcal{F}_t\right]$$

$$\Pi_t\left[X\right] = \Pi_t\left[X\right] = S_t^0 E^0\left[\left.\Phi\left(1, \frac{S_T^1}{S_T^0}\right)\right| \mathcal{F}_t\right]$$

Important example cnt'd

Proposition: For a claim of the form

$$X = \Phi\left[S_T^0, S_T^1\right],$$

where Φ is homogeneous, we have

$$\Pi_t[X] = S_t^0 E^0\left[\varphi(Z_T)|\,\mathcal{F}_t\right]$$

where

$$\varphi(z) = \Phi[1, z], \quad Z_t = \frac{S_t^1}{S_t^0}$$

Exchange option

Consider an exchange option, i.e. a claim X given by

$$X = \max\left[S_T^1 - S_T^0,\ 0\right]$$

Since $\Phi(x, y) = \max[x - y, 0]$ is homogeneous we obtain

$$\Pi_t[X] = S_t^0 E^0\left[\max\left[Z_T - 1,\ 0\right]\middle|\ \mathcal{F}_t\right]$$

- This is a European Call on Z with strike price K.
- Zero interest rate.
- Piece of cake!
- If S^0 and S^1 are both GBM, then so is Z, and the price will be given by the Black-Scholes formula.

Identifying the Girsanov Transformation

Assume the Q-dynamics of S are known as

$$dS_t = r_t S_t dt + S_t \sigma_t dW_t^Q$$

$$L_t = \frac{S_t}{S_0 B_t}$$

From this we immediately have

$$dL_t = L_t \sigma_t dW_t^Q.$$

and we can summarize.

Theorem: The Girsanov kernel is given by the numeraire volatility σ_t, i.e.

$$dL_t = L_t \sigma_t dW_t^Q.$$

Recap on zero coupon bonds

Recall: A zero coupon T-bond is a contract which gives you the claim

$$X \equiv 1$$

at time T.

The price process $\Pi_t[1]$ is denoted by $p(t, T)$.

Allowing a stochastic short rate r_t we have

$$dB_t = r_t B_t dt.$$

This gives us

$$B_t = e^{\int_0^t r_s ds},$$

and using standard risk neutral valuation we have

$$p(t, T) = E^Q \left[e^{-\int_t^T r_s ds} \middle| \mathcal{F}_t \right]$$

Note:

$$p(T, T) = 1$$

The forward measure Q^T

- Consider a fixed T.
- Choose the bond price process $p(t,T)$ as numeraire.
- The corresponding martingale measure is denoted by Q^T and referred to as "the T-forward measure".

For any T claim X we obtain

$$\Pi_t[X] = p(t,T) E^{Q^T} \left[\left. \frac{\Pi_T[X]}{p(T,T)} \right| \mathcal{F}_t \right]$$

We have

$$\Pi_T[X] = X, \quad p(T,T) = 1$$

Theorem: For any T-claim X we have

$$\Pi_t[X] = p(t,T) E^{Q^T} [X | \mathcal{F}_t]$$

A general option pricing formula

European call on asset S with strike price K and maturity T.

$$X = \max\left[S_T - K,\ 0\right]$$

Write X as

$$X = (S_T - K) \cdot I\left\{S_T \geq K\right\} = S_T I\left\{S_T \geq K\right\} - K I\left\{S_T \geq K\right\}$$

Use Q^S on the first term and Q^T on the second.

$$\Pi_0\left[X\right] = S_0 \cdot Q^S\left[S_T \geq K\right] - K \cdot p(0,T) \cdot Q^T\left[S_T \geq K\right]$$

Chapter 19

Stochastic Control Theory

Tomas Björk

Contents

1. Dynamic programming.
2. Investment theory.

1. Dynamic Programming

- The basic idea.
- Deriving the HJB equation.
- The verification theorem.
- The linear quadratic regulator.

Problem Formulation

$$\max_{u} \quad E\left[\int_0^T F(t, X_t, u_t)dt + \Phi(X_T)\right]$$

subject to

$$\begin{aligned} dX_t &= \mu\left(t, X_t, u_t\right)dt + \sigma\left(t, X_t, u_t\right)dW_t \\ X_0 &= x_0, \\ u_t &\in U(t, X_t), \quad \forall t. \end{aligned}$$

We will only consider **feedback control laws**, i.e. controls of the form

$$u_t = \mathbf{u}(t, X_t)$$

Terminology:

$$\begin{aligned} X &= \text{state variable} \\ u &= \text{control variable} \\ U &= \text{control constraint} \end{aligned}$$

Note: No state space constraints.

Main idea

- Embedd the problem above in a family of problems indexed by starting point in time and space.
- Tie all these problems together by a PDE–the Hamilton Jacobi Bellman equation.
- The control problem is reduced to the problem of solving the deterministic HJB equation.

NOTE:
For simplicity of notation we assume that X, W, and u are scalar.

Some notation

- For any fixed number $u \in R$, the functions μ^u and σ^u are defined by

$$
\begin{aligned}
\mu^u(t,x) &= \mu(t,x,u), \\
\sigma^u(t,x) &= \sigma(t,x,u),
\end{aligned}
$$

- For any control law $\mathbf{u}$, the functions $\mu^{\mathbf{u}}$, $\sigma^{\mathbf{u}}$, and $F^{\mathbf{u}}(t,x)$ are defined by

$$
\begin{aligned}
\mu^{\mathbf{u}}(t,x) &= \mu(t,x,\mathbf{u}(t,x)), \\
\sigma^{\mathbf{u}}(t,x) &= \sigma(t,x,\mathbf{u}(t,x)), \\
F^{\mathbf{u}}(t,x) &= F(t,x,\mathbf{u}(t,x)).
\end{aligned}
$$

More notation

- For any fixed number $u \in R$, the partial differential operator $\mathcal{A}^u$ is defined by

$$\mathcal{A}^u = \mathcal{A}^u = \mu^u(t,x)\frac{\partial}{\partial x} + \frac{1}{2}\left[\sigma^u(t,x)\right]^2\frac{\partial^2}{\partial x^2}.$$

- For any control law $\mathbf{u}$, the partial differential operator $\mathcal{A}^{\mathbf{u}}$ is defined by

$$\mathcal{A}^{\mathbf{u}} = \mu^{\mathbf{u}}(t,x)\frac{\partial}{\partial x} + \frac{1}{2}\left[\sigma^{\mathbf{u}}(t,x)\right]^2\frac{\partial^2}{\partial x^2}.$$

- For any control law $\mathbf{u}$, the process $X^{\mathbf{u}}$ is the solution of the SDE

$$dX_t^{\mathbf{u}} = \mu\left(t, X_t^{\mathbf{u}}, \mathbf{u}_t\right)dt + \sigma\left(t, X_t^{\mathbf{u}}, \mathbf{u}_t\right)dW_t,$$

where

$$\mathbf{u}_t = \mathbf{u}(t, X_t^{\mathbf{u}})$$

Multi dimensional notation

- For any fixed vector $u \in R^k$, the functions μ^u, σ^u and C^u are defined by

$$\begin{aligned}
\mu^u(t,x) &= \mu(t,x,u), \\
\sigma^u(t,x) &= \sigma(t,x,u), \\
C^u(t,x) &= \sigma(t,x,u)\sigma(t,x,u)'.
\end{aligned}$$

- For any control law $\mathbf{u}$, the functions $\mu^{\mathbf{u}}$, $\sigma^{\mathbf{u}}$, $C^{\mathbf{u}}(t,x)$ and $F^{\mathbf{u}}(t,x)$ are defined by

$$\begin{aligned}
\mu^{\mathbf{u}}(t,x) &= \mu(t,x,\mathbf{u}(t,x)), \\
\sigma^{\mathbf{u}}(t,x) &= \sigma(t,x,\mathbf{u}(t,x)), \\
C^{\mathbf{u}}(t,x) &= \sigma(t,x,\mathbf{u}(t,x))\sigma(t,x,\mathbf{u}(t,x))', \\
F^{\mathbf{u}}(t,x) &= F(t,x,\mathbf{u}(t,x)).
\end{aligned}$$

More multi dimensional notation

- For any fixed vector $u \in R^k$, the partial differential operator $\mathcal{A}^u$ is defined by

$$\mathcal{A}^u = \sum_{i=1}^{n} \mu_i^u(t,x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} C_{ij}^u(t,x) \frac{\partial^2}{\partial x_i \partial x_j}.$$

- For any control law $\mathbf{u}$, the partial differential operator $\mathcal{A}^{\mathbf{u}}$ is defined by

$$\mathcal{A}^{\mathbf{u}} = \sum_{i=1}^{n} \mu_i^{\mathbf{u}}(t,x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} C_{ij}^{\mathbf{u}}(t,x) \frac{\partial^2}{\partial x_i \partial x_j}.$$

- For any control law $\mathbf{u}$, the process $X^{\mathbf{u}}$ is the solution of the SDE

$$dX_t^{\mathbf{u}} = \mu\left(t, X_t^{\mathbf{u}}, \mathbf{u}_t\right) dt + \sigma\left(t, X_t^{\mathbf{u}}, \mathbf{u}_t\right) dW_t,$$

where

$$\mathbf{u}_t = \mathbf{u}(t, X_t^{\mathbf{u}})$$

Embedding the problem

For every fixed (t, x) the control problem $\mathcal{P}_{t,x}$ is defined as the problem to maximize

$$E_{t,x}\left[\int_t^T F(s, X_s^{\mathbf{u}}, u_s)ds + \Phi\left(X_T^{\mathbf{u}}\right)\right],$$

given the dynamics

$$\begin{aligned} dX_s^{\mathbf{u}} &= \mu\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right)ds + \sigma\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right)dW_s, \\ X_t &= x, \end{aligned}$$

and the constraints

$$\mathbf{u}(s, y) \in U, \quad \forall (s, y) \in [t, T] \times R.$$

The original problem was $\mathcal{P}_{0,x_0}$.

The optimal value function

- The **value function**

$$\mathcal{J} : R_+ \times R \times \mathcal{U} \to R$$

is defined by

$$\mathcal{J}(t,x,\mathbf{u}) = E\left[\int_t^T F(s, X_s^{\mathbf{u}}, \mathbf{u}_s)ds + \Phi\left(X_T^{\mathbf{u}}\right)\right]$$

given the dynamics above.

- The **optimal value function**

$$V : R_+ \times R \to R$$

is defined by

$$V(t,x) = \sup_{\mathbf{u}\in\mathcal{U}} \mathcal{J}(t,x,\mathbf{u}).$$

- We want to derive a PDE for V.

Assumptions

We assume:

- There exists an optimal control law $\hat{\mathbf{u}}$.
- The optimal value function V is regular in the sense that $V \in C^{1,2}$.
- A number of limiting procedures in the following arguments can be justified.

Bellman Optimality Principle

Theorem: If a control law $\hat{\mathbf{u}}$ is optimal for the time interval $[t, T]$ then it is also optimal for all smaller intervals $[s, T]$ where $s \geq t$.

Proof: Exercise. ∎

Basic strategy

To derive the PDE do as follows:

- Fix $(t, x) \in (0, T) \times R$.
- Choose a real number h (interpreted as a "small" time increment).
- Choose an arbitrary control law $\mathbf{u}$ on the time inerval $[t, t+h]$.

Now define the control law $\mathbf{u}^{\star}$ by

$$\mathbf{u}^{\star}(s, y) = \begin{cases} \mathbf{u}(s, y), & (s, y) \in [t, t+h] \times R \\ \hat{\mathbf{u}}(s, y), & (s, y) \in (t+h, T] \times R. \end{cases}$$

In other words, if we use $\mathbf{u}^{\star}$ then we use the arbitrary control $\mathbf{u}$ during the time interval $[t, t+h]$, and then we switch to the optimal control law during the rest of the time period.

Basic idea

The whole idea of DynP boils down to the following procedure.

- Given the point (t, x) above, we consider the following two strategies over the time interval $[t, T]$:

  I: Use the optimal law $\hat{\mathbf{u}}$.

  II: Use the control law $\mathbf{u}^{\star}$ defined above.

- Compute the expected utilities obtained by the respective strategies.

- Using the obvious fact that $\hat{\mathbf{u}}$ is least as good as $\mathbf{u}^{\star}$, and letting h tend to zero, we obtain our fundamental PDE.

Strategy values

Expected utility for $\hat{\mathbf{u}}$:

$$\mathcal{J}(t, x, \hat{\mathbf{u}}) = V(t, x)$$

Expected utility for $\mathbf{u}^{\star}$:

- The expected utility for $[t, t+h)$ is given by

$$E_{t,x}\left[\int_t^{t+h} F\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right) ds\right].$$

- Conditional expected utility over $[t+h, T]$, given (t, x):

$$E_{t,x}\left[V(t+h, X_{t+h}^{\mathbf{u}})\right].$$

- Total expected utility for Strategy II is

$$E_{t,x}\left[\int_t^{t+h} F\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right) ds + V(t+h, X_{t+h}^{\mathbf{u}})\right].$$

Comparing strategies

We have trivially

$$V(t,x) \geq E_{t,x}\left[\int_t^{t+h} F\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right) ds + V(t+h, X_{t+h}^{\mathbf{u}})\right].$$

Remark
We have equality above if and only if the control law $\mathbf{u}$ is the optimal law $\hat{\mathbf{u}}$.

Now use Itô to obtain

$$V(t+h, X_{t+h}^{\mathbf{u}}) = V(t,x)$$

$$+ \int_t^{t+h} \left\{ \frac{\partial V}{\partial t}(s, X_s^{\mathbf{u}}) + \mathcal{A}^{\mathbf{u}} V(s, X_s^{\mathbf{u}}) \right\} ds$$

$$+ \int_t^{t+h} V_x(s, X_s^{\mathbf{u}}) \sigma^{\mathbf{u}} dW_s,$$

and plug into the formula above.

We obtain

$$E_{t,x}\left[\int_t^{t+h}\left\{F\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right)+\frac{\partial V}{\partial t}(s, X_s^{\mathbf{u}})+\mathcal{A}^{\mathbf{u}}V(s, X_s^{\mathbf{u}})\right\}ds\right]\leq 0.$$

Going to the limit:
Divide by h, move h within the expectation and let h tend to zero.
We get

$$F(t, x, u)+\frac{\partial V}{\partial t}(t, x)+\mathcal{A}^{u}V(t, x)\leq 0,$$

Recall

$$F(t,x,u) + \frac{\partial V}{\partial t}(t,x) + \mathcal{A}^u V(t,x) \leq 0,$$

This holds for all $u = \mathbf{u}(t,x)$, with equality if and only if $\mathbf{u} = \hat{\mathbf{u}}$.

We thus obtain the **HJB equation**

$$\frac{\partial V}{\partial t}(t,x) + \sup_{u \in U} \left\{ F(t,x,u) + \mathcal{A}^u V(t,x) \right\} = 0.$$

The HJB equation

Theorem:
Under suitable regularity assumptions the follwing hold:

I: V satisfies the Hamilton–Jacobi–Bellman equation

$$\begin{aligned}\frac{\partial V}{\partial t}(t,x) + \sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u V(t,x)\} &= 0, \\ V(T,x) &= \Phi(x),\end{aligned}$$

II: For each $(t,x) \in [0,T] \times R$ the supremum in the HJB equation above is attained by $u = \hat{\mathbf{u}}(t,x)$, i.e. by the optimal control.

Note: We have only treated the scalar case, but the extension to the multidimensional case is obvious.

Logic and problem

Note: We have shown that **if** V is the optimal value function, and **if** V is regular enough, **then** V satisfies the HJB equation. The HJB eqn is thus derived as a **necessary** condition, and requires strong *ad hoc* regularity assumptions, alternatively the use of viscosity solutions techniques.

Problem: Suppose we have solved the HJB equation. Have we then found the optimal value function and the optimal control law? In other words, is HJB a **sufficient** condition for optimality.

Answer: Yes! This follows from the **Verification Theorem**.

The Verification Theorem

Suppose that we have two functions $H(t,x)$ and $g(t,x)$, such that

- H is sufficiently integrable, and solves the HJB equation

$$\begin{cases} \frac{\partial H}{\partial t}(t,x) + \sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u H(t,x)\} & = & 0, \\ H(T,x) & = & \Phi(x), \end{cases}$$

- For each fixed (t,x), the supremum in the expression

$$\sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u H(t,x)\}$$

  is attained by the choice $u = g(t,x)$.

Then the following hold.

1. The optimal value function V to the control problem is given by
$$V(t,x) = H(t,x).$$
2. There exists an optimal control law $\hat{\mathbf{u}}$, and in fact
$$\hat{\mathbf{u}}(t,x) = g(t,x)$$

.

Handling the HJB equation

1. Consider the HJB equation for V.
2. Fix $(t, x) \in [0, T] \times R^n$ and solve, the static optimization problem
$$\max_{u \in U} \quad [F(t, x, u) + \mathcal{A}^u V(t, x)] \,.$$
Here u is the only variable, whereas t and x are fixed parameters. The functions F, μ, σ and V are considered as given.
3. The optimal $\hat{u}$, will depend on t and x, and on the function V and its partial derivatives. We thus write $\hat{u}$ as
$$\hat{\mathbf{u}} = \hat{\mathbf{u}}\,(t, x; V)\,. \tag{5}$$
4. The function $\hat{\mathbf{u}}\,(t, x; V)$ is our candidate for the optimal control law, but since we do not know V this description is incomplete. Therefore we substitute the expression for $\hat{u}$ into the PDE , giving us the highly nonlinear (why?) PDE
$$\begin{aligned} \frac{\partial V}{\partial t}(t, x) + F^{\hat{\mathbf{u}}}(t, x) + \mathcal{A}^{\hat{\mathbf{u}}}\,(t, x)\, V(t, x) &= 0, \\ V(T, x) &= \Phi(x). \end{aligned}$$
5. Now we solve the PDE above! Then we put the solution V into expression (**??**). Using the verification theorem we can identify V as the optimal value function, and $\hat{u}$ as the optimal control law.

Making an Ansatz

- The hard work of dynamic programming consists in solving the highly nonlinear HJB equation

- There are no general analytic methods available for this, so the number of known optimal control problems with an analytic solution is very small indeed.

- In an actual case one usually tries to **guess** a solution, i.e. we typically make a parameterized **Ansatz** for V then use the PDE in order to identify the parameters.

- **Hint:** V often inherits some structural properties from the boundary function Φ as well as from the instantaneous utility function F.

- Most of the known solved control problems have, to some extent, been "rigged" in order to be analytically solvable.

The Linear Quadratic Regulator

$$\min_{u \in R} \quad E\left[\int_0^T \left\{QX_t^2 + Ru_t^2\right\} dt + HX_T^2\right],$$

with dynamics

$$dX_t = \{AX_t + Bu_t\} dt + CdW_t.$$

We want to control a vehicle in such a way that it stays close to the origin (the terms Qx^2 and Hx^2) while at the same time keeping the "energy" Ru^2 small.

Here $X_t \in R$ and $\mathbf{u}_t \in R$, and we impose no control constraints on u.

The real numbers Q, R, H, A, B and C are assumed to be known. We assume that R is strictly positive.

Handling the Problem

The HJB equation becomes

$$\begin{cases} \dfrac{\partial V}{\partial t}(t,x) & + \quad \inf_{u \in R} \left\{ Qx^2 + Ru^2 + V_x(t,x)\left[Ax + Bu\right] \right\} \\ & + \quad \frac{1}{2}\frac{\partial^2 V}{\partial x^2}(t,x) C^2 = 0, \\ V(T,x) & = \quad Hx^2. \end{cases}$$

For each fixed choice of (t, x) we now have to solve the static unconstrained optimization problem to minimize

$$Qx^2 + Ru^2 + V_x(t,x)\left[Ax + Bu\right].$$

The problem was:

$$\min_{u} \quad Qx^2 + Ru^2 + V_x(t,x)\left[Ax + Bu\right].$$

Since $R > 0$ we set the u-derivative to zero and obtain

$$2Ru = -V_x B,$$

which gives us the optimal u as

$$\hat{u} = -\frac{1}{2}\frac{B}{R}V_x.$$

Note: This is our candidate of optimal control law, but it depends on the unknown function V.

We now make an educated guess about the structure of V.

From the boundary function Hx^2 and the term Qx^2 in the cost function we make the Ansatz

$$V(t,x) = P(t)x^2 + q(t),$$

where $P(t)$ and $q(t)$ are deterministic functions.

With this trial solution we have,

$$\begin{aligned} \frac{\partial V}{\partial t}(t,x) &= \dot{P}x^2 + \dot{q}, \\ V_x(t,x) &= 2Px, \\ V_{xx}(t,x) &= 2P \\ \hat{u} &= -\frac{B}{R}Px. \end{aligned}$$

Inserting these expressions into the HJB equation we get

$$x^2 \left\{ \dot{P} + Q - \frac{B^2}{R}P^2 + 2AP \right\}$$
$$+\dot{q}PC^2 = 0.$$

We thus get the following ODE for P

$$\begin{cases} \dot{P} &= \frac{B^2}{R}P^2 - 2AP - Q, \\ P(T) &= H. \end{cases}$$

and we can integrate directly for q:

$$\begin{cases} \dot{q} &= -C^2 P, \\ q(T) &= 0. \end{cases}$$

The is ODE for P is a **Riccati equation**. The equation for q can then be integrated directly.

Final Result for LQ:

$$\begin{aligned} V(t,x) &= P(t)x^2 + \int_t^T C^2 P(s)ds, \\ \hat{\mathbf{u}}(t,x) &= -\frac{B}{R}P(t)x. \end{aligned}$$

2. Investment Theory

- Problem formulation.
- An extension of HJB.
- The simplest consumption-investment problem.
- The Merton fund separation results.

Recap of Basic Facts

We consider a market with n assets.

$$
\begin{aligned}
S_t^i &= \text{price of asset No } i, \\
h_t^i &= \text{units of asset No } i \text{ in portfolio} \\
w_t^i &= \text{portfolio weight on asset No } i \\
X_t &= \text{portfolio value} \\
c_t &= \text{consumption rate}
\end{aligned}
$$

We have the relations

$$X_t = \sum_{i=1}^{n} h_t^i S_t^i, \quad w_t^i = \frac{h_t^i S_t^i}{X_t}, \quad \sum_{i=1}^{n} w_t^i = 1.$$

Basic equation:
Dynamics of self financing portfolio in terms of relative weights

$$dX_t = X_t \sum_{i=1}^{n} w_t^i \frac{dS_t^i}{S_t^i} - c_t dt$$

Simplest model

Assume a scalar risky asset and a constant short rate.

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t \\ dB_t &= r B_t dt \end{aligned}$$

We want to maximize expected utility over time

$$\max_{w^0, w^1, c} \quad E\left[\int_0^T F(t, c_t) dt\right]$$

Dynamics

$$dX_t = X_t \left[w_t^0 r + w_t^1 \alpha \right] dt - c_t dt + w_t^1 \sigma X_t dW_t,$$

Constraints

$$\begin{aligned} c_t &\geq 0, \ \forall t \geq 0, \\ w_t^0 + w_t^1 &= 1, \ \forall t \geq 0. \end{aligned}$$

Nonsense!

What are the problems?

- We can obtain unlimited utility by simply consuming arbitrary large amounts.
- The wealth will go negative, but there is nothing in the problem formulations which prohibits this.
- We would like to impose a constratin of type $X_t \geq 0$ but this is a **state constraint** and DynP does not allow this.

Good News:
DynP can be generalized to handle (some) problems of this kind.

Generalized problem

Let D be a nice open subset of $[0,T] \times R^n$ and consider the following problem.

$$\max_{u \in U} \quad E\left[\int_0^\tau F(s, X_s^{\mathbf{u}}, \mathbf{u}_s)ds + \Phi\left(\tau, X_\tau^{\mathbf{u}}\right)\right].$$

Dynamics:

$$\begin{aligned} dX_t &= \mu\left(t, X_t, u_t\right) dt + \sigma\left(t, X_t, u_t\right) dW_t, \\ X_0 &= x_0, \end{aligned}$$

The **stopping time** τ is defined by

$$\tau = \inf\left\{t \geq 0 \ |(t, X_t) \in \partial D\right\} \wedge T.$$

Generalized HJB

Theorem: Given enough regularity the follwing hold.

1. The optimal value function satisfies

$$\begin{cases} \frac{\partial V}{\partial t}(t,x) + \sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u V(t,x)\} & = & 0, & \forall (t,x) \in D \\ V(t,x) & = & \Phi(t,x), & \forall (t,x) \in \partial D. \end{cases}$$

2. We have an obvious verification theorem.

Reformulated problem

$$\max_{c \geq 0,\ w \in R} \quad E\left[\int_0^{\tau} F(t, c_t)dt\right]$$

The "ruin time" τ is defined by

$$\tau = \inf\left\{t \geq 0 \ |X_t = 0\right\} \wedge T.$$

Notation:

$$\begin{aligned} w^1 &= w, \\ w^0 &= 1 - w \end{aligned}$$

Thus no constraint on w.

Dynamics

$$dX_t = w_t \left[\alpha - r\right] X_t dt + \left(rX_t - c_t\right) dt + w\sigma X_t dW_t,$$

HJB Equation

$$\begin{aligned}
\frac{\partial V}{\partial t} + \sup_{c \geq 0, w \in R} \left\{ F(t,c) + wx(\alpha - r)\frac{\partial V}{\partial x} + (rx - c)\frac{\partial V}{\partial x} + \frac{1}{2}x^2 w^2 \sigma^2 \frac{\partial^2 V}{\partial x^2} \right\} &= 0, \\
V(T,x) &= 0, \\
V(t,0) &= 0.
\end{aligned}$$

We now specialize (why?) to

$$F(t,c) = e^{-\delta t} c^{\gamma},$$

and for simplicity we assume that

$$\Phi = 0,$$

so we have to maximize

$$e^{-\delta t} c^{\gamma} + wx(\alpha - r)\frac{\partial V}{\partial x} + (rx - c)\frac{\partial V}{\partial x} + \frac{1}{2}x^2 w^2 \sigma^2 \frac{\partial^2 V}{\partial x^2},$$

Analysis of the HJB Equation

In the embedded static problem we maximize, over c and w,

$$e^{-\delta t}c^{\gamma} + wx(\alpha - r)V_x + (rx - c)V_x + \frac{1}{2}x^2w^2\sigma^2V_{xx},$$

First order conditions:

$$\begin{aligned} \gamma c^{\gamma-1} &= e^{\delta t}V_x, \\ w &= \frac{-V_x}{x \cdot V_{xx}} \cdot \frac{\alpha - r}{\sigma^2}, \end{aligned}$$

Ansatz:

$$V(t,x) = e^{-\delta t}h(t)x^{\gamma},$$

Because of the boundary conditions, we must demand that

$$h(T) = 0. \tag{6}$$

Given a V of this form we have (using $\cdot$ to denote the time derivative)

$$\begin{aligned} V_t &= e^{-\delta t}\dot{h}x^{\gamma} - \delta e^{-\delta t} h x^{\gamma}, \\ V_x &= \gamma e^{-\delta t} h x^{\gamma - 1}, \\ V_{xx} &= \gamma(\gamma - 1) e^{-\delta t} h x^{\gamma - 2}. \end{aligned}$$

giving us

$$\begin{aligned} \widehat{w}(t,x) &= \frac{\alpha - r}{\sigma^2(1-\gamma)}, \\ \widehat{c}(t,x) &= x h(t)^{-1/(1-\gamma)}. \end{aligned}$$

Plug all this into HJB!

After rearrangements we obtain

$$x^{\gamma}\left\{\dot{h}(t)+Ah(t)+Bh(t)^{-\gamma/(1-\gamma)}\right\}=0,$$

where the constants A and B are given by

$$\begin{aligned} A &= \frac{\gamma(\alpha-r)^2}{\sigma^2(1-\gamma)}+r\gamma-\frac{1}{2}\frac{\gamma(\alpha-r)^2}{\sigma^2(1-\gamma)}-\delta \\ B &= 1-\gamma. \end{aligned}$$

If this equation is to hold for all x and all t, then we see that h must solve the ODE

$$\begin{aligned} \dot{h}(t)+Ah(t)+Bh(t)^{-\gamma/(1-\gamma)} &= 0, \\ h(T) &= 0. \end{aligned}$$

An equation of this kind is known as a **Bernoulli equation**, and it can be solved explicitly.

We are done.

Merton's Mutal Fund Theorems

We consider n risky assets with dynamics

$$dS_i = S_i\alpha_i dt + S_i\sigma_i dW, \quad i = 1, \ldots, n$$

where W is Wiener in R^n. On vector form:

$$dS = D(S)\alpha dt + D(S)\sigma dW.$$

where

$$\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \sigma = \begin{bmatrix} - \sigma_1 - \\ \vdots \\ - \sigma_n - \end{bmatrix}$$

$D(S)$ is the diagonal matrix

$$D(S) = diag[S_1, \ldots, S_n].$$

Assumption: We assume that the volatility matrix σ is invertible. This assumption is **very** important.

The case with a risk free asset

We consider the standard model above, i.e.

$$dS = D(S)\alpha dt + D(S)\sigma dW(t),$$

We also assume the risk free asset B with dynamics

$$dB = rBdt.$$

We denote $B = S_0$ and consider portfolio weights $(w_0, w_1, \ldots, w_n)'$ where $\sum_0^n w_i = 1$. We then eliminate w_0 by the relation

$$w_0 = 1 - \sum_1^n w_i,$$

and use the letter w to denote the portfolio weight vector for the risky assets only. Thus we use the notation

$$w = (w_1, \ldots, w_n)',$$

Note: $w \in R^n$ without constraints.

HJB

We obtain

$$dX = X \cdot w'(\alpha - re)dt + (rX - c)dt + X \cdot w'\sigma dW,$$

where $e = (1, 1, \ldots, 1)'$.

The HJB equation now becomes

$$\begin{cases} V_t(t,x) + \sup_{c \geq 0, w \in R^n} \{F(t,c) + \mathcal{A}^{c,w}V(t,x)\} &= 0, \\ V(T,x) &= 0, \\ V(t,0) &= 0, \end{cases}$$

where

$$\begin{aligned} \mathcal{A}^c V &= xw'(\alpha - re)V_x(t,x) + (rx - c)V_x(t,x) \\ &+ \frac{1}{2}x^2 w'\Sigma w V_{xx}(t,x). \end{aligned}$$

First order conditions

We maximize

$$F(t,c) + xw'(\alpha - re)V_x + (rx - c)V_x + \frac{1}{2}x^2 w' \Sigma w V_{xx}$$

with $c \geq 0$ and $w \in R^n$.

The first order conditions are

$$\begin{aligned} F_c &= V_x, \\ \hat{w} &= -\frac{V_x}{xV_{xx}} \Sigma^{-1}(\alpha - re), \end{aligned}$$

This has a simple geometrically and economic interpretation.

Interpretation

We had

$$\hat{w}_t = -\frac{V_x(t, X_t)}{X_t V_{xx}(t, X_t)} \Sigma^{-1}(\alpha - re),$$

so we can write

$$\hat{w}_t = Y_t \cdot h$$

where the scalar process Y is given by

$$Y_t = -\frac{V_x(t, X_t)}{X_t V_{xx}(t, X_t)}$$

and the fixed vector h is given by

$$h = \Sigma^{-1}(\alpha - re).$$

This means that the portfolio weight vector ω for the risky assets are always a(random) multiple of the fixed vector h.

Mutual Fund Separation Theorem

1. The optimal portfolio consists of an allocation between two fixed mutual funds w^0 and w^f.

2. The fund w^0 consists only of the risk free asset.

3. The fund w^f consists only of the risky assets, and is determined by the vector

$$h = \Sigma^{-1}(\alpha - re).$$

Formal problem

$$\max_{c,w} \quad E\left[\int_0^{\tau} F(t, c_t)dt\right]$$

given the dynamics

$$dX = Xw'\alpha dt - cdt + Xw'\sigma dW.$$

and constraints

$$e'w = 1, \quad c \geq 0.$$

Assumptions:

- The vector α and the matrix σ are constant and deterministic.
- The volatility matrix σ has full rank so $\sigma\sigma'$ is positive definite and invertible.

Note: S does not turn up in the X-dynamics so V is of the form

$$V(t, x, s) = V(t, x)$$

The HJB equation is

$$\begin{cases} V_t(t,x) + \sup_{e'w=1,\ c\geq 0} \{F(t,c) + \mathcal{A}^{c,w}V(t,x)\} & = & 0, \\ V(T,x) & = & 0, \\ V(t,0) & = & 0. \end{cases}$$

where

$$\mathcal{A}^{c,w}V = xw'\alpha V_x - cV_x + \frac{1}{2}x^2 w'\Sigma w\ V_{xx},$$

The matrix Σ is given by

$$\Sigma = \sigma\sigma'.$$

The HJB equation is

$$\begin{cases} V_t + \sup_{w'e=1,\ c\geq 0} \left\{ F(t,c) + (xw'\alpha - c)V_x + \frac{1}{2}x^2 w'\Sigma w V_{xx} \right\} & = & 0, \\ V(T,x) & = & 0, \\ V(t,0) & = & 0. \end{cases}$$

where $\Sigma = \sigma\sigma'$.

If we relax the constraint $w'e = 1$, the Lagrange function for the static optimization problem is given by

$$L = F(t,c) + (xw'\alpha - c)V_x + \frac{1}{2}x^2 w'\Sigma w V_{xx} + \lambda\left(1 - w'e\right).$$

$$
\begin{aligned}
L &= F(t,c) + (xw'\alpha - c)V_x \\
&+ \frac{1}{2}x^2 w'\Sigma w V_{xx} + \lambda\left(1 - w'e\right).
\end{aligned}
$$

The first order condition for c is

$$F_c = V_x.$$

The first order condition for w is

$$x\alpha' V_x + x^2 V_{xx} w'\Sigma = \lambda e',$$

so we can solve for w in order to obtain

$$\hat{w} = \Sigma^{-1}\left[\frac{\lambda}{x^2 V_{xx}}e - \frac{xV_x}{x^2 V_{xx}}\alpha\right].$$

Using the relation $e'w = 1$ this gives λ as

$$\lambda = \frac{x^2 V_{xx} + xV_x e'\Sigma^{-1}\alpha}{e'\Sigma^{-1}e},$$

Inserting λ gives us, after some manipulation,

$$\hat{w} = \frac{1}{e'\Sigma^{-1}e}\Sigma^{-1}e + \frac{V_x}{xV_{xx}}\Sigma^{-1}\left[\frac{e'\Sigma^{-1}\alpha}{e'\Sigma^{-1}e}e - \alpha\right].$$

We can write this as

$$\hat{w}(t) = g + Y(t)h,$$

where the fixed vectors g and h are given by

$$\begin{aligned} g &= \frac{1}{e'\Sigma^{-1}e}\Sigma^{-1}e, \\ h &= \Sigma^{-1}\left[\frac{e'\Sigma^{-1}\alpha}{e'\Sigma^{-1}e}e - \alpha\right], \end{aligned}$$

whereas Y is given by

$$Y(t) = \frac{V_x(t, X(t))}{X(t)V_{xx}(t, X(t))}.$$

We had

$$\hat{w}(t) = g + Y(t)h,$$

Thus we see that the optimal portfolio is moving stochastically along the one-dimensional "optimal portfolio line"

$$g + sh,$$

in the $(n-1)$-dimensional "portfolio hyperplane" Δ, where

$$\Delta = \{w \in R^n \, | e'w = 1\}.$$

If we fix two points on the optimal portfolio line, say $w^a = g + ah$ and $w^b = g + bh$, then any point w on the line can be written as an affine combination of the basis points w^a and w^b. An easy calculation shows that if $w^s = g + sh$ then we can write

$$w^s = \mu w^a + (1-\mu) w^b,$$

where

$$\mu = \frac{s-b}{a-b}.$$

Mutual Fund Theorem

There exists a family of mutual funds, given by $w^s = g + sh$, such that

1. For each fixed s the portfolio w^s stays fixed over time.

2. For fixed a, b with $a \neq b$ the optimal portfolio $\hat{\mathbf{w}}(t)$ is, obtained by allocating all resources between the fixed funds w^a and w^b, i.e.

$$\hat{w}(t) = \mu^a(t)w^a + \mu^b(t)w^b,$$

Chapter 20

The Martingale Approach to Optimal Investment Theory

Tomas Björk

Contents

- Decoupling the wealth profile from the portfolio choice.
- Lagrange relaxation.
- Solving the general wealth problem.
- Example: Log utility.
- Example: The numeraire portfolio.

Problem Formulation

Standard model with internal filtration

$$\begin{aligned} dS_t &= D(S_t)\alpha_t dt + D(S_t)\sigma_t dW_t, \\ dB_t &= rB_t dt. \end{aligned}$$

Assumptions:

- Drift and diffusion terms are allowed to be arbitrary adapted processes.
- The market is **complete**.
- We have a given initial wealth x_0

Problem:

$$\max_{h \in \mathcal{H}} \quad E^P \left[\Phi(X_T) \right]$$

where

$$\mathcal{H} = \{\text{self financing portfolios}\}$$

given the initial wealth $X_0 = x_0$.

Some observations

- In a complete market, there is a unique martingale measure Q.

- Every claim Z satisfying the budget constraint

$$e^{-rT}E^{Q}[Z] = x_0,$$

  is attainable by an $h \in \mathcal{H}$ and vice versa.

- We can thus write our problem as

$$\max_{Z} \quad E^{P}[\Phi(Z)]$$

  subject to the constraint

$$e^{-rT}E^{Q}[Z] = x_0.$$

- We can forget the wealth dynamics!

Basic Ideas

Our problem was

$$\max_{Z} \quad E^P[\Phi(Z)]$$

subject to

$$e^{-rT}E^Q[Z] = x_0.$$

Idea I:

We can **decouple** the optimal portfolio problem into:

1. Finding the optimal wealth profile $\hat{Z}$.
2. Given $\hat{Z}$, find the replicating portfolio.

Idea II:

- Rewrite the constraint under the measure P.
- Use Lagrangian techniques to relax the constraint.

Lagrange formulation

Problem:

$$\max_{Z} \quad E^P \left[\Phi(Z) \right]$$

subject to

$$e^{-rT} E^P \left[L_T Z \right] = x_0.$$

Here L is the likelihood process, i.e.

$$L_t = \frac{dQ}{dP}, \quad \text{on } \mathcal{F}_t, \quad 0 \le t \le T$$

The Lagrangian of the problem is

$$\mathcal{L} = E^P \left[\Phi(Z) \right] + \lambda \left\{ x_0 - e^{-rT} E^P \left[L_T Z \right] \right\}$$

i.e.

$$\mathcal{L} = E^P \left[\Phi(Z) - \lambda e^{-rT} L_T Z \right] + \lambda x_0$$

The optimal wealth profile

Given enough convexity and regularity we now expect, given the dual variable λ, to find the optimal Z by maximizing

$$\mathcal{L} = E^P \left[\Phi(Z) - \lambda e^{-rT} L_T Z \right] + \lambda x_0$$

over unconstrained Z, i.e. to maximize

$$\int_\Omega \left\{ \Phi(Z(\omega)) - \lambda e^{-rT} L_T(\omega) Z(\omega) \right\} dP(\omega)$$

This is a trivial problem!

We can simply maximize $Z(\omega)$ for each ω separately.

$$\max_z \quad \left\{ \Phi(z) - \lambda e^{-rT} L_T z \right\}$$

The optimal wealth profile

Our problem:

$$\max_{z} \quad \left\{ \Phi(z) - \lambda e^{-rT} L_T z \right\}$$

First order condition

$$\Phi'(z) = \lambda e^{-rT} L_T$$

The optimal Z is thus given by

$$\hat{Z} = G\left(\lambda e^{-rT} L_T\right)$$

where

$$G(y) = \left[\Phi'\right]^{-1}(y).$$

The dual varaiable λ is determined by the constraint

$$e^{-rT} E^P \left[L_T \hat{Z} \right] = x_0.$$

Example – log utility

Assume that

$$\Phi(x) = \ln(x)$$

Then

$$g(y) = \frac{1}{y}$$

Thus

$$\hat{Z} = G\left(\lambda e^{-rT} L_T\right) = \frac{1}{\lambda} e^{rT} L_T^{-1}$$

Finally λ is determined by

$$e^{-rT} E^P \left[L_T \hat{Z} \right] = x_0.$$

i.e.

$$e^{-rT} E^P \left[L_T \frac{1}{\lambda} e^{rT} L_T^{-1} \right] = x_0.$$

so $\lambda = x_0^{-1}$ and

$$\hat{Z} = x_0 e^{rT} L_T^{-1}$$

The optimal wealth process

- We have computed the optimal **terminal** wealth profile
$$\widehat{Z} = \widehat{X}_T = x_0 e^{rT} L_T^{-1}$$

- What does the optimal wealth **process** $\widehat{X}_t$ look like?

We have (why?)

$$\widehat{X}_t = e^{-r(T-t)} E^Q \left[\widehat{X}_T \middle| \mathcal{F}_t \right]$$

so we obtain

$$\widehat{X}_t = x_0 e^{rt} E^Q \left[L_T^{-1} \middle| \mathcal{F}_t \right]$$

But L^{-1} is a Q-martingale (why?) so we obtain

$$\widehat{X}_t = x_0 e^{rt} L_t^{-1}.$$

The Optimal Portfolio

- We have computed the optimal wealth process.
- How do we compute the optimal portfolio?

Assume for simplicity that we have a standard Black-Scholes model

$$
\begin{aligned}
dS_t &= \mu S_t dt + \sigma S_t dW_t, \\
dB_t &= r B_t dt
\end{aligned}
$$

Recall that

$$\widehat{X}_t = x_0 e^{rt} L_t^{-1}.$$

Basic Program

1. Use Ito and the formula for $\widehat{X}_t$ to compute $d\widehat{X}_t$ like

$$d\widehat{X}_t = \widehat{X}_t(\quad)dt + \widehat{X}_t\beta_t dW_t$$

where we do not care about ( ).

2. Recall that

$$d\widehat{X}_t = \widehat{X}_t\left\{(1-\hat{u}_t)\frac{dB_t}{B_t} + \hat{u}_t\frac{dS_t}{S_t}\right\}$$

which we write as

$$d\widehat{X}_t = \widehat{X}_t\{\quad\}dt + \widehat{X}_t\hat{u}_t\sigma dW_t$$

3. We can identify $\hat{u}$ as

$$\hat{u}_t = \frac{\beta_t}{\sigma}$$

We recall

$$\widehat{X}_t = x_0 e^{rt} L_t^{-1}.$$

We also recall that

$$dL_t = L_t \varphi dW_t,$$

where

$$\varphi = \frac{r - \mu}{\sigma}$$

From this we have

$$dL_t^{-1} = \varphi^2 L_t^{-1} dt - L_t^{-1} \varphi dW_t$$

and we obtain

$$\widehat{X}_t = \widehat{X}_t \left\{ \quad \right\} dt - \widehat{X}_t \varphi dW_t$$

Result: The optimal portfolio is given by

$$\hat{u}_t = \frac{\mu - r}{\sigma^2}$$

Note that $\hat{u}$ is a "myopic" portfolio in the sense that it does not depend on the time horizon T.

A Digression: The Numeraire Portfolio

Standard approach:

- Choose a fixed numeraire (portfolio) N.
- Find the corresponding martingale measure, i.e. find Q^N s.t.

$$\frac{B}{N}, \quad \text{and} \quad \frac{S}{N}$$

  are Q^N -martingales.

Alternative approach:

- Choose a fixed measure $Q \sim P$.
- Find numeraire N such that $Q = Q^N$.

Special case:

- Set $Q = P$
- Find numeraire N such that $Q^N = P$ i.e. such that

$$\frac{B}{N}, \quad \text{and} \quad \frac{S}{N}$$

  are Q^N-martingales under the **objective** measure P.
- This N is called the **numeraire portfolio**.

Log utility and the numeraire portfolio

Definition:
The **growth optimal portfolio** (GOP) is the portfolio which is optimal for log utility (for arbitrary terminal date T.

Theorem:
Assume that X is GOP. Then X is the numeraire portfolio.

Proof:
We have to show that the process

$$Y_t = \frac{S_t}{X_t}$$

is a P martingale.

We have

$$\frac{S_t}{X_t} = x_0^{-1} e^{-rt} S_t L_t$$

which is a P martingale, since $x_0^{-1} e^{-rt} S_t$ is a Q martingale.