Continuous Time Finance

based on the

original 2017 slides by Tomas Björk

slightly reworked, extended and stylized by
Peter Spreij

2026

Contents of the course

Textbook: Tomas Björk: *Arbitrage Theory in Continuous Time*, (3rd ed.), Oxford University Press, 2009. Still available at amazon.com.

- Stochastic Calculus (Chapters 4–5).
- Black-Scholes (Chapters 6–7).
- Completeness and hedging (Chapters 8–9).
- The martingale approach (Chapters 10–12).
- Incomplete markets (Chapter 15).
- Dividends (Chapter 16).
- Currency derivatives (Chapter 17).
- Stochastic Control Theory (Chapter 19).
- Martingale Methods for Optimal Investment (Chapters 20).

Lecture 0

a very concise, intuitive,
introduction to Stochastic Calculus

This lecture is not given during the course; it serves as background material that students are supposed to master.

Stochastic Integrals

(Chapters 4–5)

Typical Setup

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

S_t = price, at time t, per unit of underlying asset

Consider a fixed **financial derivative**, a financial product specified in terms of an underlying asset, e.g. a European call option.

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

What is needed

1. Mathematical model for the underlying price process.
   Example: The Black-Scholes model.

2. Mathematical techniques to handle the price dynamics. Main tool: 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, not the whole history. All past information is already incorporated into today's stock prices. (Market efficiency).

Note the following concepts, and their differences.

Stochastic variable, choosing a **number** at random

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

Notation

Here is some basic notation.

$$X_t = \text{any random process},$$
$$\mathrm{d}t = \text{(infinitesimal) small time step},$$
$$\mathrm{d}X_t = X_{t+\mathrm{d}t} - X_t$$

- We sometimes write $X(t)$ instead of X_t, when this is notationally more convenient.
- $\mathrm{d}X_t$ is called the **increment** of X over the interval $[t, t + \mathrm{d}t]$.
- For any fixed interval $[t, t + \mathrm{d}t]$, the increment $\mathrm{d}X_t$ is considered a stochastic variable.
- If the increments $\mathrm{d}X_s$ and $\mathrm{d}X_t$, over the disjoint intervals $[s, s + \mathrm{d}s]$ and $[t, t + \mathrm{d}t]$ are independent random variables, 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** (named after the famous mathematician Norbert Wiener) if it has the following properties

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

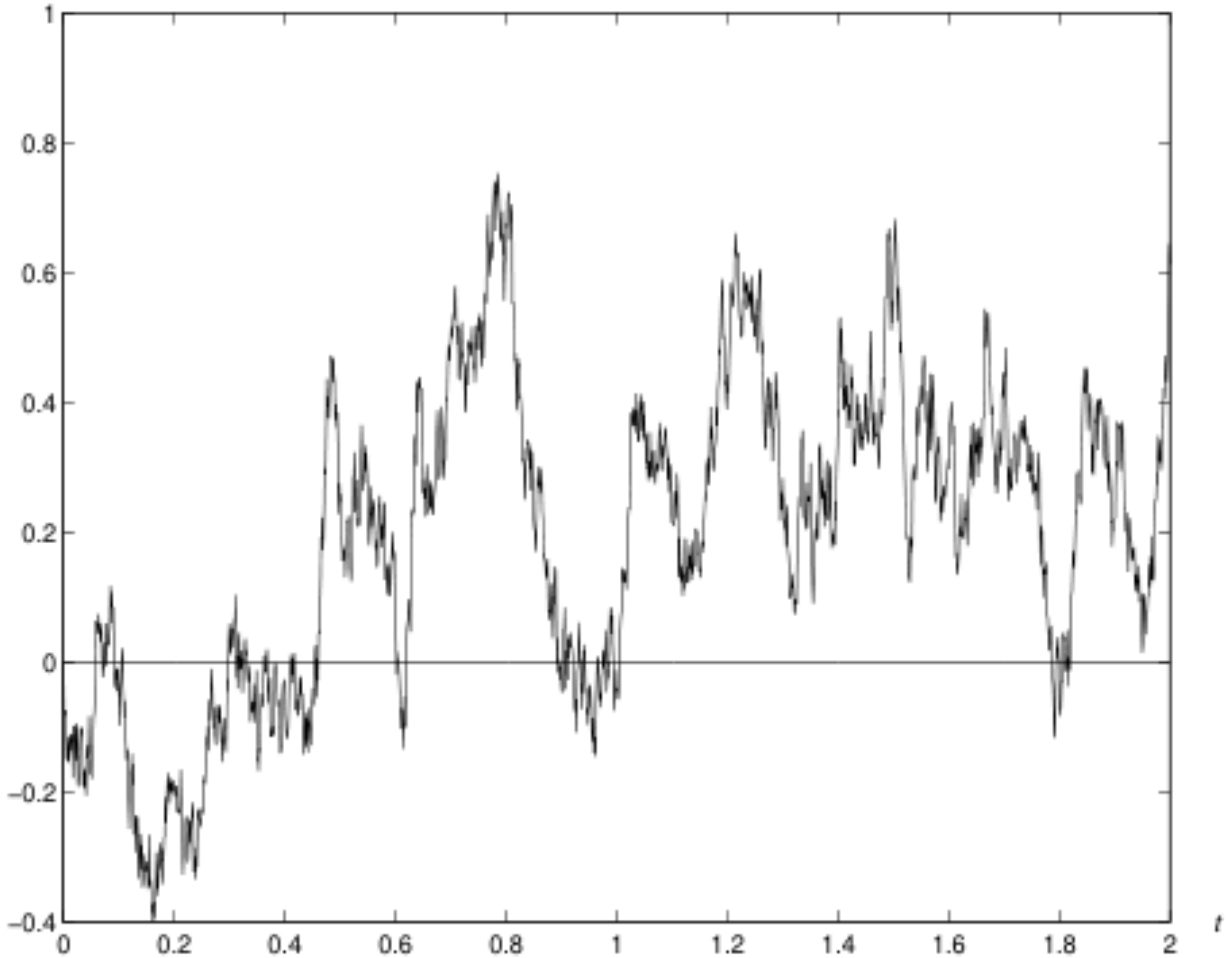
$$\begin{aligned} W_t - W_s &\sim N[0, t-s] \\ \mathbb{E}[W_t - W_s] &= 0, \\ \mathrm{Var}\,[W_t - W_s] &= t - s \end{aligned}$$

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

The Wiener process is a model for a continuous random walk.

Note: In Hull (another popular book), a Wiener process is typically denoted by Z instead of W.

A Wiener Trajectory, a typical example



Important Fact

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

Proof. Hard, not given.

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

$$\mathrm{d}X_t = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t,$$

where μ and σ are constants.

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

$$\begin{aligned} X_t - X_0 &= \mu \cdot t + \sigma \cdot (W_t - W_0), \\ X_t &= X_0 + \mu t + \sigma W_t. \end{aligned}$$

It follows that

$$X_t \sim N[(X_0 + \mu t, \sigma^2 t).$$

Itô processes

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

$$\mathrm{d}X_t = \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t,$$

where μ_t and σ_t are random processes.

Informally you can think of $\mathrm{d}W_t$ as a random variable of the form

$$\mathrm{d}W_t \sim N(0, \mathrm{d}t).$$

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

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

Example. The Black-Scholes model

Price dynamics (Geometric Brownian Motion),

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t.$$

Terminology. An equation with a $\mathrm{d}t$ term on the right as well as a $\mathrm{d}W_t$ term is called a stochastic differential equation (SDE).

Simple analysis. Assume that $\sigma = 0$. Then $\mathrm{d}S_t = \mu S_t \,\mathrm{d}t$. Divide, assume this make sense, by $\mathrm{d}t$!

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = \mu S_t,\ S_0 = s_0 \text{ (a given value)}.$$

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. Is this true?

Intuitive Economic Interpretation

Rewrite,

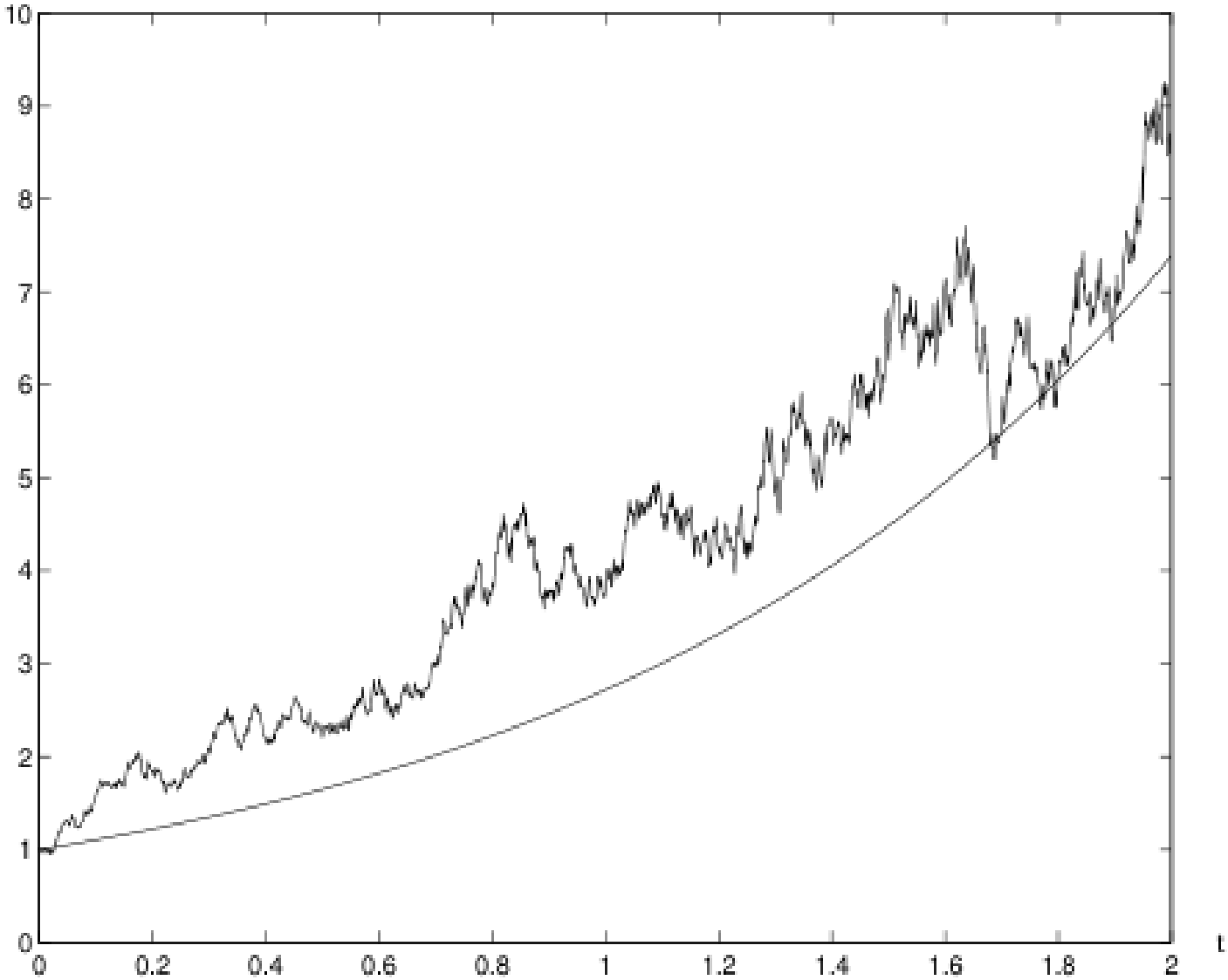
$$\frac{\mathrm{d}S_t}{S_t} = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t.$$

Over a small time interval $[t, t+\mathrm{d}t]$ this means:

$$\text{Return of } S = \text{mean return}(\mu) + \sigma \times (\text{Gaussian random disturbance}).$$

- The asset **return** is a random walk (with drift).
- μ = mean rate of return per unit time.
- σ = volatility. Large σ = large random fluctuations, small σ = small random fluctuations.
- The returns are normal, and hence,
- the stock price is lognormal.

A GBM Trajectory



Stochastic Differentials and Integrals, first steps

Consider an expression of the form

$$\begin{aligned} \mathrm{d}X_t &= \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t, \\ X_0 &= x_0. \end{aligned}$$

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

Answer. Write the equation on integrated form as

$$X_t = x_0 + \int_0^t \mu_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s.$$

But then, how is this interpreted?

Stochastic Integrals

Recall

$$X_t = x_0 + \int_0^t \mu_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s.$$

We distinguish two terms.

- The first one is

  $$\int_0^t \mu_s \, \mathrm{d}s.$$

  This is a standard Riemann integral for each μ-trajectory.
- The second one is

  $$\int_0^t \sigma_s \, \mathrm{d}W_s,$$

  a **Stochastic integral**, also called an **Itô integral**. This can **not** be interpreted as a Stieljes integral for each trajectory. We need a new theory for this.

Information

Consider a Wiener process W.

(Informal) Definition. $\mathcal{F}_t^W$ is "The information generated by W over the interval $[0, t]$". The collection $\{\mathcal{F}_t^W,\ t \geq 0\}$ is called a filtration.

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

Example. For the stochastic variable Z, defined by

$$Z = \int_0^5 W_s \, \mathrm{d}s,$$

we have $Z \in \mathcal{F}_5^W$. We do **not** have $Z \in \mathcal{F}_4^W$.

Adapted Processes

Let W be a Wiener process.

Definition. A process X is **adapted** to the filtration

$$\left\{\mathcal{F}_t^W : t \geq 0\right\}$$

if

$$X_t \in \mathcal{F}_t^W, \quad \forall t \geq 0.$$

"An adapted process does not look into the future." Adapted processes are nice integrands for stochastic integrals.

Examples

- The process

$$X_t = \int_0^t W_s \, \mathrm{d}s,$$

  is adapted.

- The process

$$X_t = \sup_{s \leq t} W_s$$

  is adapted.

- The process

$$X_t = \sup_{s \leq t+1} W_s$$

  is **not** adapted.

The Itô Integral

We will define the Itô integral

$$\int_a^b g_s \, \mathrm{d}W_s$$

for processes g satisfying the following.

- The process g is adapted.
- The process g satisfies

$$\mathbb{E}\left[\int_a^b g_s^2 \, \mathrm{d}s\right] < \infty.$$

The definition and construction will be done in two steps.

Simple Integrands

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

- g is adapted;
- there exists deterministic points $t_0, \ldots, t_n$ with $a = t_0 < t_1 < \ldots < t_n = b$ such that g is piecewise constant (but may still be random), i.e.

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

For simple g we define the simple integral

$$\int_a^b g_s \, \mathrm{d}W_s = \sum_{k=0}^{n-1} g(t_k)[W(t_{k+1}) - W(t_k)].$$

Note the *forward* increments $W(t_{k+1}) - W(t_k)$ at time t_k.

Properties of the simple integral

Theorem. For simple g the following relations hold.

- The expected value is given by

$$\mathbb{E}\left[\int_a^b g_s \,\mathrm{d}W_s\right] = 0.$$

- The second moment (variance) is given by

$$\mathbb{E}\left[\left(\int_a^b g_s \,\mathrm{d}W_s\right)^2\right] = \int_a^b \mathbb{E}\left[g_s^2\right] \,\mathrm{d}s.$$

- We have

$$\int_a^b g_s \,\mathrm{d}W_s \in \mathcal{F}_b^W.$$

General Case

For a general g we do as follows.

1. Approximate g with a sequence of simple g_n such that
$$\int_a^b \mathbb{E}\left[\{g_n(s) - g(s)\}^2\right] \mathrm{d}s \to 0.$$
2. For each n the integral
$$\int_a^b g_n(s)\,\mathrm{d}W(s)$$
is a well defined stochastic variable Z_n.
3. One can show that the Z_n sequence converges to a limiting stochastic variable Z, $\mathbb{E}(Z_n - Z)^2 \to 0$.
4. We define $\int_a^b g(s)\,\mathrm{d}W(s)$ (in short $\int_a^b g\,\mathrm{d}W$) by the above limit Z,
$$\int_a^b g(s)\,\mathrm{d}W(s) =: Z = \lim_{n\to\infty} \int_a^b g_n(s)\,\mathrm{d}W(s).$$

Properties of the general integral

Theorem. For general g following relations hold.

- The expected value is given by

$$\mathbb{E}\left[\int_a^b g_s \,\mathrm{d}W_s\right] = 0.$$

- We do in fact have the conditional expectation

$$\mathbb{E}\left[\int_a^b g_s \,\mathrm{d}W_s \mid \mathcal{F}_a\right] = 0.$$

  Later we will call this the martingale property of the stochastic integral.

Properties of the general integral continued

- The second moment is given by

$$\mathbb{E}\left[\left(\int_a^b g_s \,\mathrm{d}W_s\right)^2\right] = \int_a^b \mathbb{E}\left[g_s^2\right] \,\mathrm{d}s.$$

- We have

$$\int_a^b g_s \,\mathrm{d}W_s \in \mathcal{F}_b^W$$

Martingales

Definition. An adapted process is a **martingale** if

$$\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s, \quad \forall s \leq t.$$

One can see a martingale as "a process without drift". Conditional expectation of a future value is the current one.

Proposition. For any g (sufficiently integrable) the process

$$X_t = \int_0^t g_s \, \mathrm{d}W_s$$

is a martingale.

Proposition: If X has dynamics

$$\mathrm{d}X_t = \mu_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}W_t,$$

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

Stochastic Calculus

(Chapters 4–5)

Stochastic Calculus, basic question

Consider the **General Model**

$$\mathrm{d}X_t = \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t.$$

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

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

Problem. What does $\mathrm{d}f(t, X_t)$ look like?

The answer is given by the **Itô formula**. But we provide an intuitive argument only. The formal proof is very hard.

A close up of the Wiener process

Consider an "infinitesimal" Wiener increment

$$\mathrm{d}W_t = W_{t+\mathrm{d}t} - W_t.$$

We have seen before

$$\begin{aligned} \mathrm{d}W_t &\sim N[0, \mathrm{d}t], \\ \mathbb{E}[\mathrm{d}W_t] = 0,\ &\mathbb{V}\mathrm{ar}[\mathrm{d}W_t] = \mathrm{d}t. \end{aligned}$$

From this one can show

$$\mathbb{E}[(\mathrm{d}W_t)^2] = \mathrm{d}t,\ \mathbb{V}\mathrm{ar}[(\mathrm{d}W_t)^2] = 2(\mathrm{d}t)^2.$$

Some observations

Recall

$$\mathbb{E}[(\mathrm{d}W_t)^2] = \mathrm{d}t,\ \mathbb{V}\mathrm{ar}[(\mathrm{d}W_t)^2] = 2(\mathrm{d}t)^2.$$

Important observations.

- Both $\mathbb{E}[(\mathrm{d}W_t)^2]$ and $\mathbb{V}\mathrm{ar}[(\mathrm{d}W_t)^2]$ are very small when $\mathrm{d}t$ is treated as a small number.
- $\mathbb{V}\mathrm{ar}[(\mathrm{d}W_t)^2]$ is negligable compared to $\mathbb{E}[(\mathrm{d}W_t)^2]$.
- Thus $(\mathrm{d}W_t)^2$ is **deterministic**.

We thus conclude, *at least intuitively*, that

$$(\mathrm{d}W_t)^2 = \mathrm{d}t.$$

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

Multiplication table.

The important observations lead to the following result. **Theorem.** We have the multiplication table

$$\begin{aligned}(\mathrm{d}t)^2 &= 0,\\ \mathrm{d}W_t \cdot \mathrm{d}t &= 0,\\ (\mathrm{d}W_t)^2 &= \mathrm{d}t.\end{aligned}$$

Deriving the Itô formula I

Start from

$$\begin{aligned} \mathrm{d}X_t &= \mu_t \,\mathrm{d}t + \sigma_t \mathrm{d}W_t, \\ Z_t &= f(t, X_t). \end{aligned}$$

We want to compute $\mathrm{d}f(t, X_t)$.

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

$$\begin{aligned} \mathrm{d}f = &\frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial x}\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(\mathrm{d}t)^2 \\ &+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\mathrm{d}X_t)^2 + \frac{\partial^2 f}{\partial t \partial x}\,\mathrm{d}t \cdot \mathrm{d}X_t \end{aligned}$$

Plug in the expression for $\mathrm{d}X$, expand, and use the multiplication table!

Deriving the Itô formula II

First computations are

$$
\begin{aligned}
\mathrm{d}f &= \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial x}[\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W] + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(\mathrm{d}t)^2 \\
&\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}[\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W]^2 + \frac{\partial^2 f}{\partial t \partial x}\,\mathrm{d}t \cdot [\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W] \\
&= \frac{\partial f}{\partial t}\,\mathrm{d}t + \mu\frac{\partial f}{\partial x}\,\mathrm{d}t + \sigma\frac{\partial f}{\partial x}\,\mathrm{d}W + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(\,\mathrm{d}t)^2 \\
&\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}[\mu^2(\mathrm{d}t)^2 + \sigma^2(\mathrm{d}W)^2 + 2\mu\sigma\,\mathrm{d}t \cdot \mathrm{d}W] \\
&\quad + \mu\frac{\partial^2 f}{\partial t \partial x}(\mathrm{d}t)^2 + \sigma\frac{\partial^2 f}{\partial t \partial x}\,\mathrm{d}t \cdot \mathrm{d}W.
\end{aligned}
$$

The Itô Formula

Using the multiplication table this reduces to

$$\mathrm{d}f = \left\{ \frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} \right\} \mathrm{d}t + \sigma \frac{\partial f}{\partial x} \mathrm{d}W.$$

We summarize the result as

Theorem. With X having dynamics as on slide 34 we have

$$\mathrm{d}f(t, X_t) = \left\{ \frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} \right\} \mathrm{d}t + \sigma \frac{\partial f}{\partial x} \, \mathrm{d}W_t.$$

Alternatively

$$\mathrm{d}f(t, X_t) = \frac{\partial f}{\partial t} \, \mathrm{d}t + \frac{\partial f}{\partial x} \mathrm{d}X_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (\mathrm{d}X_t)^2.$$

The multidimensional Itô Formula

Let X is a multidimensional process with dynamics (all quantities of appropriate dimensions)

$$\mathrm{d}X = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W,$$

and $Z_t = f(t, X_t)$. Then

$$\begin{aligned}\mathrm{d}Z_t &= \frac{\partial f}{\partial t}\mathrm{d}t + \nabla_X f^\top \mathrm{d}X_t + \tfrac{1}{2}\mathrm{d}X_t^\top (H_X f)\,\mathrm{d}X_t \\ &= \left\{\frac{\partial f}{\partial t} + \nabla_X f^\top \mu_t + \tfrac{1}{2}\mathrm{Tr}[\sigma_t^\top H_X f\,\sigma_t]\right\}\mathrm{d}t + \nabla_X f^\top \sigma_t \mathrm{d}W_t.\end{aligned}$$

Here $\nabla_X f$ is the gradient of f, the vector of partial derivatives of f w.r.t. the x_i variables, and $H_X f$ the Hessian, the matrix of second order partial derivatives.

The twodimensional Itô Formula

Here X is a two dimensional process, $X = (X^1, X^2)$ with

$$\mathrm{d}X^i = \mu^i \,\mathrm{d}t + \sigma^i \,\mathrm{d}W,\ i = 1, 2,$$

and W a one-dimensional Wiener process. We explicitly write down the gradient and the Hessian. We have

$$\nabla_X f = (f_{x_1}, f_{x_2})^\top$$
$$H_X f = \begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{pmatrix}.$$

Hence

$$\begin{aligned} \mathrm{d}Z_t = \frac{\partial f}{\partial t}\mathrm{d}t + f_{x_1}\,\mathrm{d}X^1 + f_{x_2}\,\mathrm{d}X^2 \\ + \tfrac{1}{2}\left(f_{x_1x_1}(\mathrm{d}X^1)^2 + 2f_{x_1x_2}(\mathrm{d}X^1)(\mathrm{d}X^2) + f_{x_2x_2}(\mathrm{d}X^2)^2\right). \end{aligned}$$

Special case of the two-dimensional Itô Formula

Let now $f(t,x) = x^1x^2$, the product. Then $Z_t = X_t^1 X_t^2$, and we have the Itô *product rule*,

$$\begin{aligned}\mathrm{d}Z_t &= X_t^1\,\mathrm{d}X_t^2 + X_t^2\,\mathrm{d}X_t^1 + \mathrm{d}X_1^1\,\mathrm{d}X_t^2 \\ &= (X_t^1\mu_t^2 + X_t^2\mu_t^1 + \sigma_t^1\sigma_t^2)\,\mathrm{d}t + (X_t^1\sigma_t^2 + X_t^2\sigma_t^1)\,\mathrm{d}W_t.\end{aligned}$$

Verify that this product formula follows from the general two-dimensional Itô formula.

In discrete time, one has for sequences $x_0^i, x_1^i, \ldots$ the analog

$$\Delta(x^1x^2)_n = x_{n-1}^1\Delta x_n^2 + x_{n-1}^2\Delta x_n^1 + \Delta x_n^1\Delta x_n^2,\ n \geq 1,$$

where we used the generic notation $\Delta y_n = y_n - y_{n-1}$. Observe the similarities with the first equation for $\mathrm{d}Z_t$.

Very special cases of the two-dimensional Itô Formula

If (at least) one of the processes X^1, X^2 has no Wiener term in its differential, the product term $\mathrm{d}X_1^1\,\mathrm{d}X_t^2$ vanishes, in which case we get the ordinary product rule

$$\mathrm{d}(X_t^1 X_t^2) = X_t^1\,\mathrm{d}X_t^2 + X_t^2\,\mathrm{d}X_t^1.$$

Suppose that $X^1 = X$, $X^2 = \frac{1}{Y}$ and that the dynamics of Y don't contain a Wiener process. Then ordinary calculus gives (with Y^2 now denoting the square of Y) $\mathrm{d}\frac{1}{Y} = -\frac{1}{Y^2}\,\mathrm{d}Y$. Hence

$$\mathrm{d}\frac{X}{Y} = \frac{\mathrm{d}X}{Y} - \frac{X}{Y^2}\,\mathrm{d}Y.$$

Example: GBM

Consider

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t.$$

We smell something exponential! This leads to a natural Ansatz,

$$S_t = e^{Z_t},\ Z_t = \log S_t.$$

Itô on $f(t,s) = \log s$ gives us

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

This brings us

$$\begin{aligned}\mathrm{d}Z_t &= \frac{1}{S_t}\mathrm{d}S_t - \frac{1}{2}\frac{1}{S_t^2}(\mathrm{d}S_t)^2 \\ &= (\mu - \frac{1}{2}\sigma^2)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t.\end{aligned}$$

Example GBM continued

Recall

$$\mathrm{d}Z_t = (\mu - \frac{1}{2}\sigma^2)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t.$$

Integrate! This yields

$$\begin{aligned} Z_t - Z_0 &= \int_0^t (\mu - \frac{1}{2}\sigma^2)\,\mathrm{d}s + \sigma \int_0^t \mathrm{d}W_s \\ &= (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t. \end{aligned}$$

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

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

Since W_t is $N[0, t]$, we see that S_t has a lognormal distribution, as the exponential has a $N\left((\mu - \frac{1}{2}\sigma^2))t, \sigma^2 t\right)$ distribution.

A Useful Trick

Problem. Compute $\mathbb{E}[Z(T)]$.

- Use Itô to get an equation of the type

$$\mathrm{d}Z(t) = \mu_Z(t)\,\mathrm{d}t + \sigma_Z(t)\,\mathrm{d}W_t.$$

- Integrate,

$$Z(T) = z_0 + \int_0^T \mu_Z(t)\,\mathrm{d}t + \int_0^T \sigma_Z(t)\,\mathrm{d}W_t$$

- Take expectations, and use that Itô integrals have expectation zero,

$$\mathbb{E}[Z(T)] = z_0 + \int_0^T \mathbb{E}[\mu_Z(t)]\,\mathrm{d}t + 0.$$

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

The connection SDE to PDE

Given are $\mu(t,x)$, $\sigma(t,x)$, $\Phi(x)$, T.

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

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

where $\mathcal{A}$ is defined by

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

Interesting enough, Stochastic Calculus helps to find a solution to this seemingly unrelated problem.

Using Itô calculus

Assume that F solves the PDE and fix the point (t, x).

Define the process X by (assume this can be done)

$$\begin{aligned} \mathrm{d}X_s &= \mu(s, X_s)\,\mathrm{d}t + \sigma(s, X_s)\,\mathrm{d}W_s, \\ X_t &= x. \end{aligned}$$

One can show, under appropriate assumptions, that X is then a Markov process.

Apply Itô to the process $F(t, X_t)$ and verify the result,

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

Using Itô calculus continued

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

$$\Phi(X_T) = F(t, x) + \int_t^T \sigma(s, X_s) \frac{\partial F}{\partial x}(s, X_s) \, \mathrm{d}W_s.$$

Take conditional expectation given $\mathcal{F}_t$. By the Markov property this reduces to an expectation given X_t only. One gets in often used notation to express that we know $X_t = x$,

$$F(t, x) = \mathbb{E}_{t,x}[\Phi(X_T)].$$

Next we change the PDE (we add a term $-rF$) to

$$\frac{\partial F}{\partial t}(t, x) + \mathcal{A}F(t, x) - rF = 0,$$
$$F(T, x) = \Phi(x).$$

The solution to this equation will be a variation on the one we just obtained.

Feynman-Kac formula

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

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

is given by

$$F(t,x) = e^{-r(T-t)}\mathbb{E}_{t,x}[\Phi(X_T)],$$

where X satisfies the SDE

$$\begin{aligned} \mathrm{d}X_s &= \mu(s,X_s)\,\mathrm{d}t + \sigma(s,X_s)\,\mathrm{d}W_s, \\ X_t &= x. \end{aligned}$$

Lecture 1

Black-Scholes

(Chapters 6-7)

Contents of Lecture 1

- Introduction.
- Portfolio theory.
- Deriving the Black-Scholes PDE
- Risk neutral valuation

Introduction

European Call Option, A contract written on paper with the text:

The holder of this paper has the right, not the obligation,

to buy

1 ACME INC

on the date

June 30, 2027

(or another fixed future date, if you wish)

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
- CDOs
- CDSs

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 Z at time T.

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

Natural answers (correct?):

1. Price = Discounted present value of future payouts,

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

   when the interest rate is r.

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** the underlying price.
- Calls for consistent pricing.
- Relative pricing.

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

Portfolio Theory

Portfolios

We consider a market with N assets.

$$S_t^i = \text{price at } t, \text{of asset No. } i,\ i = 1, \ldots, N.$$

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

$$h_t = (h_t^1, \ldots, h_t^N), x$$

where h_t^i = number of units of asset i.

Furthermore, V_t is the market value of the portfolio,

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

The portfolio is typically of the form $h_t = h(t, S_t)$, i.e. today's portfolio is based on today's prices, (sometimes also on prices from the past).

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 (note the Itô differentials)

$$\mathrm{d}V_t = \sum_{i=1}^{N} h_t^i \, \mathrm{d}S_t^i.$$

Interpret!

See Appendix B of Lecture 2 (slide 96 and further) for details and motivation from discrete time.

Accept this definition for the time being.

Relative weights

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

$$\omega_t^i := \frac{h_t^i S_t^i}{V_t} \quad \text{and then} \quad h_t^i = \omega_t^i \frac{V_t}{S_t^i}.$$

Insert this into the self financing condition

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

to obtain the **Portfolio dynamics:**

$$\mathrm{d}V_t = V_t \sum_{i=1}^{N} \omega_t^i \frac{\mathrm{d}S_t^i}{S_t^i} \qquad \left(\text{equivalent to } \frac{\mathrm{d}V_t}{V_t} = \sum_{i=1}^{N} \omega_t^i \frac{\mathrm{d}S_t^i}{S_t^i} \right).$$

Interpret! (see also slide 103)

Deriving the Black-Scholes PDE

Back to Financial Derivatives

Consider the Black-Scholes model (W is Wiener process, a Brownian motion, r is the interest rate),

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t,$$
$$\mathrm{d}B_t = rB_t \,\mathrm{d}t \text{ (bank account).}$$

It results that $B_t = e^{rt}B_0$, where usually $B_0 = 1$ (a normalization).

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, Z, given by

$$Z = (S_T - K)^+ = \max\{S_T - K, 0\}.$$

More general, one can take

$$Z = \Phi(S_T) \text{for some contract function } \Phi.$$

Main problem: What is a "reasonable" price, $\Pi_t[Z]$, for 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), i.e., a *viable* market.

Arbitrage

The (relative) portfolio ω is an **arbitrage portfolio** if

- the portfolio strategy is self financing;
- $V_0 = 0$;
- $V_T > 0$ with probability one.
  (Or, weaker, $\mathbb{P}(V_T \geq 0) = 1$, *and* $\mathbb{P}(V_T > 0) > 0$. *See slide 198.)*

Moral:

- Arbitrage = Free Lunch.
- No arbitrage possibilities in an efficient market (arbitrage possibilities exist only in a market with "wrong" prices).

Arbitrage test *(fundamental idea)*

Suppose that a portfolio ω is self financing with dynamics

$$\mathrm{d}V_t = kV_t\,\mathrm{d}t.$$

Note the following.

- 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

$$k = r \quad \text{(because, otherwise, } \ldots\text{)}$$

Remark: k and r may depend on t. Then we get $k_t = r_t$.

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

$$k = r \quad (\text{or } k_t = r_t).$$

$\longrightarrow$ **End of lecture 1a** $\longleftarrow$

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. (True?)
- 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 of Lecture 2, slide 104 and further.

Formalized program a la Merton (outline)

- Assume that the derivative price is of the form $\Pi_t[Z] = f(t, S_t)$.
- Form a **self financing** portfolio based on the underlying S and the derivative f, with portfolio dynamics with relative weights ω_t^S, ω_t^f (recall the general case from slide 62). The value process V satisfies

$$\mathrm{d}V_t = V_t \left\{ \omega_t^S \cdot \frac{\mathrm{d}S_t}{S_t} + \omega_t^f \cdot \frac{\mathrm{d}f}{f} \right\}.$$

- Choose ω^S and ω^f such that the $\mathrm{d}W$-term is wiped out. This will gives us something like

$$\mathrm{d}V_t = V_t k \, \mathrm{d}t.$$

- Absence of arbitrage implies

$$k = r.$$

- This relation will say something about f, as we shall see.

Back to Black-Scholes

The model is

$$\begin{aligned} \mathrm{d}S_t &= \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \\ \Pi_t[Z] &= f(t, S_t). \end{aligned}$$

Itô's formula gives us the f dynamics as

$$\mathrm{d}f = \left\{ \frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial s} + \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial s^2} \right\} \mathrm{d}t + \sigma S \frac{\partial f}{\partial s} \,\mathrm{d}W.$$

Write this as

$$\mathrm{d}f = \mu_f \cdot f \,\mathrm{d}t + \sigma_f \cdot f \,\mathrm{d}W,$$

where (both expressed in terms of partial derivatives of f)

$$\mu_f = \frac{\frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial s} + \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial s^2}}{f},$$

$$\sigma_f = \frac{\sigma S \frac{\partial f}{\partial s}}{f}.$$

Black-Scholes continued

Recall from the previous slide $\mathrm{d}f = \mu_f f\,\mathrm{d}t + \sigma_f f\,\mathrm{d}W$, and compute

$$\begin{aligned}\mathrm{d}V &= V\left[\omega^S\frac{\mathrm{d}S}{S} + \omega^f\frac{\mathrm{d}f}{f}\right]\\ &= V\left[\omega^S(\mu\mathrm{d}t + \sigma\mathrm{d}W) + \omega^f(\mu_f\,\mathrm{d}t + \sigma_f\,\mathrm{d}W)\right]\\ &= V\left[\omega^S\mu + \omega^f\mu_f\right]\mathrm{d}t + V\left[\omega^S\sigma + \omega^f\sigma_f\right]\mathrm{d}W.\end{aligned}$$

Now we kill the $\mathrm{d}W$-term! We choose (ω^S, ω^f) such that

$$\omega^S\sigma + \omega^f\sigma_f = 0,$$

$$\omega^S + \omega^f = 1 \text{ as relative weights sum to 1.}$$

This is a linear system with solution (when $\sigma_f \neq \sigma$)

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

Plug this into $\mathrm{d}V$!

Black-Scholes continued

We obtain

$$\mathrm{d}V = V\left[\omega^S \mu + \omega^f \mu_f\right] \mathrm{d}t.$$

This is a risk free "synthetic bank" with short rate $\omega^S \mu + \omega^f \mu_F$.

Absence of arbitrage implies

$$\omega^S \mu + \omega^f \mu_f = r.$$

Plug in (you verify the computations!) the expressions, that all involve partial derivatives as on slides 71 and 72, for ω^S, ω^f, μ_f and simplify.

This will give us the following result.

Black-Scholes' PDE

The price $\Pi_t[Z]$ of the derivative is given by

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

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

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

There is a **unique solution** (this is a **Theorem!**) to this PDE, and so there is a **unique arbitrage free** price process for the contract.

Black-Scholes' PDE continued

Recall the PDE from the previous slide, now in simplified notation without arguments t, s,

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

- 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 to compute the price

- 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!

A miracle ??

We'll come back this later.

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 for a European Call

The ingredients are

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

The PDE with $\Phi(S_T) = (S_T - K)^+$ is solved by

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

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

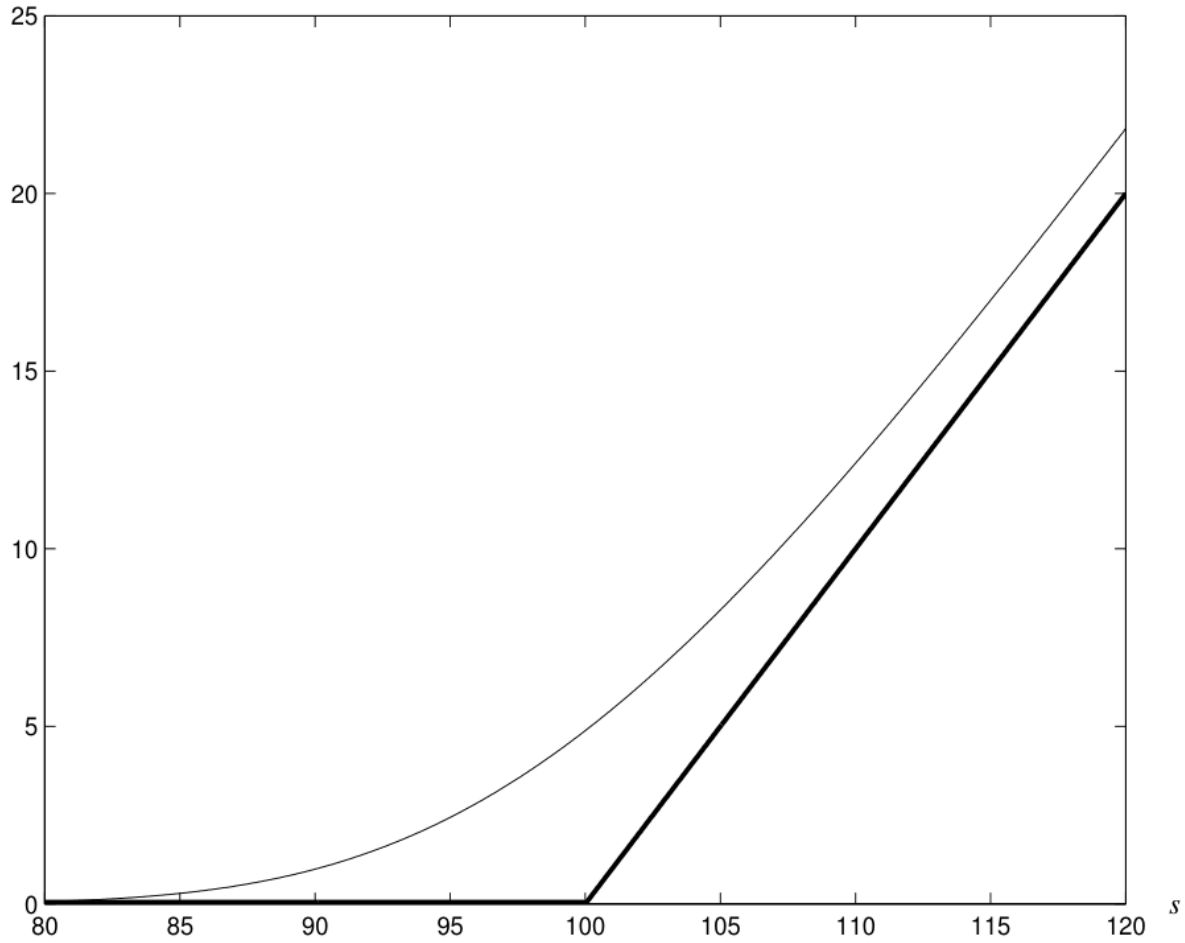
$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\log\left(\frac{s}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)\right],$$
$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

This formula comes out of the blue for the time being; but this f solves the Black-Scholes PDE (check!). See slide 85 for $t = 0$.

Black-Scholes properties in a picture

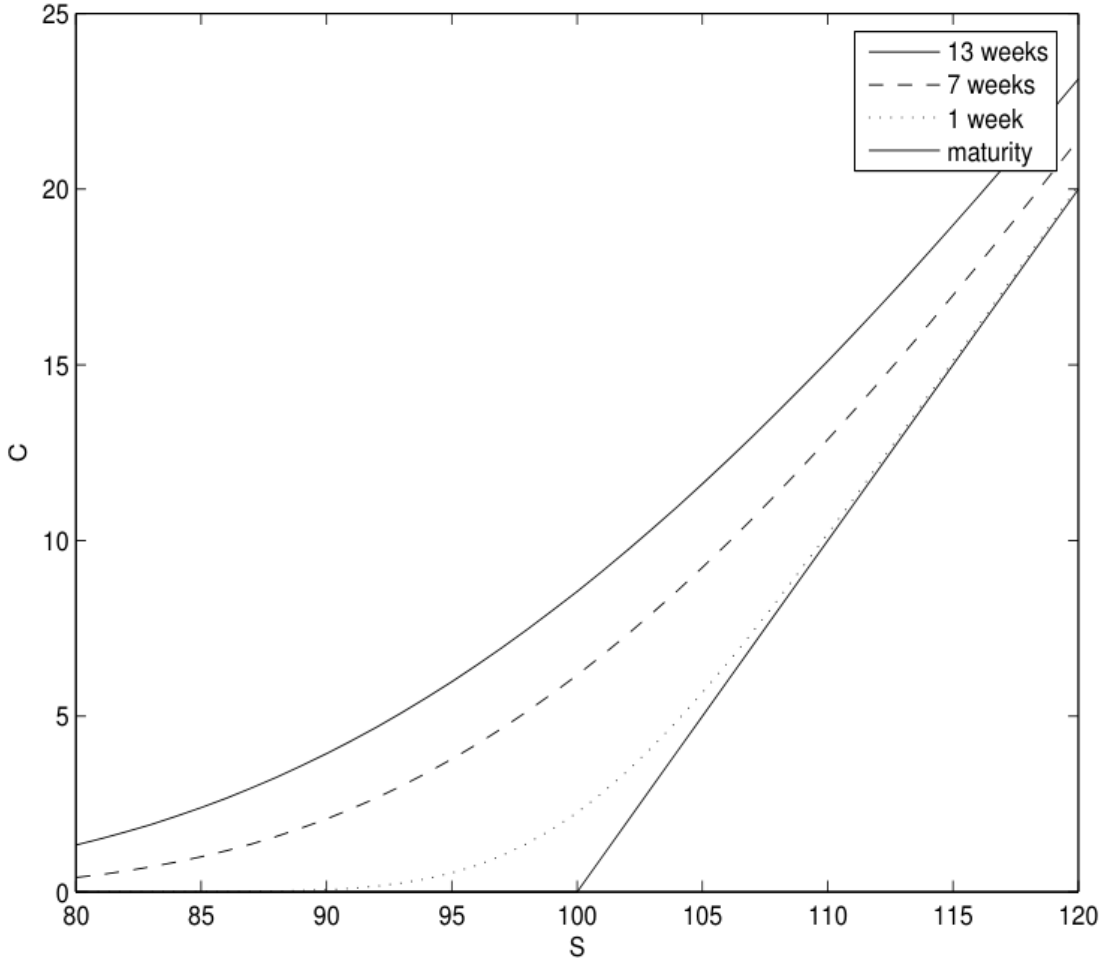
Take the European Call with,

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

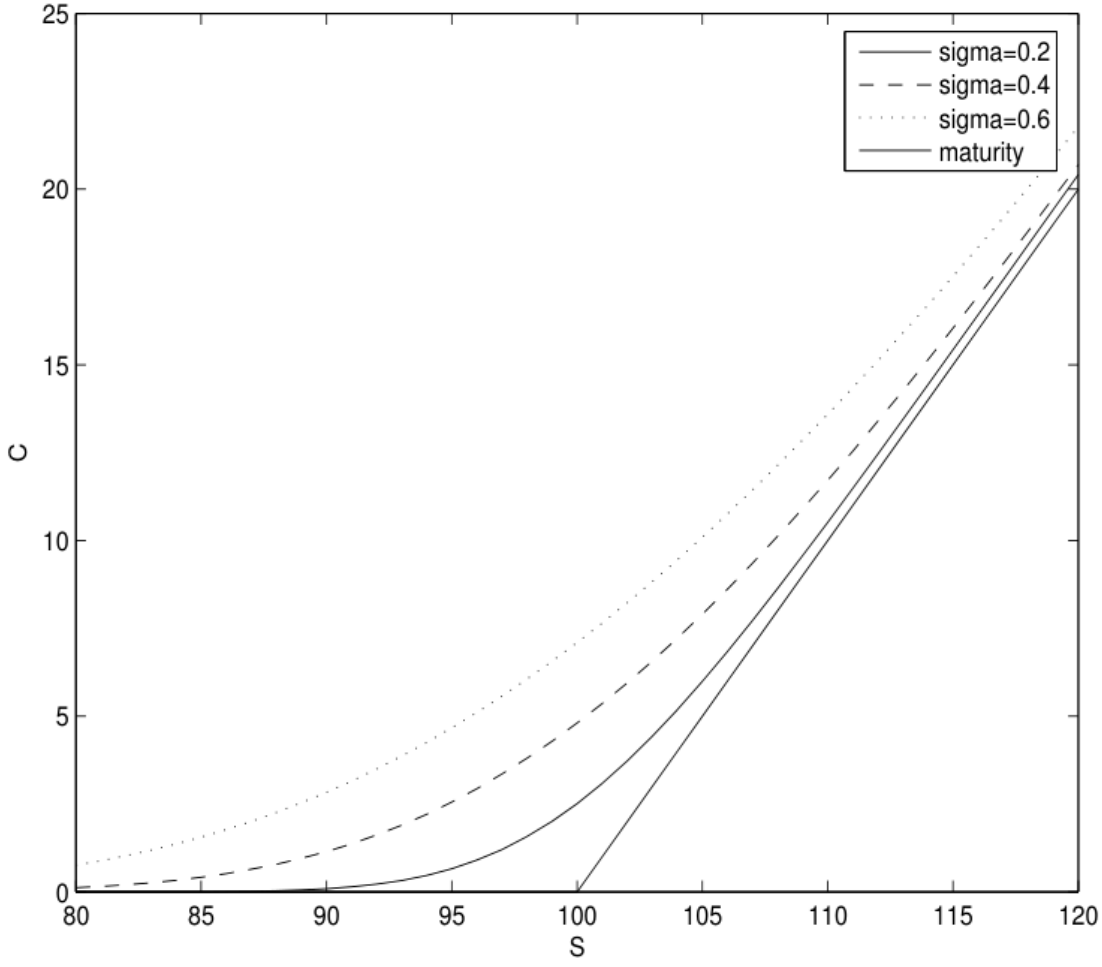


(f versus stock price s)

Dependence on Time to Maturity



Dependence on Volatility



Risk Neutral Valuation

Risk neutral valuation

Apply Feynman-Kac (slide 47) to the Black-Scholes PDE,

$$\Pi[t; X] = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,s}[X].$$

Notation: $\mathbb{E}^{\mathbb{Q}}_{t,s}[\cdot]$ = conditional expectation at time t, with $S_t = s$, under the measure $\mathbb{Q}$.

This is related to the $\mathbb{Q}$-dynamics, $W^{\mathbb{Q}}$ is Wiener under $\mathbb{Q}$,

$$\begin{cases} \mathrm{d}S_t = rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W^{\mathbb{Q}}_t, \\ \mathrm{d}B_t = rB_t\,\mathrm{d}t. \end{cases}$$

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

Note: $\mathbb{P} \sim \mathbb{Q}$ (Girsanov, equivalence of the two probability measures on $\mathcal{F}_T$. See slide 190 for the underlying theory.)

Concrete formulas I

From the $\mathbb{Q}$ dynamics of S_t one obtains that $\log(S_T/S_0)$ has the normal $N((r-\frac{1}{2}\sigma^2)T, \sigma^2 T)$ distribution.

At $t=0$ we have, with $S_0 = s$,

$$\Pi[0;\Phi] = e^{-rT}\int_{-\infty}^{\infty}\Phi(se^z)f(z)\,\mathrm{d}z$$

with

$$f(z) = \frac{1}{\sigma\sqrt{2\pi T}}\exp\left(-\frac{\left(z-(r-\frac{1}{2}\sigma^2)T\right)^2}{2\sigma^2 T}\right),$$

the density of the $N((r-\frac{1}{2}\sigma^2)T, \sigma^2 T)$ distribution.

Concrete formulas II

For the European call with strike K one gets

$$\begin{aligned}\Pi[0;\Phi] &= e^{-rT}\int_{-\infty}^{\infty}(se^z-K)^+ f(z)\,\mathrm{d}z\\ &= e^{-rT}\int_{\log\frac{K}{s}}^{\infty} se^z f(z)\mathrm{d}z - K\int_{\log\frac{K}{s}}^{\infty} f(z)\,\mathrm{d}z.\end{aligned}$$

Exercise. Compute explicitly the integrals with f as on slide 84 to arrive at the formula on slide 78.

Interpretation of the risk adjusted measure

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

$$s = S_0 = e^{-rt}\mathbb{E}^{\mathbb{Q}}[S_t].$$

- In our model this means that

$$\mathrm{d}S_t = rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t^{\mathbb{Q}}.$$

- The risk adjusted probabilities can be intrepreted as probabilities in a fictitious 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".

Quick mentioning of Properties of $\mathbb{Q}$ (More on this in Lecture 5)

- $\mathbb{P} \sim \mathbb{Q}$ (Girsanov)
- For the price process π of **any** traded asset, derivative or underlying, the process

$$Z_t = \frac{\pi_t}{B_t}, \, (0 \leq t \leq T)$$

  is a $\mathbb{Q}$-*martingale* (details later).
- Under $\mathbb{Q}$, the price process π of **any** traded asset, derivative or underlying, has r as its local rate of return: i.e.,

$$\mathrm{d}\pi_t = r\pi_t \, \mathrm{d}t + \sigma_\pi \pi_t \, \mathrm{d}W_t^{\mathbb{Q}}$$

- The *volatility* of π (i.e., σ_π) is the *same* under $\mathbb{Q}$ as under $\mathbb{P}$.

A Preview of Equivalent Martingale Measures (EMMs)

Remark. *More on this on slides 207 and further.*

Consider a market, under an objective probability measure $\mathbb{P}$, with underlying assets $B, S^1, \ldots, S^N$.

Definition. A probability measure $\mathbb{Q}$ on $\mathcal{F}_T$ is called an **equivalent martingale measure** if

- $\mathbb{P} \sim \mathbb{Q}$,
- For every i, the process $Z_t^i = \frac{S_t^i}{B_t}$, $(0 \leq t \leq T)$, is a $\mathbb{Q}$-martingale.

Theorem [FTAP 1 – first Fundamental Theorem of Asset Pricing]. The market is arbitrage free **iff** there exists an equivalent martingale measure.

$\longrightarrow$ **End of lecture 1b** $\longleftarrow$

Lecture 2

Appendices to Lecture 1

Appendix A: Black-Scholes vs Binomial, [book: Chapter 2]

Consider a **binomial model** (if you know this) without interest for a stock price.

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

$$u = e^{\sigma\sqrt{\Delta t}},\ d = e^{-\sigma\sqrt{\Delta t}}, \quad \Delta t = T/n \quad (\sigma > 0,\ d < 1 < u)$$

- 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 (n divisions of $[0, T]$ with $T = n\Delta t$).
- More precisely, with n divisions of $[0, T]$, the number of periods in the binomial model, with $T = n\Delta T$,

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

where $Z_i \in \{u, d\}$.

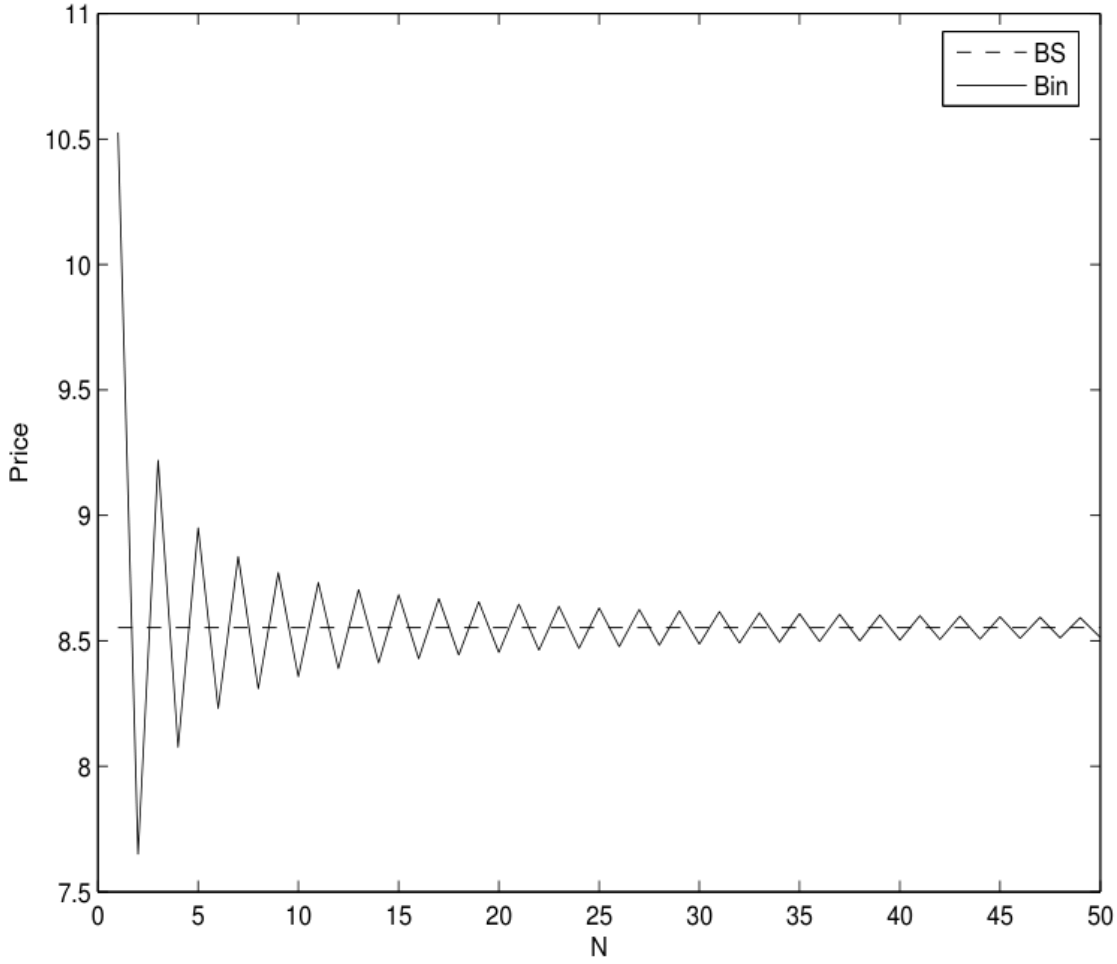
- In S_T, the number of u's and d's matters only, not the order (like successes/failures in Binomial trials).

Recall (this is the *Cox-Ross-Rubinstein model)*

Recall $S_T = S_0 Z_1 Z_2 \cdots Z_n$ and let $n \to \infty$.

- The stock price at time T will be a *product* of a large number of i.i.d. random variables.
- The **log-return** will be a large **sum** of i.i.d. variables: $\log S_T = \log S_0 + \sum_{i=1}^{n} \log Z_i$.
- The Central Limit Theorem will kick in (details omitted).
- In the limit, log-returns will be **normally** distributed.
- Hence, stock prices will be **lognormally** distributed.
- We are in the Black-Scholes model.
- The binomial price will **converge** to the Black-Scholes price. This is a **Theorem**, we give no proof/details.

Binomial convergence to Black-Scholes



Binomial $\sim$ Black-Scholes

The intuition, if you have it, from the Binomial model and its properties carry 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 (this notion comes later).
- *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.

(These statements are actually theorems. Take them for granted.)

Remark. Binomial models have been used in practice (even in Excel).

Appendix B: Portfolio theory *(this is a copy of slide 60)*

We consider a market with N assets.

$$S_t^i = \text{price at } t, \text{of asset No } i,\ i = 1, \ldots, N.$$

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

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

where h_t^i = number of units of asset i.

Furthermore, V_t = the market value of the portfolio,

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

The portfolio is typically of the form $h_t = h(t, S_t)$, i.e. today's portfolio is based on today's prices, (sometimes also on prices from the past).

Self financing portfolios

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

- there is no 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 $\mathrm{d}V_t$ (the one already presented on slide 61) for a *self financing portfolio*.

Approach. We analyze in *discrete time*, and then go to the continuous time *limit*, in a heuristic way.

Discrete time portfolios

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

Price vector process:

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

Portfolio process:

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

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

Value process:

$$V_n = \sum_{i=1}^{N} h_n^i S_n^i = h_n S_n \quad \text{(inner product notation)}.$$

The self financing condition in discrete time

- At time $n-1$ we buy the portfolio h_{n-1} at the price S_{n-1}.
- At time n this portfolio is worth $h_{n-1}S_n$ [prices jump from S_{n-1} to 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, a useful lemma

Recall

$$V_n = h_n S_n.$$

Definition. For any sequence $x_0, x_1, x_2, \ldots$ we define the difference sequence Δx_n by $\Delta x_n = x_n - x_{n-1}$, $n \geq 1$.

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

Lemma. For any pair of sequences $x_0, x_1, x_2, \ldots$ and $y_0, y_1, y_2, \ldots$ we have the relation, for $n \geq 1$,

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

(This is an abbreviation of Abel's summation formula: $x_n y_n - x_0 y_0 = \sum x_{i-1}\Delta y_i + \sum y_i \Delta x_i$*, like integration by parts!)*

Proof. Do it yourself.

The self financing condition in a proposition

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 as a budget contraint,

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

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

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

Portfolios in continuous time

Price process: $S_t^i =$ price at $t \in [0, T]$, of asset No. i.
Portfolio: $h_t = (h_t^1, \ldots, h_t^N)$, value process $V_t = \sum_{i=1}^N h_t^i S_t^i$.

From the self financing condition in discrete time, see slide 61,

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

we are led to the following definition *(by analogy!)*.
Definition. The portfolio h is self financing if and only if (alluded to on page 54)

$$\mathrm{d}V_t = \sum_{i=1}^N h_t^i \, \mathrm{d}S_t^i.$$

(It is needed that the S^i are "semimartingales" or diffusions, the Itô differentials need to make sense.)

Relative weights [this slide is a copy of slide 62]

Definition. ω_t^i = relative portfolio weight on asset No i. We have

$$\omega_t^i = \frac{h_t^i S_t^i}{V_t} \quad \text{and then} \quad h_t^i = \omega_t^i \frac{V_t}{S_t^i}.$$

Insert this into the self financing condition

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

to obtain the **Portfolio dynamics**

$$\mathrm{d}V_t = V_t \sum_{i=1}^{N} \omega_t^i \frac{\mathrm{d}S_t^i}{S_t^i} \qquad \left(\text{equivalent to } \frac{\mathrm{d}V_t}{V_t} = \sum_{i=1}^{N} \omega_t^i \frac{\mathrm{d}S_t^i}{S_t^i} \right).$$

Interpret!

Appendix C: The original Black-Scholes PDE argument

Recall the BS model, $\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t$.

Consider the following portfolio.

- *Short* ("borrow") one unit of the derivative, with pricing function $f(t, s)$: *you have* -1 *as a quantity*.
- Hold x units (or x_t at time t; later we find a "good" x) of the underlying S.

The portfolio value is given by ($x_t = x$, shorthand notation):

$$V = -f(t, S_t) + x_t S_t.$$

The object is to choose x_t such that the portfolio is risk free for an infinitesimal interval of length $\mathrm{d}t$.

Manipulate with V_t I

We have $\mathrm{d}V = -\mathrm{d}f + x_t\,\mathrm{d}S$ (self financing!) and from Itô (for $\mathrm{d}f$, use $\mathrm{d}S_t$) we obtain

$$\begin{aligned}
\mathrm{d}V &= -\left\{\frac{\partial f}{\partial t} + \mu S\frac{\partial f}{\partial s} + \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial s^2}\right\}\mathrm{d}t \\
&\quad - \sigma S\frac{\partial f}{\partial s}\mathrm{d}W + \underbrace{x\mu S\,\mathrm{d}t + x\sigma S\,\mathrm{d}W}_{x_t\,\mathrm{d}S_t} \\
&= \left[x\mu S - \frac{\partial f}{\partial t} - \mu S\frac{\partial f}{\partial s} - \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial s^2}\right]\mathrm{d}t \\
&\quad + \sigma S\left[x - \frac{\partial f}{\partial s}\right]\mathrm{d}W.
\end{aligned}$$

We obtain a risk free portfolio if we choose x as

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

Manipulate with V_t II

Then we have, after simplification,

$$\mathrm{d}V = \left[-\frac{\partial f}{\partial t} - \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial s^2}\right]\mathrm{d}t.$$

Using $V = -f + xS$ and x as above, the return $\mathrm{d}V/V$ is thus given by

$$\frac{\mathrm{d}V}{V} = \frac{-\frac{\partial f}{\partial t} - \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial s^2}}{-f + S\frac{\partial f}{\partial s}}\,\mathrm{d}t.$$

We had *(previous slide)*

$$\frac{\mathrm{d}V}{V} = \frac{-\frac{\partial f}{\partial t} - \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial s^2}}{-f + S\frac{\partial f}{\partial s}}\,\mathrm{d}t.$$

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

$$\frac{-\frac{\partial f}{\partial t} - \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial s^2}}{-f + S\frac{\partial f}{\partial s}} = r\,! \quad \text{(recall slide 67)}$$

Simplifying this expression gives us the Black-Scholes PDE:

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

$\longleftrightarrow$ **End of lecture 2a** $\longleftarrow$

Completeness and Hedging

(Chapters 8-9)

Problems around Standard Black-Scholes

- We **assumed** that the derivative was traded. How do we price OTC (over the counter) products?
- Why is the option price independent of the expected rate of return α (previously we used *mu* instead of α) of the underlying stock? Note that the model is now

$$\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t.$$

- 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**.

Hedging and completeness

[The contents of this page will return on slide 224.]

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

$$V_T^h = X, \quad \mathbb{P} - \text{a.s.} \qquad (\text{Notation: } V_t^h = V_t = h_t S_t, \ t \leq T).$$

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 (this is the law of one price for a reachable claim),

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

The fair price

The rough argument is as follows. If $\Pi_t(X) < V_t^h$, you sell the portfolio, buy the claim and put the surplus aside. At time T you sell the claim, and buy the portfolio back, making a profit with net cost equal to zero.

More precisely, consider a replicable claim X which we want to sell at $t = 0$. A similar argument works for $0 < t < T$.

- Compute the price $\Pi_0[X]$ and sell X at a slightly (well) *higher* price [Suppose you are able to do that].
- Buy the hedging portfolio and *invest the surplus* in the bank.
- Wait until expiration date T.
- The liabilities stemming from X are 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 $X = \Phi(S_T)$ (with an "arbitrary" function Φ). Recall from slide 102 the self financing condition for *relative* portfolios u^B, u^S and note that

$$V_t = h_t^B B_t + h_t^S S_t \Rightarrow u_t^B = \frac{h_t^B B_t}{V_t}.$$

Our aim is to find processes V, and u^B, u^S such that

$$\mathrm{d}V_t = V_t \left\{ u_t^B \frac{\mathrm{d}B_t}{B_t} + u_t^S \frac{\mathrm{d}S_t}{S_t} \right\},$$

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

i.e., (recall $\mathrm{d}B_t = rB_t\,\mathrm{d}t$, $\mathrm{d}S_t = \alpha S_t\,\mathrm{d}t + \sigma S_t \mathrm{d}W_t$)

$$\mathrm{d}V_t = V_t \left\{ u_t^B r + u_t^S \alpha \right\} \mathrm{d}t + V_t u_t^S \sigma\,\mathrm{d}W_t,$$

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

Heuristics

Let us **assume** that X is replicated by (u^B, u^S) with value process V. **Ansatz** *(reasonable, based on $X = \Phi(S_T)$ and X is Markov)*:

$$V_t = F(t, S_t) \quad \text{(for some } F \text{ to be found)}.$$

Itô gives us

$$\mathrm{d}V = \left\{ F_t + \alpha S F_s + \frac{1}{2}\sigma^2 S^2 F_{ss} \right\} \mathrm{d}t + \sigma S F_s \, \mathrm{d}W.$$

Write this as

$$\mathrm{d}V = V \left\{ \frac{F_t + \alpha S F_s + \frac{1}{2}\sigma^2 S^2 F_{ss}}{V} \right\} \mathrm{d}t + V \left(\frac{S F_s}{V} \right) \sigma \mathrm{d}W. \, (*)$$

Compare with (the $\mathrm{d}W$ and $\mathrm{d}t$ coefficients should then coincide)

$$\mathrm{d}V = V \left\{ u^B r + u^S \alpha \right\} \mathrm{d}t + V u^S \sigma \, \mathrm{d}W.$$

Finding u^B and u^S

Define u^S by *(now the time index t and S_t are explicitly written)*:

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

This gives us the equation, from $(*)$ on the previous slide,

$$\mathrm{d}V = V\left\{\frac{F_t + \frac{1}{2}\sigma^2 S^2 F_{ss}}{rF} r + u^S \alpha\right\}\mathrm{d}t + V u^S \sigma \mathrm{d}W.$$

Again, compare to

$$\mathrm{d}V = V\left\{u^B r + u^S \alpha\right\}\mathrm{d}t + V u^S \sigma\, \mathrm{d}W.$$

Natural choice for u^B is given by *(we match the $\mathrm{d}t$ terms)*:

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

The final steps

The relation $u^B + u^S = 1$ with the formulas for u^B and u^S on slide 114 gives us the Black-Scholes PDE

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

The condition

$$V_T = \Phi(S_T)$$

gives us the boundary condition

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

Moral. The model is complete and we even have *explicit formulas* for the replicating portfolio, as we know the relative portfolios u^B and u^S, and then also the h^B and h^S, see slide 112.

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 $X = \Phi(S_T)$ can be replicated by the relative portfolio (use u^B on slide 112 and the PDE):

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

The corresponding absolute portfolio is given by (recall $h_t^B = u_t^B \frac{V_t}{B_t}$, $V_t = F(t,S_t)$)

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

and the value process V^h is given by

$$V_t^h = F(t,S_t) \quad \textit{(see also book Lemma 8.4)}.$$

Notes

- Completeness explains unique price (use the PDE) – the claim is *superfluous* in the sense that there is nothing "new" compared to S and B in the market; use the PDE).
- Replicating the claim $\mathbb{P}$–a.s. $\iff$ Replicating the claim $\mathbb{Q}$–a.s. for any $\mathbb{Q} \sim \mathbb{P}$.
- Thus (Girsanov) it will *not* depend on the drift α (determined by the measure $\mathbb{P}$) 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.
- The proof is *only* given for claims of the type $\Phi(S_T)$ and under the *Ansatz* $V_t = F(t, S_t)$. A more general result is true. "Any" claim can be hedged.

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.

$R =$ number of random sources,

$N =$ number of risky assets.

Intuition: If N is large, compared to R, you have lots of possibilities of forming **clever** portfolios (for instance to annihilate the random sources and find arbitrage opportunities). 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

$$N \geq R.$$

(Compare to solve $Ax = b$, with $A \in \mathbb{R}^{m \times n}$ etc. When (unique) solution? If $m \leq n$ (if you ignore "rank" conditions)).

Example. The Black-Scholes model has $R = N = 1$. So, it is arbitrage free and complete.

$\longrightarrow$ **End of lecture 2b** $\longleftarrow$

Lecture 3

Parity Relations I

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

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

Proof. Linearity of mathematical expectation (see Feynman-Kac, slide 47), or use a no-arbitrage argument for which you note that $\Pi_T[\alpha\Phi + \beta\Psi] = \alpha\Pi_T[\Phi] + \beta\Pi_T[\Psi]$, as $\Pi_T[\Phi] = \Phi(S_T)$ etc.

Consider the following "basic" contract functions:

$$\begin{aligned} \Phi_S(x) &= x, \quad \text{(think of } x \text{ as a value of } S_T\text{)} \\ \Phi_B(x) &\equiv 1, \\ \Phi_{C,K}(x) &= \max\{x - K, 0\}. \end{aligned}$$

Parity Relations II

These basic functions lead to prices

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

with the notation in the last equation to emphasize dependence of the call-option price on the relevant quantities.

If we have for more options with strikes K_i the compound pay-off

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

then linearity extends to

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

Parity Relations III

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

- α shares of the underlying stock.
- β *zero coupon* T*-bonds* with face value €1 (these have a pay off at time T of €1, and hence their value at time t is $e^{-t(T-t)}$, see further on slide 273).
- γ_i European call options with strikes K_i, all maturing at T.

Put-Call Parity

Consider a European *put* contract,

$$\Phi_{P,K}(s) = \max\{K - s, 0\}.$$

It is easy to see (draw a figure, or do some simple algebra) that

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

We immediately get the so-called

Put-call parity,

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

Thus you can construct a synthetic put option, using a buy-and-hold portfolio (with a call option). See also Proposition 9.3 in the book.

Delta Hedging, (name has to do with the "Greeks", see further down)

Consider a fixed claim $X = \Phi(S_T)$, with pricing function

$$F(t, s) \quad (\text{or } F(t, S_t) \text{ with } S_t = s).$$

Setup. We are at time t, and have a short position in the contract (a "debt" in the contract).

Goal. Offset the risk in the derivative by buying (or selling) the (highly correlated) underlying, find how much to buy/sell.

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

This calls for *differentiation*, derivatives *in the sense of calculus*.

Formal Analysis

Let

$$\begin{aligned} -1 &= \text{number of units of the derivative product} \\ &\quad \text{(the minus because of the } \textit{short} \text{ position)}, \\ x &= \text{number of units of the underlying}, \\ s &= \text{today's stock price}, \\ t &= \text{today's date, so } (S_t = s). \end{aligned}$$

Value of the portfolio at time t is

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

A **delta hedge** is characterized by the property that

$$\frac{\partial V}{\partial s} = 0 \quad \text{(insensitivity to changes in } s\text{)}.$$

We obtain

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

Solve for x!

Result. We should have

$$\hat{x} = \frac{\partial F}{\partial s} \quad \text{(useful to compare to } h_t^S \text{ on slide 116)}.$$

shares of the underlying in the delta hedged portfolio.

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

$$\Delta = \frac{\partial F}{\partial s} \quad \text{(Recall: } F = \text{pricing function)}.$$

Result. For our goal we should thus have

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

shares of the underlying in the delta hedged portfolio.

Warning. The delta hedge must be rebalanced over time, in fact

$$\hat{x}_t = \Delta_t = \frac{\partial F}{\partial s}(t, S_t), \text{ time dependent.}$$

Application to Black Scholes

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

$$\Delta = N[d_1] = P(N(0,1) \leq d_1).$$

Remark. This is **not** a trivial result! (But it follows from slide 77, Black-Scholes case).

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

$$\Delta = N[d_1] - 1 = -P(N(0,1) > d_1) < 0.$$

Check signs and interpret!

Rebalanced Delta Hedge

- Sell one call option at time $t = 0$ at the B-S price F.
- Compute Δ and buy Δ shares. (Use the income from the sale of the option, and borrow money if necessary.)
- Wait one day (week, minute, second, or whatever is relevant). 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 and S) will match the value of the option better and better if you rebalance more frequently.

Some remarks

- Lack of perfection of the match 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

$$u_S = \frac{S \cdot \Delta}{F},$$
$$u_B = \frac{F - S \cdot \Delta}{F}.$$

Compare this to u_B on slide 116, with $\Delta = F_s(t, S_t)$.

Portfolio Delta

Assume that you have a portfolio consisting of derivatives

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

all **written on the same underlying** stock S, with

$$F_i(t, s) = \text{pricing function for } i\text{-th derivative } (S_t = s),$$

$$\Delta_i = \frac{\partial F_i}{\partial s},$$

$$h_i = \text{units of } i\text{:th derivative.}$$

The portfolio value is

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

and the *portfolio delta* becomes

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

Problem with delta-hedging

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 (see page 115).
- Thus you are sitting with the "wrong" value of delta (at a later time instant).

Moral.

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

Gamma

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

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

i.e.,

$$\Gamma = \frac{\partial^2 \Pi}{\partial s^2}.$$ (Here Π is the pricing function, often denoted $F(t, s)$.)

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

$$\Gamma = \frac{N'[d_1]}{S\sigma\sqrt{T-t}}$$ (Exercise, use the result on slide 128!)

Gamma Neutrality

Important fact. For a position in the underlying stock itself (recall $\Phi_S(x) = x$) we have

$$\Gamma = 0 \quad \text{(trivial, but useful)}.$$

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

$$\Gamma_\Pi = 0.$$

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

$\longrightarrow$ **End of lecture 3a** $\longleftarrow$

General procedure

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

$$x_F = \text{number of units of } F,$$
$$x_G = \text{number of units of } G.$$

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

The value of the entire hedged portfolio is

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

Repeat. Value of hedged portfolio:

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

We get the equations

$$\frac{\partial V}{\partial s} = 0 \,(\text{delta neutrality}),$$
$$\frac{\partial^2 V}{\partial s^2} = 0 \,(\text{gamma neutrality}),$$

i.e.,

$$\Delta_\Pi + x_F \Delta_F + x_G \Delta_G = 0,$$
$$\Gamma_\Pi + x_F \Gamma_F + x_G \Gamma_G = 0.$$

Solve for x_F and x_G! (This is a linear system, does it have a unique solution? – Yes, if G is sufficiently different from F.)

Particular Case

- In many practical cases the original portfolio Π is already delta neutral (this is experienced in practice).
- Then it is natural to use a derivative to obtain gamma-neutrality.
- But, this will destroy the delta-neutrality (for the new portfolio).
- Therefore we use the underlying stock (with zero gamma!) to delta hedge in the end (see next page).

Formal analysis I

Formally, the value V becomes

$$V = \Pi + x_F \cdot \underbrace{F}_{\text{additional derivative}} + \underbrace{x_S \cdot S}_{\text{stock needed for } \Delta\text{-neutrality}}.$$

The linear system becomes

$$\Delta_\Pi + x_F \Delta_F + x_S \Delta_S = 0,$$
$$\Gamma_\Pi + x_F \Gamma_F + x_S \Gamma_S = 0.$$

We have

$$\Delta_\Pi = 0 \,(\text{assumption was } \Pi \text{ is } \Delta\text{-neutral}),$$
$$\Delta_S = 1,$$
$$\Gamma_S = 0.$$

i.e.,

$$\Delta_\Pi + x_F \Delta_F + x_S = 0,$$
$$\Gamma_\Pi + x_F \Gamma_F = 0.$$

Formal analysis II

The solution to the linear system is

$$\begin{cases} x_F &= -\dfrac{\Gamma_\Pi}{\Gamma_F}, \\ x_S &= \dfrac{\Delta_F \Gamma_\Pi}{\Gamma_F} - \Delta_\Pi. \end{cases}$$

This portfolio is Delta- and Gamma-neutral.

Further Greeks

There are more sensitivity parameters, briefly mentioned below.

$$\Theta = \frac{\partial \Pi}{\partial t},$$
$$\mathcal{V} = \frac{\partial \Pi}{\partial \sigma} \text{ (N.B. } \mathcal{V} \text{ is pronounced "Vega"),}$$
$$\rho = \frac{\partial \Pi}{\partial r}.$$

Note the following.

- 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**, as σ is a model parameter.

The Martingale Approach
I: Mathematics

(Chapters 10-12)

a purely theoretical part of the lectures

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 (which you probably know already) 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. *They form your working knowledge, part of your toolkit.*

Contents

- Events and sigma-algebras
- Conditional expectations
- Measure changes
- The Martingale Representation Theorem
- The Girsanov Theorem

Events and sigma-algebras

Events and sigma-algebras (book: Section A.2)

Suppose we have a *sample space* Ω (a sample space contains all "relevant" outcomes of an experiment). If $A \subseteq \Omega$, one would like to compute $\mathbb{P}(A)$, the probability that A occurs.

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

For technical reasons the class $\mathcal{F}$ must be a **sigma-algebra** (also σ-algebra), which means that $\mathcal{F}$ is closed under the usual set theoretic operations like complements, **countable** intersections and **countable** unions. Subsets $A \in \mathcal{F}$ are called *events*.

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.

Some formalities

Definition. A σ-algebra on Ω is a collection $\mathcal{F}$ of subsets of Ω with the following properties.

1. It holds that $\emptyset \in \mathcal{F}$.
2. If $A \in \mathcal{F}$ then its complement $A^c \in \mathcal{F}$.
3. If $A_1, A_2, \ldots$ is a sequence of sets in $\mathcal{F}$, then also the union $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

It follows that *all* manipulations with sets (unions, complements, intersection, etc.) in $\mathcal{F}$ result in a set still belonging to $\mathcal{F}$, as long as there are at most countable operations with them.

Definition. A probability measure on $\mathcal{F}$ is a mapping $\mathbb{P} : \mathcal{F} \to [0, 1]$ with the following properties.

1. $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$.
2. If $A_1, A_2, \ldots$ is a sequence of *disjoint* sets in $\mathcal{F}$, then $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

A **Probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$.

Borel sets

Exercise. Let for all i in an arbitrary index set I, $\mathcal{F}_i$ be a σ-algebra on Ω. Then also $\cap_{i\in I}\mathcal{F}_i$ is a σ-algebra on Ω, smaller than any of the $\mathcal{F}_i$.

Note. If all $\mathcal{F}_i$ have a certain property, then $\cap_{i\in I}\mathcal{F}_i$ is the smallest σ-algebra having that property.

Definition. The **Borel algebra** $\mathcal{B}$ is the *smallest* sigma-algebra on $\mathbb{R}$ which *contains all intervals* (alternatively, contains all open sets, or contains all closed sets — these are topological concepts). 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 $\mathbb{R}$ 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 $\mathbb{R}$".

Random variables (Section B.1)

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

$$X : \Omega \to \mathbb{R},$$

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 $\mathbb{P}(X \in B)$. Instead 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 $\mathcal{F}$". So, it may happen that $X \in \mathcal{F}$, but $X \notin \mathcal{G}$. The information in $\mathcal{G}$ is then not sufficient to determine values of X.

$\longrightarrow$ **End of lecture 3b** $\longleftarrow$

Lecture 4

Conditional Expectation

Conditional Expectation (Section B.5), Indicator functions

Let $F \in \mathcal{F}$. We use the notation $\mathbf{1}_F$ (also I_F) for the function on Ω defined by

$$\mathbf{1}_F(\omega) = \begin{cases} 1 & \text{if } \omega \in F \\ 0 & \text{if } \omega \notin F. \end{cases}$$

In what follows, let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. This means $X \in \mathcal{F}$ and $\mathbb{E}|X| < \infty$.

Conditional Expectation, existence

Theorem. There exists a "unique" random variable $\hat{X} \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ such that

$$\mathbb{E}[\hat{X}\mathbf{1}_G] = \mathbb{E}[X\mathbf{1}_G] \text{ for all } G \in \mathcal{G}.$$

Remarks.

- Very important in the theorem is the σ-algebra $\mathcal{G}$ in $\hat{X} \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$, $\hat{X}$ is measurable on the smaller σ-algebra $\mathcal{G}$.
- $\hat{X}$ is "unique" means that another $\hat{X}' \in \mathcal{G}$ with the displayed property satisfies $\mathbb{P}(\hat{X} = \hat{X}') = 1$.
- Any $\hat{X}$ as in the theorem is called a version of the conditional expectation of X given G.
- Often used notations are $\hat{X} = \mathbb{E}[X \mid \mathcal{G}]$ and $\hat{X} = \mathbb{E}_{\mathcal{G}}[X]$ or $\hat{X} = \mathbb{E}_{\mathcal{G}}X$.

Main properties

Proposition 1: Assume that $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, and that $\mathcal{G} \subseteq \mathcal{F}$.

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

  $$\mathbb{E}[X \mid \mathcal{G}] \in \mathcal{G} \quad \text{(by definition)}.$$

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

  $$\mathbb{E}[XY \mid \mathcal{G}] = Y\mathbb{E}[X \mid \mathcal{G}].$$

  In particular we have $\mathbb{E}[Y \mid \mathcal{G}] = Y$.
- If $\mathcal{H} \subseteq \mathcal{G}$ then we have the "law of iterated expectations"

  $$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}] \quad \text{(analogy with iterated projections)}.$$

- In particular we have

  $$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]].$$

Conditional Expectation, characterization

The Proposition on slide 153 contains everything that we will need to know about conditional expectations within this course. But helpful for understanding may be the following characterization. The notation $Y \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ means that $Y \in \mathcal{G}$ and $\mathbb{E}[Y^2] < \infty$.

Proposition. If $\mathbb{E}[X^2] < \infty$, then for all $Y \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ one has

$$\mathbb{E}[(X - Y)^2] = \mathbb{E}[(X - \hat{X})^2] + \mathbb{E}[(\hat{X} - Y)^2],$$
$$\mathbb{E}[(X - Y)^2] \geq \mathbb{E}[(X - \hat{X})^2].$$

Remark. For the equality, think of 'Pythagoras' and ordinary projections; you may see in a picture $X - \hat{X} \perp \hat{X} - Y]$. The inequality says that $\hat{X}$ is the best approximation in $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ of X. It is a *best guess* of $X \in \mathcal{F}$, if one only posesses the information $\mathcal{G}$.

The proof (you try!) uses properties on slide 153.

Measure changes

Changing Measures (Section B.6)

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

$$L \geq 0 \text{ and } \mathbb{E}^{\mathbb{P}}[L] = 1,$$

where we write $\mathbb{E}^{\mathbb{P}}$ to denote that the expectation here is taken under the probability measure $\mathbb{P}$.

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

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[L \cdot I_A],$$

where the random variable I_A (also written as $\mathbf{1}_A$) is the indicator of A.

Towards a new probability measure

Recall the definition $\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[L \cdot I_A]$.

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

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

We also see that if $A \cap B = \emptyset$ then $I_{A \cup B} = I_A + I_B$ and hence

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

which immediately extends to finite disjoint unions.

Furthermore, it holds that (use the rule $Z = 0$, a.s. $\Rightarrow \mathbb{E}Z = 0$ applied to $Z = L \cdot I_A$)

$$\mathbb{P}(A) = 0 \quad \Rightarrow \quad \mathbb{Q}(A) = 0$$

We have thus more or less proved the following result.

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

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

then $\mathbb{Q}$ will be a *probability measure* (for which you need countable additivity, this holds true!) on $\mathcal{F}$ with the property that

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

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

Absolute Continuity

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

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

We write this as

$$\mathbb{Q} \ll \mathbb{P}.$$

If $\mathbb{Q} \ll \mathbb{P}$ and $\mathbb{P} \ll \mathbb{Q}$ then we say that $\mathbb{P}$ and $\mathbb{Q}$ are **equivalent** and write

$$\mathbb{Q} \sim \mathbb{P}.$$

Remark. $\mathbb{Q} \sim \mathbb{P}$ does *NOT* mean $\mathbb{Q} = \mathbb{P}$!

Equivalent measures

From the definition it immediately follows that $\mathbb{P}$ and $\mathbb{Q}$ are equivalent if and only if

$$\mathbb{P}(A) = 0 \quad \Leftrightarrow \quad \mathbb{Q}(A) = 0$$

or, equivalently (look at complements!),

$$\mathbb{P}(A) = 1 \quad \Leftrightarrow \quad \mathbb{Q}(A) = 1 \quad \text{(look at complements)}$$

(Other equivalent statements: $\mathbb{P}(A) > 0 \Leftrightarrow \mathbb{Q}(A) > 0$, $\mathbb{P}(A) < 1 \Leftrightarrow \mathbb{Q}(A) < 1$)

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

Simple examples:

- All non degenerate Gaussian distributions on $\mathbb{R}$ are equivalent.
- If $\mathbb{P}$ is Gaussian on $\mathbb{R}$ and $\mathbb{Q}$ is exponential then $\mathbb{Q} \ll \mathbb{P}$, but (explain why) *not* the other way around.

Absolute Continuity continued

We have seen that if we are given $\mathbb{P}$ and define $\mathbb{Q}$ by

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[L \cdot I_A] \qquad (*)$$

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

A natural question is now if **all** measures $\mathbb{Q} \ll \mathbb{P}$ are obtained in this way (that is, by formula $(*)$ for some L). The answer is yes, and the precise (non-trivial) result, the Radon-Nikodym theorem follows below. The proof of it is difficult and therefore omitted.

The Radon-Nikodym Theorem

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

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

(Here "unique" means that for another L' satisfying 1–4, one has $\mathbb{P}(L = L') = 1$.)

The random variable L is denoted as

$$L = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \quad \text{on } \mathcal{F}$$

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

Note on notation

$\mathbb{E}^{\mathbb{P}}[X]$ is often written as the *Lebesgue integral* $\int X \, \mathrm{d}\mathbb{P}$.

$$\text{Then} \quad \mathbb{E}^{\mathbb{P}}[L I_A] = \int L I_A \, \mathrm{d}\mathbb{P}. \tag{1}$$

$$\text{But also} \quad \mathbb{Q}(A) = \mathbb{E}^{\mathbb{Q}}[I_A] = \int I_A \, \mathrm{d}\mathbb{Q}. \tag{2}$$

Now *divide formally* (but mathematical *nonsensical*) by $\mathrm{d}\mathbb{P}$ (and multiply):

$$\mathrm{d}\mathbb{Q} = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \mathrm{d}\mathbb{P} = L \, \mathrm{d}\mathbb{P}$$

and we understand that (1) and (2) should be equal, and they are.

Additional facts

One can show that $\mathbb{Q}(L > 0) = 1$, equivalent to $\mathbb{Q}(L = 0) = 0$. To see the latter, one uses the simple computation

$$\mathbb{Q}(L = 0) = \mathbb{E}^{\mathbb{P}}[I_{\{L=0\}} L] = \mathbb{E}^{\mathbb{P}}[0] = 0.$$

But $\mathbb{P}(L = 0)$ can in principle be any number in $[0, 1]$.
Nevertheless we have the following result.

Proposition. Assume $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}_T$ with R-N derivative L. Then $\mathbb{P}(L = 0)$ iff $\mathbb{P} \ll \mathbb{Q}$ (and hence $\mathbb{P} \sim \mathbb{Q}$), and then $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}} = \frac{1}{L}$.

The proof is a nice exercise to do.

A simple example I

The Radon-Nikodym derivative L is intuitively the *local scale factor* between $\mathbb{P}$ and $\mathbb{Q}$, which will be explained now for a simple case.

If the sample space Ω is finite so $\Omega = \{\omega_1, \ldots, \omega_n\}$ then $\mathbb{P}$ is determined by the probabilities $p_1, \ldots, p_n$ where

$$p_i = \mathbb{P}(\omega_i) \quad i = 1, \ldots, n.$$

Now consider a measure $\mathbb{Q}$ with probabilities

$$q_i = \mathbb{Q}(\omega_i) \quad i = 1, \ldots, n.$$

If $\mathbb{Q} \ll \mathbb{P}$ this simply says that

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

and it is easy to see that the Radon-Nikodym derivative $L = \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$ (recall it is a function on Ω) is given by

$$L(\omega_i) = \frac{q_i}{p_i},\ i = 1, \ldots, n \text{ provided } p_i > 0.$$

A simple example II

Note that it follows that $q_i = p_i L(\omega_i)$, provided that $p_i > 0$. What if there is an i such that $p_i = 0$? By assumption, we also have $q_i = 0$. In this case we can now define the ratio q_i/p_i, or rather $L(\omega_i)$, arbitrarily, e.g. zero or 1345 if you wish, which now also leads to the same relation $q_i = L(\omega_i)p_i$.

z If $p_1, \ldots, p_n$ as well as $q_1, \ldots, q_n$ are all positive, then we see that $\mathbb{Q} \sim \mathbb{P}$, so we also have the R-N derivative $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}$. In fact

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}} = \frac{1}{L} = \left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)^{-1}$$

as could be expected.

Another example

Here $\mathbb{P}$ and $\mathbb{Q}$ are two possible distributions of a random variable with densities $f_{\mathbb{P}}$ and $f_{\mathbb{Q}}$. Assume $f_{\mathbb{P}} > 0$ and $f_{\mathbb{Q}} > 0$ on $\mathbb{R}$. Then $\mathbb{P} \sim \mathbb{Q}$ (seen as probability measures on $\mathcal{F} = \mathcal{B}$, the Borel sets of $\mathbb{R}$), and

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}(x) = \frac{f_{\mathbb{Q}}(x)}{f_{\mathbb{P}}(x)}, \qquad \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}(x) = \frac{f_{\mathbb{P}}(x)}{f_{\mathbb{Q}}(x)}.$$

Exercise. Compute explicitly for two normal densities $f_{\mathbb{P}}, f_{\mathbb{Q}}$ in the case where the variances are the same σ^2.

$\longrightarrow$ **End of lecture 4a** $\longleftarrow$

Computing expected values

A main use of Radon-Nikodym derivatives is for the computation of expected values, assuming that all of them exist.

Suppose therefore that $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}$ and that $X \in \mathcal{F}$. With $L = \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$ on $\mathcal{F}$ then have the following result.

Proposition 3: With notation as above we have

$$\mathbb{E}^{\mathbb{Q}}[X] = \mathbb{E}^{\mathbb{P}}[L \cdot X], \text{ also denoted}$$

$$\int X \, \mathrm{d}\mathbb{Q} = \int LX \, \mathrm{d}\mathbb{P} \quad \left(\text{write } L = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \text{ etc...} \right).$$

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

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

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 $\mathbb{P}$ and $\mathbb{Q}$ with $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}$ and with

$$L^{\mathcal{F}} = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \quad \text{on } \mathcal{F}.$$

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

$$\mathbb{E}^{\mathbb{Q}}[X \mid \mathcal{G}] = \frac{\mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} X \mid \mathcal{G}]}{\mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \mid \mathcal{G}]}.$$

Note the denominator, different from 1, as in the result $\mathbb{E}^{\mathbb{Q}}[X] = \mathbb{E}^{\mathbb{P}}[LX]$; see book Proposition B.41.

Dependence on the σ-algebra I

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

$$L^{\mathcal{F}} = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \quad \text{on } \mathcal{F}.$$

Now consider a smaller σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, and note that also $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{G}$! Our problem is to find the Radon-Nikodym (R-N) derivative

$$L^{\mathcal{G}} = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \quad \text{on } \mathcal{G}.$$

We recall (slide 162) that $L^{\mathcal{G}}$ is characterized as follows.

1. $\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[L^{\mathcal{G}} \cdot I_A], \quad \forall A \in \mathcal{G},$
2. $L^{\mathcal{G}} \geq 0,$
3. $\mathbb{E}^{\mathbb{P}}[L^{\mathcal{G}}] = 1,$
4. $L^{\mathcal{G}} \in \mathcal{G}$ (crucial property!).

Dependence on the σ-algebra II

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

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

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

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

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 $L^{\mathcal{F}} \in \mathcal{G}$ (so $L^{\mathcal{G}} \neq L^{\mathcal{F}}$ in general).

Dependence on the σ-algebra III

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

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \cdot I_A] = \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \cdot I_A \mid \mathcal{G}].\right]$$

Consider the term in brackets,

$$\mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \cdot I_A \mid \mathcal{G}].$$

Since $A \in \mathcal{G}$ we have, see property on slide 153 ,

$$\mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \cdot I_A \mid \mathcal{G}] = \mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \mid \mathcal{G}] I_A.$$

We thus obtain

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \mid \mathcal{G}] I_A].$$

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

Let us now define $L_{\mathcal{G}}$ by

$$L_{\mathcal{G}} = \mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \mid \mathcal{G}].$$

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

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

It is "easy" to see (use properties of slide 153 when needed) that also points 2-3 are satisfied so we have then proved the following result.

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

$$L^{\mathcal{G}} = \mathbb{E}^{\mathbb{P}}[L^{\mathcal{F}} \mid \mathcal{G}].$$

Here, the point was that we needed $L^{\mathcal{G}}$ to be $\mathcal{G}$-measurable.

The likelihood process on a filtered space

We now consider the case when we have a probability measure $\mathbb{P}$ 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 usual *interpretation* is 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 $\mathbb{Q}$, and that for some fixed T, we have $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}_T$. We define the random variable L_T by

$$L_T = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \quad \text{on } \mathcal{F}_T.$$

Since $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}_T$ we also have $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}_t$ for all $t \leq T$ (as on slide 170) and we define

$$L_t = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \quad \text{on } \mathcal{F}_t, \quad 0 \leq t \leq T.$$

For every t we have $L_t \in \mathcal{F}_t$ (part of R-N theorem), so L is an **adapted** process, known as the **likelihood process**.

Intermezzo, adapted processes, martingales

Definition. A process X is **adapted** (to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$) if for every $t \geq 0$, X_t is $\mathcal{F}_t$-measurable.

Definition. A process X is a $(\mathbb{P}, \mathcal{F}_t)$-martingale (in short: $\mathbb{P}$-martingale, or martingale) if

(i) it is adapted (to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$),

(ii) $\mathbb{E}[|X_t|] < \infty, \quad \forall t \geq 0,$

(iii) $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s, \quad \forall t \geq s \geq 0$ (martingale property).

Here $\mathbb{E} = \mathbb{E}^{\mathbb{P}}$, expectation under the probability measure $\mathbb{P}$.

The L process is a $\mathbb{P}$-martingale

We recall that

$$L_t = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \quad \text{on } \mathcal{F}_t, \quad 0 \leq t \leq T.$$

Since $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$ we can use Proposition 5 (slide 173) to deduce that

$$L_s = \mathbb{E}^{\mathbb{P}}[L_t \mid \mathcal{F}_s], \quad s \leq t \leq T$$

and we have thus proved the following result.

Proposition: Given the assumptions above, the likelihood process L is a $\mathbb{P}$-martingale.

Where are we heading, and why do we have to know this?

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

For this we need define the proper likelihood process L and, since L is a $\mathbb{P}$-martingale, we have the following natural questions:

- What does a martingale look like in a Wiener driven framework (like the Black-Scholes setting)?
- Suppose that we have a $\mathbb{P}$-Wiener process W and then change measure from $\mathbb{P}$ to $\mathbb{Q}$. What are the properties of this W under the *new* measure $\mathbb{Q}$?

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

The BS framework under $\mathbb{P}$ and $\mathbb{Q}$

Recall the BS framework, with

$$\mathrm{d}S = \mu S\,\mathrm{d}t + \sigma S\,\mathrm{d}W \text{ (under } \mathbb{P}\text{) and}$$
$$\mathrm{d}S = rS\,\mathrm{d}t + \sigma S\,\mathrm{d}W^{\mathbb{Q}} \text{ (under } \mathbb{Q}\text{).}$$

Note that these are two different SDEs for the *same* process S, due to two different Wiener processes, resulting from two different probability measures.

We shall see that $\mathbb{P} \sim \mathbb{Q}$ on $\mathcal{F}_T$ (and then also on $\mathcal{F}_t$, $t \leq T$).

$\longrightarrow$ **End of lecture 4b** $\longleftarrow$

Lecture 5

The Martingale Representation Theorem

(Book: Section 11.1)

A question

Suppose that we have a Wiener process W under the measure $\mathbb{P}$. We recall (from general Itô theory) that if h is adapted (and integrable enough: $\mathbb{E}\int_0^T h_s^2\,\mathrm{d}s < \infty$) and if the process X is defined by

$$X_t = x_0 + \int_0^t h_s\,\mathrm{d}W_s,$$

then X is 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

$$X_t = x_0 + \int_0^t h_s\,\mathrm{d}W_s,$$

for some adapted process h?

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

Answer: No, but...

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

$$X_t = \begin{cases} 0 & \text{for } 0 \leq t < 1, \\ Z & \text{for } t \geq 1. \end{cases}$$

where Z is a random variable, **independent** of W, with $\mathbb{E}[Z] = 0$.

X is then a martingale (you compute $\mathbb{E}[X_t \mid \mathcal{F}_s]$ for different $t > s$), but it should be clear that it cannot be written as

$$X_t = x_0 + \int_0^t h_s \, \mathrm{d}W_s$$

for any process h, see also next page.

Which filtration to take? X adapted requires $Z \in \mathcal{F}_t$ for $t \geq 1$. Simplest case: $\mathcal{F}_t = \{\emptyset, \Omega\}$ for $t < 1$, and $\mathcal{F}_t = \sigma(Z)$ for $t \geq 1$. But then W is not adapted to this filtration.

Intuition

The intuitive reason why we cannot write

$$X_t = x_0 + \int_0^t h_s \, \mathrm{d}W_s$$

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 "unrelated".

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

$$\mathcal{F}_t = \sigma\{W_s,\ s \leq t\},\ t \geq 0.$$

Note. The X of slide 182 is **NOT** adapted to this filtration, and the filtration there is not generated by W.

The Martingale Representation Theorem

Theorem. Let W be a $\mathbb{P}$-Wiener process and assume that the filtation is the **internal** one i.e.,

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

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

$$X_t = x + \int_0^t h_s \, \mathrm{d}W_s,$$

i.e., in differential notation,

$$\mathrm{d}X_t = h_t \, \mathrm{d}W_t.$$

Proof: Hard. This is a very deep result.

Crucial in this theorem is that X is adapted to this special filtration.

The theorem is not constructive

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

$$X_t = x + \int_0^t h_s \, \mathrm{d}W_s,$$

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

The Girsanov Theorem

Book: Sections 11.2, 11.3

Setup

Let W be a $\mathbb{P}$-Wiener process and fix a time horizon T. Suppose that we want to change measure from $\mathbb{P}$ to $\mathbb{Q}$ on $\mathcal{F}_T$, such that $\mathbb{Q} \ll \mathbb{P}$ or even $\mathbb{Q} \sim \mathbb{P}$. For this we need a $\mathbb{P}$-martingale L with $L_0 = 1$ to use as a likelihood process (see slide 176).

A natural way of constructing this is to choose a process g and then define L by

$$\begin{cases} \mathrm{d}L_t = g_t \, \mathrm{d}W_t & \text{(to get a martingale)}, \\ L_0 = 1 & \text{(and then } \mathbb{E}L_t = 1\text{)}. \end{cases}$$

However this definition does not guarantee that $L \geq 0$, so we make a useful adjustment. We choose a process ϕ and define (under integrability conditions) L by *(take above $g_t = L_t\phi_t$ for some ϕ_t)*

$$\begin{aligned} \mathrm{d}L_t &= L_t\phi_t \, \mathrm{d}W_t, \\ L_0 &= 1. \end{aligned}$$

Derivation with the Itô formula

The process L will be a martingale, under an integrability condition, and we easily obtain

$$L_t = e^{\int_0^t \phi_s \, \mathrm{d}W_s - \frac{1}{2}\int_0^t \phi_s^2 \, \mathrm{d}s}.$$

Indeed, apply Itô's formula to this L_t by noting

$$\begin{aligned} L_t &= f(X_t), \text{ with } f(x) = e^x, \\ X_t &= \int_0^t \phi_s \, \mathrm{d}W_s - \frac{1}{2}\int_0^t \phi_s^2 \, \mathrm{d}s, \text{ and} \\ (\mathrm{d}X_t)^2 &= \phi_t^2 \, \mathrm{d}t. \end{aligned}$$

Complete the computations to see that it results that L_t satisfies the SDE on the previous slide.

Measure change, from $\mathbb{P}$ to $\mathbb{Q}$

By the exponential form of L_t we are guaranteed that $L \geq 0$, even $L > 0$. We now change measure from $\mathbb{P}$ to $\mathbb{Q}$ on $\mathcal{F}_T$ by setting

$$\mathrm{d}\mathbb{Q} = L_T \, \mathrm{d}\mathbb{P}.$$

By strict positivity $\mathbb{P}(L_T = 0) = 0$, and hence we even have $\mathbb{Q} \sim \mathbb{P}$ on $\mathcal{F}_T$, see slide 164. Note that we also have, see slide 176, $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}_t$ (even $\mathbb{Q} \sim \mathbb{P}$), and thus

$$\mathrm{d}\mathbb{Q} = L_t \, \mathrm{d}\mathbb{P} \text{ on } \mathcal{F}_t, \ 0 \leq t \leq T,$$
$$L_t = \mathbb{E}^{\mathbb{P}}[L_T \mid \mathcal{F}_t] \, .$$

A main problem is to find out what the properties of W are, under the new measure $\mathbb{Q}$. This problem is resolved by the **Girsanov Theorem**.

The Girsanov Theorem

Let W be a $\mathbb{P}$-Wiener process. Fix a time horizon T.

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

$$\begin{cases} \mathrm{d}L_t = L_t \phi_t \, \mathrm{d}W_t, \quad (*) \\ L_0 = 1. \end{cases}$$

Assume that $\mathbb{E}^{\mathbb{P}}[L_T] = 1$ (this is the case if L is a $\mathbb{P}$-martingale), and define a new measure $\mathbb{Q}$ on $\mathcal{F}_T$ by

$$\mathrm{d}\mathbb{Q} = L_T \mathrm{d}\mathbb{P} \text{ on } \mathcal{F}_T.$$

Then $\mathbb{Q} \ll \mathbb{P}$ on $\mathcal{F}_T$ (and hence also on all $\mathcal{F}_t$ for $t \leq T$) and the process $W^{\mathbb{Q}}$, defined by

$$W_t^{\mathbb{Q}} = W_t - \int_0^t \phi_s \, \mathrm{d}s$$

is $\mathbb{Q}$-Wiener. We can also write this as

$$\mathrm{d}W_t = \phi_t \, \mathrm{d}t + \mathrm{d}W_t^{\mathbb{Q}}.$$

Remarks

- We have seen before that in Girsanov's theorem also

$$\mathrm{d}\mathbb{Q} = L_t \, \mathrm{d}\mathbb{P} \text{ on } \mathcal{F}_t,\ 0 \leq t \leq T,$$

  and that

$$L_t = \mathbb{E}^{\mathbb{P}}[L_T \mid \mathcal{F}_t] \,.$$

- We state without proof that in fact, with (*), L is a martingale ($t \leq T$) if and only if $\mathbb{E}^{\mathbb{P}}[L_T] = 1$.

Changing the drift in an SDE (Section 11.5)

Suppose that we have a process X with $\mathbb{P}$ dynamics

$$(**) \qquad \mathrm{d}X_t = \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t,$$

where μ and σ are adapted processes and W is $\mathbb{P}$-Wiener.

We now do a Girsanov Transformation as above, and the question is what the $\mathbb{Q}$-dynamics of X look like.

From the Girsanov Theorem, see slide 190, we have

$$\mathrm{d}W_t = \phi_t \,\mathrm{d}t + \,\mathrm{d}W_t^{\mathbb{Q}},$$

and substituting this into the $\mathbb{P}$-dynamics (**), we obtain the $\mathbb{Q}$ dynamics as

$$\mathrm{d}X_t = (\mu_t + \sigma_t\phi_t)\,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t^{\mathbb{Q}}.$$

Moral. The drift changes but the diffusion is unaffected, meaning that we keep on having the *same* σ_t in front of the new Brownian motion $W^{\mathbb{Q}}$.

The Converse Girsanov Theorem

Let W be a $\mathbb{P}$-Wiener process. Fix a time horizon T.

Theorem. Assume the following.

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

$$L_t = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \text{ on } \mathcal{F}_t, \;\; 0 \leq t \leq T.$$

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

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

Then there exists a process ϕ such that

$$\begin{cases} \mathrm{d}L_t = L_t \phi_t \, \mathrm{d}W_t \\ L_0 = 1 \end{cases}$$

Note. L is a $\mathbb{P}$-martingale, see (slide 176).

$\longrightarrow$ **End of Lecture 5a** $\longleftarrow$

The Martingale Approach II: Pricing and Hedging

(Chapters 10-12)

Financial Markets, a recap

Price Processes.

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

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

$$\begin{aligned} \mathrm{d}S_t^0 &= rS_t^0\,\mathrm{d}t, \\ \mathrm{d}S_t^1 &= \alpha S_t^1\,\mathrm{d}t + \sigma S_t^1\,\mathrm{d}W_t. \end{aligned}$$

Portfolio. $h_t = \left[h_t^0, \ldots, h_t^N\right]$, where
$h_t^i =$ number of shares of asset i at time t.

Value process:

$$V_t^h = \sum_{i=0}^{N} h_t^i S_t^i =: h_t S_t \text{ (inner product notation).}$$

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

$$\mathrm{d}V_t^h = \sum_{i=0}^{N} h_t^i \,\mathrm{d}S_t^i.$$

Major insight *(from general theory)*. If the price process S is a **martingale**, and if h is **self-financing**, then V^h is a **martingale**. *(This needs technical assumptions).*

Remark. This simple observation is in fact the basis of the theory that follows. But first a digression that helps to understand this.

Digression in discrete time

Let the price and value processes (S and V, both adapted) evolve in discrete time. Assume that h is another adapted process, and the relation

$$\Delta V_t = h_{t-1}\,\Delta X_t.$$

Then, from properties of conditional expectation (slide 153)

$$\mathbb{E}[\Delta V_t \mid \mathcal{F}_{t-1}] = \mathbb{E}[h_{t-1}\Delta X_t \mid \mathcal{F}_{t-1}] = h_{t-1}\,\mathbb{E}[\Delta X_t \mid \mathcal{F}_{t-1}].$$

If X is a martingale, then $\mathbb{E}[\Delta X_t \mid \mathcal{F}_{t-1}] = 0$ and then also $\mathbb{E}[\Delta V_t \mid \mathcal{F}_{t-1}] = 0$, meaning that V is a martingale as well.

The property that V becomes a martingale when X is a martingale carries over to continuous time, when $\mathrm{d}V_t = h_t\,\mathrm{d}X_t$ for an adapted process h (subject to integrability conditions).

Arbitrage

A self financing portfolio h (with $V = V^h$) is an **arbitrage strategy** (weaker definition than the earlier one on slide 66) on $[0, T]$ if

$$V_0^h = 0,$$
$$\mathbb{P}\Big(V_T^h \geq 0\Big) = 1,$$
$$\mathbb{P}\Big(V_T^h > 0\Big) > 0.$$

Remark. The last two requirements imply $\mathbb{E}^{\mathbb{P}}[V_T^h] > 0$.

Main Question. Find conditions for the market to be free of arbitrage.

Basic Idea

Remark. A useful fact is the following. If the market is free of arbitrage, and h is a self-financing portfolio with $\mathbb{P}(V_T^h \geq 0) = 1$, then $\mathbb{P}(V_T^h > 0) = 0$, equivalently $\mathbb{P}(V_T^h = 0) = 1$.

Assume there are **no** arbitrage possibilities.

- Then there should exist some **fair pricing system** on the market.
- Formalization. The martingale property holds for some process under some measure.
- Basic Idea. If $S^0, \ldots, S^N$ are martingales (*but does this happen?*), then the market is free of arbitrage.

First attempt to apply the basic idea

Proposition. If all price processes $S_t^0, \ldots, S_t^N$ are $\mathbb{P}$-martingales *(which is not realistic)*, then the market is free of arbitrage.

Proof. Assume that h is an arbitrage strategy. Since

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

V^h is a $\mathbb{P}$-martingale too, so (because then expectations remain the same), $\mathbb{E}^{\mathbb{P}}[V_t^h] = \mathbb{E}^{\mathbb{P}}[V_0^h]$. Hence

$$V_0^h = \mathbb{E}^{\mathbb{P}}[V_T^h] > 0 \quad \text{(see slide 198)}.$$

But this contradicts $V_0^h = 0$, so arbitrage cannot exist.

The argumentation is correct, but useless (see next slide).
Yet, as we'll see, there is a point in the argumentation.

The stock price is not a $\mathbb{P}$-martingale (in general)

Example. Black-Scholes dynamics under $\mathbb{P}$:

$$\begin{aligned} \mathrm{d}S_t &= \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \\ \mathrm{d}B_t &= rB_t \,\mathrm{d}t. \end{aligned}$$

For S and B to be martingales, one would have to assume that $\alpha = r = 0$. This is *not* realistic.

We now try to improve on this result.

Normalization: S_0 as Numéraire

Choose one asset, say S^0 as **numéraire** (accounting unit) and normalize all prices with this asset.

The **normalized price vector** is given by

$$Z_t = \frac{S_t}{S_t^0}, \text{ with elements } Z_t^i = \frac{S_t^i}{S_t^0}, \ i = 0, \ldots, N.$$

In particular, $Z_t^0 = 1$, $\forall t \leq T$, certainly a martingale!

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

$$V_t^Z = \frac{V_t}{S_t^0} = \sum_{i=0}^{N} h_t^i Z_t^i =: h_t Z_t.$$

Insight. The arbitrage and self financing concepts are independent of the accounting unit (numéraire).

Invariance of Numéraire

Proposition. One can show (see the book) that

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

$$\mathrm{d}V_t^h = h_t\,\mathrm{d}S_t \iff \mathrm{d}V_t^Z = h_t\,\mathrm{d}Z_t,$$

so we can just talk of self-financing.

This leads to the **Insight**,

- If h self-financing then

$$\mathrm{d}V_t^Z = \sum_1^N h_t^i\,\mathrm{d}Z_t^i \quad \text{(note: we don't need } i = 0 \text{ here)}.$$

- Thus, if the **normalized** price process Z is a $\mathbb{P}$-martingale, then V^Z is a martingale, *same argument as before*.

Invariance of the self financing property

Proposition. A portfolio h is self financing in the original market $S^0, \ldots, S^N$ if and only if it is self financing in the normalized market $Z^0, \ldots, Z^N$.
That is:

$$\mathrm{d}V_t^h = \sum_{i=0}^{N} h_t^i \, \mathrm{d}S_t^i \quad \Longleftrightarrow \quad \mathrm{d}V_t^Z = \sum_{i=0}^{N} h_t^i \, \mathrm{d}Z_t^i.$$

Proof. Use

$$V_t^Z = \frac{V_t^h}{S_t^0}$$

and "apply Itô" to the ratio. Recall that S^0 is not random, the "ordinary" quotient rule essentially applies. Details omitted.

Second Attempt

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

True result, but still fairly useless (argument as on slide 200).

Example. (Black-Scholes)

$$\begin{aligned} \mathrm{d}S_t &= \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \\ \mathrm{d}B_t &= rB_t \,\mathrm{d}t. \end{aligned}$$

Use the quotient rule for differentiation of $Z_t^1 = \frac{S_t}{B_t}$ to get (check!)

$$\begin{aligned} \mathrm{d}Z_t^1 &= (\alpha - r)Z_t^1 \,\mathrm{d}t + \sigma Z_t^1 \,\mathrm{d}W_t, \\ \mathrm{d}Z_t^0 &= 0 \,\mathrm{d}t. \end{aligned}$$

We would have to assume "risk-neutrality", i.e. that $\alpha = r$, to have Z^1 is a $\mathbb{P}$-martingale. But, in principle, $\alpha \neq r$. Later we'll see what happens under the risk neutral measure $\mathbb{Q}$.

Arbitrage revisited

Recall, as noted on slide 160, that $\mathbb{P} \sim \mathbb{Q}$ iff one, and then each, of the following holds true,

$$\mathbb{P}(A) = 0 \iff \mathbb{Q}(A) = 0, \textit{or}$$
$$\mathbb{P}(A) = 1 \iff \mathbb{Q}(A) = 1, \textit{or}$$
$$\mathbb{P}(A) > 0 \iff \mathbb{Q}(A) > 0.$$

Recall from slide 198 that h is an arbitrage if (We write V for V^h)

- h is self financing
- $V_0 = 0$.
- $V_T \geq 0$, $\mathbb{P}$-a.s. and $\mathbb{P}(V_T > 0) > 0$.

The last item is equivalent to

$$V_T \geq 0, \mathbb{Q} - a.s. \text{ and } \mathbb{Q}(V_T > 0) > 0.$$

This leads to the **major insight** that arbitrage is **invariant** under an **equivalent change of measure**!

.

(Equivalent) Martingale Measures

Consider the market $S^0, \ldots, S^N$.

Definition. A probability measure $\mathbb{Q} \sim \mathbb{P}$ is called an **Equivalent Martingale Measure** (EMM) for the market $S^0, \ldots, S^N$, with numéraire S^0, if the normalized price processes

$$Z^0, \ldots, Z^N, \; Z^i_t = \frac{S^i_t}{S^0_t},$$

are $\mathbb{Q}$-martingales.

Since $Z^0 \equiv 1$ is always a martingale, the condition in the display reduces to $Z^1, \ldots, Z^N$ are $\mathbb{Q}$-martingales.

Other names for an EMM are *Risk Neutral Measure*, *Pricing Measure*.

First Fundamental Theorem of Asset Pricing

Theorem [FTAP 1]. The market is free of arbitrage if and only if there exists a martingale measure $\mathbb{Q}$.

Remark. An EMM is in general not unique. Uniqueness is related to completeness, as we shall see later.

Proof of FTAP 1. The implication $\Rightarrow$ is very hard, but the implication $\Leftarrow$ is relatively easy. See below.

This theorem was already announced on slide 89.

Comments

- We shall see that it is not hard (but you still have to know quite some math) to prove that existence of an EMM implies no arbitrage.
- The other implication is technically very hard, even in a usually much simpler discrete time setting.
- For discrete time and finite sample space Ω the hard part follows "easily" from the *separation theorem* for convex sets that uses *separating hyperplanes* (you don't have to know this).
- For discrete time and more general sample spaces we need the *Hahn-Banach Theorem* (formulated as an infinite dimensional version of the separation theorem) from Functional Analysis,
- For continuous time the proof becomes technically very demanding, mainly due to topological problems in function spaces. See the textbook.

$\longrightarrow$ **End of Lecture 5b** $\longleftarrow$

Lecture 6

Existence of EMM implies no arbitrage

Proof. *This is a proof by contradiction.* Recall the statement of FTAP 1 ("NA $\iff$ EMM"), see slide 208.

Assume that there exists an EMM $\mathbb{Q}$, and assume existence of an arbitrage opportunity, $\mathbb{P}(V_T \geq 0) = 1$ and $\mathbb{P}(V_T > 0) > 0$.
Then, since $\mathbb{P} \sim \mathbb{Q}$ we also have $\mathbb{Q}(V_T \geq 0) = 1$ and $\mathbb{Q}(V_T > 0) > 0$.

It then follows (result from probability theory) that $\mathbb{E}^{\mathbb{Q}}(V_T) > 0$!
Moreover, all these statements are also true for V_T^Z instead of V_T, as one normalizes by a strictly positive process Z^0.

Proof continued

Recall

$$\mathrm{d}V_t^Z = \sum_1^N h_t^i \, \mathrm{d}Z_t^i.$$

Consider the following chain of implications.

$$\mathbb{Q} \text{ is a martingale measure}$$
$$\Downarrow$$
$$V^Z \text{ is a } \mathbb{Q}\text{-martingale (\textit{Itô theory, as } } Z \text{ \textit{is a } } \mathbb{Q}\text{-\textit{martingale by definition}})$$
$$\Downarrow$$
$$V_0 \underset{S_0^0=1}{=} V_0^Z \underset{\text{martingale}}{=} \underset{\text{previous page}}{\mathbb{E}^{\mathbb{Q}}[V_T^Z]} > 0.$$

This contradicts $V_0 = 0$ in the definition of an arbitrage opportunity. Conclusion: there is no arbitrage when an EMM exists.

Choice of Numéraire

The numéraire 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

$$\mathrm{d}B_t = r_t B_t \,\mathrm{d}t,\ B_0 = 1.$$

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

$$B_t = e^{\int_0^t r_s \,\mathrm{d}s}$$

This generalizes $B_t = e^{rt}$, $\mathrm{d}B_t = rB_t \,\mathrm{d}t$, which we assume in most examples.

Example: The Black-Scholes Model

Recall

$$\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t,$$
$$\mathrm{d}B_t = rB_t \,\mathrm{d}t.$$

Look for a martingale measure. We set $Z = S/B$, i.e. $Z_t = \frac{S_t}{B_t} = S_t e^{-rt}$.
Standard calculus (no Itô correction terms are needed as B_t contains no Brownian term) applies to differentiate the product $S_t e^{-rt}$. One obtains

$$\mathrm{d}Z_t = Z_t(\alpha - r)\,\mathrm{d}t + Z_t\sigma\,\mathrm{d}W_t. \qquad (6a)$$

Next do a Girsanov transformation on $[0, T]$, L_t is the likelihood,

$$\mathrm{d}L_t = L_t\phi_t\,\mathrm{d}W_t,\ L_0 = 1,$$
$$\mathrm{d}\mathbb{Q} = L_T\,\mathrm{d}\mathbb{P},\ \text{on } \mathcal{F}_T.$$

Example continued

Apply the Girsanov theorem (see slide 190):

$$\mathrm{d}W_t = \phi_t \,\mathrm{d}t + \mathrm{d}W_t^{\mathbb{Q}}, \qquad (6b)$$

where $W^{\mathbb{Q}}$ is a $\mathbb{Q}$-Wiener process. *(whereas W is $\mathbb{P}$-Wiener)*

Insert (6b) into (6a) to obtain that the $\mathbb{Q}$-dynamics for Z are given by

$$\mathrm{d}Z_t = Z_t(\alpha - r + \sigma\phi_t)\,\mathrm{d}t + Z_t\sigma\,\mathrm{d}W_t^{\mathbb{Q}}.$$

There is a unique(!) martingale measure $\mathbb{Q}$, with Girsanov kernel given by $\phi_t = \frac{r-\alpha}{\sigma}$, and then

$$\mathrm{d}Z_t = Z_t\sigma\ \mathrm{d}W_t^{\mathbb{Q}}, \text{ a } \mathbb{Q}\text{-martingale!}$$

Conclusion

The Black-Scholes model is *free of arbitrage*, as follows from slide 212 since we have now shown that $\mathbb{Q}$ is an EMM: Z is a martingale under $\mathbb{Q}$. We have also seen there is a *unique* EMM.

To find the $\mathbb{Q}$-dynamics of S, we *insert (6b) into equation for* $\mathrm{d}S_t$ to obtain the familiar SDE

$$\mathrm{d}S_t = rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t^{\mathbb{Q}}.$$

Pricing revisited

We consider a market $B_t, S_t^1, \ldots, S_t^N$ *(many risky assets)*.

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

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

Recall, at $t = T$ the amount X is paid to the holder of the claim.

Well known **Example** (European Call Option),

$$X = (S_T - K)^+ = \max\{S_T - K, 0\}.$$

Let X be a contingent T-claim.

Classical Problem. How do we find an arbitrage free price process $\Pi_t[X]$ for X?

New Approach: use the change of measure framework.

Solution to the pricing problem

Definition. The *extended* market *(with extra asset* $\Pi_t[X]$*)* is

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

Definition. $\Pi_t[X]$ is an arbitrage free price process if the extended market is arbitrage free.

To find the arbitrage free price $\Pi_t[X]$, we need the extended market to be arbitrage free. So there must exist a martingale measure $\mathbb{Q}$ for it (FTAP 1). In particular

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

is a $\mathbb{Q}$-martingale, i.e., *it has the martingale property under* $\mathbb{Q}$.

Risk neutral pricing formula

That martingale property says

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

and so

$$\Pi_t[X] = B_t\,\mathbb{E}^{\mathbb{Q}}\left[\frac{\Pi_T[X]}{B_T}\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{B_t}{B_T}X\middle|\mathcal{F}_t\right]. \qquad (6c)$$

Since we obviously (why?) have $\Pi_T[X] = X$, we have proved the main pricing formula in the following theorem *(rewrite (6c)).* .

Theorem (Risk neutral valuation). For a T-claim X, and if $\mathrm{d}B_t = r_t B_t\,\mathrm{d}t$, the arbitrage free price is given by the formula

$$\Pi_t[X] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s\,\mathrm{d}s}X\middle|\mathcal{F}_t\right].$$

Special case, constant interest rate, and final note

Note. If $r_t = r$, then

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

a formula which we have already encountered in a slightly different form on slide 83.

Note on the underlying new method. We have *not* used the Feynman-Kac formula to arrive at the same $\Pi_t(X)$, but used a **martingale argument**. Moreover, we didn't impose that X was a simple claim of the special form $X = \Phi(S_T)$. But simple claims are still important.

Example: The Black-Scholes Model

Consider the $\mathbb{Q}$-dynamics of S *(same as before)*,

$$\mathrm{d}S_t = rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t^{\mathbb{Q}}.$$

Note. S is a Markov process under $\mathbb{Q}$!

Consider the *simple* claim

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

The pricing formula on slide 219 and the Markov property give

$$\begin{aligned}\Pi_t[X] &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\Phi(S_T) \mid \mathcal{F}_t] \\ &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_T) \mid S_t\right].\end{aligned}$$

Then, "Kolmogorov" implies that $\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_T) \mid S_t\right]$ and hence also $\Pi_t(X)$ is a function of (t and) S_t, write

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

Example continued

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

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

To derive this, use Feynman-Kac and the $\mathbb{Q}$-dynamics of S.

Problem

Recall the valuation formula

$$\Pi_t[X] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s \, \mathrm{d}s} X \middle| \mathcal{F}_t \right],$$

which (in principle) depends on $\mathbb{Q}$.

Question. What if there are several different martingale measures $\mathbb{Q}$? An EMM is in general not unique. Are there then different prices $\Pi_t[X]$, or not?

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

Hedging

Recall slide 110 for a first treatment.

Definition. A portfolio h is a **hedge** against X ("replicates X") if

- h is self financing,
- $V_T = X$, $\mathbb{P}$–a.s.: $\mathbb{P}(V_T = X) = 1$.

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

Pricing Formula. If h replicates X, then a natural way of pricing X in an arbitrage free market is, see slide 111 for a justification,

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

so $\Pi_t[X]$ is given by the value V_t^h of the hedge portfolio.

Central question is now:

When can we hedge?

A relation between economics and mathematics

Existence of hedge

$$\Updownarrow$$

Existence of stochastic integral

(Martingale Representation Theorem)

Hedgebable claim

Fix a T-claim X. If h is a hedge for X then $V_T = X$ and

- $V_T^Z = \frac{X}{B_T}$ (*recall the normalized prices* $Z_t^i = \frac{S_t^i}{B_t}$, *which are* $\mathbb{Q}$*-martingales)*,
- h is self financing, i.e.,

$$\mathrm{d}V_t^Z = \sum_1^K h_t^i \, \mathrm{d}Z_t^i \quad \text{see slide 203.}$$

Thus V^Z is a $\mathbb{Q}$-martingale, $V_t^Z = \mathbb{E}^{\mathbb{Q}}\left[V_T^Z \,|\, \mathcal{F}_t\right]$, and thus

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

if X can be hedged by h. But what if X cannot be hedged?

$\longrightarrow$ **End of Lecture 6a** $\longleftarrow$

Replication lemma

Here we reverse the argument that is used on slide 226.

Lemma. Fix T-claim X. Define the martingale M by

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

and note that then $M_T B_T = X$.

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

$$M_t = x + \sum_{i=1}^{N} \int_0^t h_s^i \, \mathrm{d}Z_s^i.$$

Then X can be replicated.

Proof I

We guess that (for replication), using that $B_t \equiv 1$ is the normalized bank account, for some h^B one has

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

Note that thus $X = V_T^Z B_T = h_T^B B_T + \sum_{i=1}^{N} h_t^i S_t^i$.

Define h^B (recall that $h^1, \ldots, h^N$ are given) by

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

Proof II

We have $M_t = V_t^Z$, and we get (by the assumption in the lemma)

$$\mathrm{d}V_t^Z = \mathrm{d}M_t = \sum_{i=1}^{N} h_t^i \, \mathrm{d}Z_t^i,$$

so the portfolio $(h^B, h^1, \ldots, h^N$ is self financing. Furthermore,

$$V_T^Z = M_T = \frac{X}{B_T},$$

which yields the hedge, as now $V_T = X$.

Second Fundamental Theorem (FTAP 2), statement

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

Theorem (FTAP 2). The market is complete **iff** there is a unique EMM $\mathbb{Q}$.

Second Fundamental Theorem (FTAP 2), proof

It is relatively easy to show that if the market is complete, then $\mathbb{Q}$ must be unique. The other implication is very hard to prove, and we will not do it, as it basically relies on duality arguments from functional analysis. Here is the "easy" part

Proof. For all $A \in \mathcal{F}_T$, the claim $\mathbf{1}_A B_T$ can be hedged, and hence has a *unique* price (law of one price). For any EMM $\mathbb{Q}$ one has

$$\Pi_t(\mathbf{1}_A B_T) = \frac{\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_A|\mathcal{F}_t]}{B_t}.$$

At $t = 0$ one gets $\Pi_0(\mathbf{1}_A B_T) = \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_A|\mathcal{F}_0] = \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_A]$, as $\mathcal{F}_0$ is the trivial σ-algebra. So, we read $\Pi_0(\mathbf{1}_A B_T) = \mathbb{Q}(A)$, and $\mathbb{Q}(A)$ is thus unique, true for every $A \in \mathcal{F}_T$. Thus $\mathbb{Q}$ is unique.

Black-Scholes Model

One has the $\mathbb{Q}$-dynamics (recall $Z_t = \frac{S_t}{B_t}$ and slide 215),

$$\mathrm{d}S_t = rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t^{\mathbb{Q}},$$
$$\mathrm{d}Z_t = Z_t\sigma\,\mathrm{d}W_t^{\mathbb{Q}}.$$

Note that $Z_t = S_0\exp(\sigma W_t^{\mathbb{Q}} - \frac{1}{2}\sigma^2 t) > 0$.

Let X be an arbitrary claim and consider the martingale

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

The representation theorem for Wiener processes, slide 184, gives the existence of a process g (given that the filtration is *also* generated by $W^{\mathbb{Q}}$, a technical issue) such that

$$M_t = M_0 + \int_0^t g_s\,\mathrm{d}W_s^{\mathbb{Q}}.$$

Thus (use $Z_t > 0$)

$$M_t = M_0 + \int_0^t h_s^1\,\mathrm{d}Z_s, \text{ with } h_t^1 = \frac{g_t}{\sigma Z_t}(\text{ recall } Z_t > 0).$$

Result. From the Lemma on slide 227 it follows that X can be replicated using the portfolio defined by

$$h_t^1 = \frac{g_t}{\sigma Z_t},$$
$$h_t^B = M_t - h_t^1 Z_t.$$

Moral. The Black Scholes model is complete.

Note. We didn't need anymore, like in earlier lectures, that X is of the special form $X = \Phi(S_T)$, but see also the next slides for special cases.

Remaining question. What is g_t?

Special Case: Simple Claims

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

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

As S is Markov, one has

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-rT}\Phi(S_T) \mid \mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}\Phi(S_T) \mid S_t\right].$$

The Kolmogorov backward equation yields $M_t = f(t, S_t)$, with f satisfying, parallel to the result for F on slide 222,

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

Exercise. Show that the above PDE follows from the one for F.

Special case continued

Apply (you do!) the Itô formula to $M_t = f(t, S_t)$ and use that f satisfies the PDE to get

$$dM_t = \sigma S_t \frac{\partial f}{\partial s} \, \mathrm{d}W_t^{\mathbb{Q}},$$

so, the answer to the question on slide 233 is

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

This gives us thus in this case explicitly the "abstract" g of the MRT, and it also gives us the replicating portfolio h,

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

Interpretation. $f(t, S_t) = V_t^Z$, the normalized price process of X.

Special case with non-normalized prices, a converse path

Let f as on the previous slides and define $F(t,s)$ by

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

so $F(t,S_t) = V_t$. Then from the expressions for h_t^B and h_t^1 on previous slide and $\frac{\partial F}{\partial s} = B_t \frac{\partial f}{\partial s}$ one gets

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

where F solves the Black-Scholes boundary problem

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

The PDE for F

Exercise. Find the PDE for F, for which you depart from the PDE for f on the slide 234 and the relations

$$\frac{\partial F}{\partial t} = rB_t f + B_t \frac{\partial f}{\partial t},$$
$$\frac{\partial F}{\partial s} = B_t \frac{\partial f}{\partial s}.$$

Main Results, a summary

- The market is arbitrage free $\Leftrightarrow$ There exists an equivalent martingale measure $\mathbb{Q}$ (FTAP 1).
- The market is complete $\Leftrightarrow$ $\mathbb{Q}$ is unique (FTAP 2).
- Every T-claim X must be priced by the formula

$$\Pi_t[X] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s \, \mathrm{d}s} X \middle| \mathcal{F}_t \right]$$

  for some choice of $\mathbb{Q}$.
- In a non-complete market, different choices of $\mathbb{Q}$ will produce *different* prices for non-hedgeable X.
- For a hedgeable claim X (recall $\Pi_t[X] = V_t$, the value process of the hedge portfolio), all choices of $\mathbb{Q}$ will produce the *same* price for X,

$$\Pi_t[X] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s \, \mathrm{d}s} \times X \middle| \mathcal{F}_t \right] = V_t.$$

Completeness and No Arbitrage, Rule of Thumb

We repeat the contents of slide 118.

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

Count the number of risky assets, and the number of random sources. Let

$R =$ number of random sources,

$N =$ number of risky assets.

Rule of thumb

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. Recall the $\mathbb{P}$-dynamics

$$\begin{aligned} \mathrm{d}S_t &= \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \\ \mathrm{d}B_t &= rB_t \,\mathrm{d}t. \end{aligned}$$

For Black-Scholes we have $N = R = 1$. Thus the Black-Scholes model is arbitrage free and complete.

Stochastic Discount Factors, pricing formulas under $\mathbb{P}$

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

$$D_t = e^{-\int_0^t r_s \, \mathrm{d}s} L_t = \frac{L_t}{B_t},$$

where

$$L_t = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \text{ on } \mathcal{F}_t.$$

There is thus a one-to-one correspondence between EMMs and SDFs, as an EMM $\mathbb{Q}$ is related to $\mathbb{P}$ by the process L.

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

Proposition. With notation as above we have a pricing formula under the measure $\mathbb{P}$ (but note that $\mathbb{Q}$ is "hidden" in D_t),

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

Proof

Proof. We use the Abstract Bayes' formula. Start from

$$Y := \frac{\Pi_t(X)}{B_t} = \mathbb{E}^{\mathbb{Q}}[\frac{X}{B_T} \mid \mathcal{F}_t].$$

From slide 169 we know that

$$\begin{aligned} Y &= \frac{\mathbb{E}^{\mathbb{P}}[\frac{X}{B_T} L_T \mid \mathcal{F}_t]}{L_t} \\ &= \frac{\mathbb{E}^{\mathbb{P}}[X D_T \mid \mathcal{F}_t]}{D_t B_t}. \end{aligned}$$

Hence, multiply by B_t to get

$$\Pi_t(X) = \frac{\mathbb{E}^{\mathbb{P}}[X D_T \mid \mathcal{F}_t]}{D_t}.$$

Martingale Property of $S \cdot D$

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

$$S_t D_t$$

is a $\mathbb{P}$-martingale.

Proof. Bayes' formula again, with the same trick. We know from the definition of EMM, $\mathbb{E}^{\mathbb{Q}}[\frac{S_T}{B_T} \mid \mathcal{F}_t] = \frac{S_t}{B_t}$, which we convert by the Abstract Bayes formula into

$$\frac{\mathbb{E}^{\mathbb{P}}[\frac{S_T}{B_T} L_T \mid \mathcal{F}_t]}{L_t} = \frac{S_t}{B_t}.$$

Use the definition of the SDF to rewrite this as

$$\frac{\mathbb{E}^{\mathbb{P}}[S_T D_T \mid \mathcal{F}_t]}{D_t B_t} = \frac{S_t}{B_t},$$

and multiply by $D_t B_t$.

⟶ End of Lecture 6b ⟵

Lecture 7

Dividends, Forwards, Futures, and Futures Options

Chapters 16 & 26

Contents

- Dividends
- Forward and futures contracts
- Futures options

Dividends

Dividends

Recal the standard Black-Scholes model,

$$\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t,$$
$$\mathrm{d}B_t = rB_t \,\mathrm{d}t.$$

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

$D_t =$ The *cumulative* dividends over the interval $[0, t]$.
[*Don't confuse this notation with a SDF!*]

Interpretation. Over the interval $[t, t + \mathrm{d}t]$ you obtain the amount $\mathrm{d}D_t$

Two cases.

- Discrete dividends (realistic but messy), not treated.
- Continuous dividends (unrealistic but easy to handle).

Portfolios and Dividends

Consider a market with N assets, and

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

Assumption: We assume that D has continuous (even differentiable) trajctories.

Definition. The **value process** V is, as before, defined by

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

Self financing portfolios in the case of dividends

Recall the value process of a portfolio h, $V_t = \sum_{i=1}^{N} h_t^i S_t^i$.

New **Definition.** The strategy h is **self financing** if

$$\mathrm{d}V_t = \sum_{i=1}^{N} h_t^i \mathrm{d}G_t^i,$$

where the **gain process** G^i is defined by

$$G_t^i = S_t^i + D_t^i,\ D_0^i = 0.$$

Note that, if $D^i = 0$ we are back in the old situtation.

Interpretation. See D^i as an extra income in addition to S^i.

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

We denote u_t^i the relative share of the portfolio value, which is invested in asset No. i (previously denoted ω_t^i),

$$u_t^i = \frac{h_t^i S_t^i}{V_t} \text{ (as before), so } h_t^i = \frac{u_t^i V_t}{S_t^i}.$$

Recall the new self financing condition,

$$\mathrm{d}V_t = \sum_{i=1}^{N} h_t^i \, \mathrm{d}G_t^i,$$

and substitute h_t^i to get

$$\mathrm{d}V_t = V_t \sum_{i=1}^{N} u_t^i \frac{\mathrm{d}G_t^i}{S_t^i}.$$

Note that without dividends we simply had $\mathrm{d}V_t = V_t \sum_{i=1}^{N} u_t^i \frac{\mathrm{d}S_t^i}{S_t^i}$.

Continuous Dividend Yield

Definition. The stock S pays a **continuous, even differentiable, dividend yield** of $q \geq 0$, if D has the form

$$\mathrm{d}D_t = qS_t\,\mathrm{d}t.$$

Question. How does the dividend affect the price of a European Call? (compared to a non-dividend stock).

Answer. The price is lower. (Why? You should be able to guess!)

Follow up question. Can we quantify this?

Black-Scholes with Continuous Dividend Yield

Consider the dynamics

$$\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \text{ as before,}$$
$$\mathrm{d}D_t = qS_t \,\mathrm{d}t \text{ (new).}$$

The gain process then satisfies (check!)

$$\mathrm{d}G_t = (\alpha + q)S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t.$$

Consider a fixed simple claim

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

and assume that, justified by the Markov property of S,

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

Standard Procedure

- Assume (as above) that the derivative price is of the form

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

- Form a SF portfolio based on underlying S and derivative F, with portfolio dynamics (now with $\mathrm{d}G$ instead of $\mathrm{d}S$)

$$\mathrm{d}V_t = V_t \left[u_t^S \cdot \frac{\mathrm{d}G_t}{S_t} + u_t^F \cdot \frac{\mathrm{d}F}{F} \right].$$

- Choose relative weights u^S and u^F such that the (hidden) $\mathrm{d}W$-term is wiped out. This gives us, for some k_t,

$$\mathrm{d}V_t = V_t k_t \, \mathrm{d}t.$$

- Absence of arbitrage implies

$$k_t = r.$$

- This relation will say something about F, as before.

The standard procedure in detail I

Recall the value dynamics (abbreviated):

$$\mathrm{d}V = V \cdot \left[u^S \frac{\mathrm{d}G}{S} + u^F \frac{dF}{F} \right],$$
$$\mathrm{d}G = S(\alpha + q)\,\mathrm{d}t + \sigma S\,\mathrm{d}W.$$

From Itô applied to $F(t, S_t)$ we obtain

$$dF = \alpha_F F\,\mathrm{d}t + \sigma_F F\,\mathrm{d}W,$$

where (in abbreviated notation)

$$\alpha_F = \frac{1}{F}\left[\frac{\partial F}{\partial t} + \alpha S \frac{\partial F}{\partial s} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial s^2} \right],$$
$$\sigma_F = \frac{1}{F} \cdot \sigma S \frac{\partial F}{\partial s}.$$

The standard procedure in detail II

Collecting the $\mathrm{d}t$ and $\mathrm{d}W$ terms in $\mathrm{d}V$ gives us

$$\mathrm{d}V = V \cdot \left[u^S(\alpha + q) + u^F \alpha_F\right] \mathrm{d}t + V \cdot \left[u^S \sigma + u^F \sigma_F\right] \mathrm{d}W.$$

To wipe out the $\mathrm{d}W$ contribution, we define u^S and u^F as the solutions to the system

$$\begin{aligned} u^S \sigma + u^F \sigma_F &= 0, \\ u^S + u^F &= 1. \end{aligned}$$

The solution is (assume $\sigma_F \neq \sigma$)

$$u^S = \frac{\sigma_F}{\sigma_F - \sigma}, \qquad u^F = \frac{-\sigma}{\sigma_F - \sigma}.$$

The standard procedure in detail III

The value dynamics now become

$$\mathrm{d}V = V \cdot \left[u^S(\alpha + q) + u^F \alpha_F \right] \mathrm{d}t.$$

Absence of arbitrage implies

$$u^S(\alpha + q) + u^F \alpha_F = r,$$

which leads to (similar to slide 72, you verify!)

$$\frac{\partial F}{\partial t} + (r - q)S\frac{\partial F}{\partial s} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial s^2} - rF = 0.$$

Note the absence of the rate of return parameter α in this PDE, as before, but that the dividend rate q figures here.

Pricing PDE

Proposition. It follows from Feynman-Kac that the pricing function F in the case of a dividend paying stock with rate q 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 for $F(t,s) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,s}\Phi(S_T)$, where under a risk neutral measure $\mathbb{Q}$ we have the dynamics

$$\mathrm{d}S_t = (r-q)S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t^{\mathbb{Q}}.$$

Note that in case $q = 0$ (no dividends), we are back in the standard situation.

Risk Neutral Valuation in presence of dividends, summary

The pricing function has the representation

$$F(t,s) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,s}[\Phi(S_T)],$$

where the $\mathbb{Q}$-dynamics of S are given by

$$\mathrm{d}S_t = (r-q)S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W^{\mathbb{Q}}_t.$$

Question. Which object is a martingale under the measure $\mathbb{Q}$? Is it $\frac{S_t}{B_t} = e^{-rt}S_t$ as before?
Answer should be NO, unless $q = 0$,

Martingale Property

Proposition. Under the martingale measure $\mathbb{Q}$ the *normalized* gain process G^Z with

$$G_t^Z = e^{-rt} S_t + \underbrace{\int_0^t e^{-ru} \mathrm{d}D_u}_{\text{extra term}},$$

is a $\mathbb{Q}$-martingale.

Proof. Exercise, in which you show that $\mathrm{d}G_t^Z$ only has $\mathrm{d}W_t^{\mathbb{Q}}$ terms, and no $\mathrm{d}t$ term.

Note. The result above holds in great generality, see slide 267.

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, including dividends. I.e., for $t = 0$,

$$S_0 = \mathbb{E}^{\mathbb{Q}} \left[\int_0^t e^{-ru} \mathrm{d}D_u + e^{-rt} S_t \right].$$

Note that now $S_0 \geq e^{-rt} \mathbb{E}^{\mathbb{Q}} S_t$, the "old" price.

Pricing formula

We will find a pricing formula for claims of the type

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

We are standing at time t, with dividend yield q and $S_t = s$.

- Suppose that you have for the claim Z the pricing function $F^0(t, s)$ for a *non* dividend paying stock.
- Denote the pricing function for the dividend paying stock by

$$F^q(t, s).$$

Proposition. With notation as above we have

$$F^q(t, s) = F^0\left(t, se^{-q(T-t)}\right).$$

The proof is Exercise 16.5.

Moral to handle the dividend case

Use your old formulas, but replace today's stock price s with $se^{-q(T-t)}$. This leads to explicit formulas in the Black-Scholes setting.

Note that $se^{-q(T-t)} \leq s$. If F^0 is increasing in s, then $F^q(t,s) \leq F^0(t,s)$.

This happens for a European call, for a put the result is reversed.

European Call on Dividend-Paying-Stock

Here we have, following the *moral* on slide 262,

$$F^q(t,s) = se^{-q(T-t)}N[d_1] - e^{-r(T-t)}KN[d_2],$$

where now, in the dividend paying situation,

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\log\left(\frac{s}{K}\right) + \left(r - q + \frac{1}{2}\sigma^2\right)(T-t)\right]$$
$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

Compare to the result on slide 78 (with other d_1 and d_2) and observe the role of q. When $q = 0$ the formula reduces to the original BS formula.

$\longrightarrow$ **End of Lecture 7a** $\longleftarrow$

Martingale Analysis

Basic task. We have a general model for the stock price S and cumulative dividends D, under $\mathbb{P}$. How do we find a martingale measure $\mathbb{Q}$, and exactly which objects will be martingales under $\mathbb{Q}$?

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** as they disappear in the bank account.
- Now apply the First Fundamental Theorem to the market (B, V) instead of the original market (B, S).
- Thus there exists a martingale measure $\mathbb{Q}$ such that normalized prices $\frac{\Pi_t}{B_t}$ are $\mathbb{Q}$-martingales for all traded assets (underlying and derivatives) without dividends.
- In particular the process $\frac{V_t}{B_t}$ is a $\mathbb{Q}$ martingale.

Next we study V.

The V Process

We assume to have one unit of stock. Let h_t denote the number of units in the bank account, where $h_0 = 0$. V is then characterized by

$$V_t = 1 \cdot S_t + h_t B_t, \tag{1}$$

$$\mathrm{d}V_t = \underbrace{\mathrm{d}S_t + \mathrm{d}D_t}_{\text{from the new SF condition}} + h_t \,\mathrm{d}B_t. \tag{2}$$

From (1) we obtain (the ordinary product rule applies to $\mathrm{d}(h_t B_t)$ as B_t has no Wiener term)

$$\mathrm{d}V_t = \mathrm{d}S_t + h_t \,\mathrm{d}B_t + B_t \,\mathrm{d}h_t.$$

Comparing this with (2) gives us

$$B_t \,\mathrm{d}h_t = \mathrm{d}D_t \text{ and then } \mathrm{d}h_t = \frac{1}{B_t}\,\mathrm{d}D_t.$$

Integrating this gives us, use $h_0 = 0$,

$$h_t = \int_0^t \frac{1}{B_s}\,\mathrm{d}D_s.$$

More on the normalized gain process

We thus have

$$V_t = S_t + B_t \int_0^t \frac{1}{B_s} \, \mathrm{d}D_s, \tag{3}$$

and the first fundamental theorem (FTAP 1) gives us the following result.

Proposition. For a market with dividends, the martingale measure $\mathbb{Q}$ is characterized by the fact that the **normalized gain process** $G_t^Z = \frac{V_t}{B_t}$ satisfies

$$G_t^Z = \frac{S_t}{B_t} + \int_0^t \frac{1}{B_s} \mathrm{d}D_s$$

and is a $\mathbb{Q}$-martingale.

Quiz (Björk). Could you have guessed the formula (3) for V?

Continuous Dividend Yield

Recall the model under $\mathbb{P}$,

$$\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t,$$
$$\mathrm{d}D_t = qS_t \,\mathrm{d}t.$$

Let $Z = S/B$. We recall from the previous slide

$$G_t^Z = \frac{S_t}{B_t} + \int_0^t \frac{1}{B_s} \,\mathrm{d}D_s = Z_t + \int_0^t qZ_s \,\mathrm{d}s.$$

Easy calculations (you do!) give us (under $\mathbb{P}$)

$$\mathrm{d}G_t^Z = Z_t(\alpha - r + q)\,\mathrm{d}t + Z_t\sigma \,\mathrm{d}W_t.$$

Apply the Girsanov transformation $\mathrm{d}\mathbb{Q} = L_T \,\mathrm{d}\mathbb{P}$ on $\mathcal{F}_T$, where

$$\mathrm{d}L_t = L_t\phi_t \,\mathrm{d}W_t \text{ (for some } \phi_t),$$

and recall

$$\mathrm{d}W_t = \phi_t \,\mathrm{d}t + \mathrm{d}W_t^{\mathbb{Q}}.$$

$\mathbb{Q}$-dynamics of G^Z and S

Insert the equation for $\mathrm{d}W_t$ into $\mathrm{d}G^Z$ to obtain the $\mathbb{Q}$ dynamics of G^Z,

$$\mathrm{d}G_t^Z = Z_t(\alpha - r + q + \sigma\phi_t)\,\mathrm{d}t + Z_t\sigma\,\mathrm{d}W_t^{\mathbb{Q}}.$$

The **Martingale condition** for G^Z is (to wipe out $\mathrm{d}t$ terms)

$$\alpha - r + q + \sigma\phi_t = 0.$$

Using the definition of $W^{\mathbb{Q}}$, we directly get

$$\mathrm{d}S_t = S_t(\alpha + \sigma\phi)\,\mathrm{d}t + S_t\sigma\,\mathrm{d}W_t^{\mathbb{Q}}.$$

Using the martingale condition this gives us the $\mathbb{Q}$-dynamics of S,

$$\mathrm{d}S_t = S_t(r - q)\,\mathrm{d}t + S_t\sigma\,\mathrm{d}W_t^{\mathbb{Q}},$$

which we have already encountered on slide 259.

Risk Neutral Valuation in the case of dividends, summary

Theorem. For a T-claim X, the price process $\Pi_t[X]$ is given by

$$\Pi_t[X] = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[X \mid \mathcal{F}_t],$$

where the $\mathbb{Q}$-dynamics of S are given by

$$\mathrm{d}S_t = (r-q)S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t^{\mathbb{Q}}.$$

Note that the formula for $\Pi_t[X]$ in the case $X = \Phi(S_T)$ is given by $F(t, S_t)$, where F satisfies the PDE on slide 258 by the Feynman-Kac formula.

Forward 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 $X \in \mathcal{F}_T$ 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 $f(t; T, X) \in \mathcal{F}_t$ 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. This happens by compensating cash flows, a swap of products that net to zero.

General Risk Neutral Formula

Suppose we have a bank account B with dynamics

$$\mathrm{d}B_t = r_t B_t \,\mathrm{d}t, \quad B_0 = 1,$$

with a (*possibly stochastic*) short rate r_t (r will be an adapted process). Then

$$B_t = e^{\int_0^t r_s \,\mathrm{d}s} \in \mathcal{F}_t.$$

Recall from slide 219 the risk neutral valuation for a T-claim X,

$$\Pi_t[X] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s \,\mathrm{d}s} \cdot X \,\middle|\, \mathcal{F}_t\right].$$

Setting $X = 1$ this gives the price, at time t, of a **zero coupon bond** maturing at T (for more on this, see the book Section 29.1 and slide 335 further down) as

$$p(t,T) := \Pi_t[1] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s \,\mathrm{d}s} \,\middle|\, \mathcal{F}_t\right] = B_t \mathbb{E}^{\mathbb{Q}}\left[B_T^{-1} \,\middle|\, \mathcal{F}_t\right].$$

Forward Price Formula

Theorem. The forward price of the claim X is given by

$$f(t,T) = \frac{1}{p(t,T)}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s\,\mathrm{d}s}\cdot X \;\middle|\; \mathcal{F}_t\right],$$

where $p(t,T)$ denotes the price at time t of a zero coupon bond maturing at time T. Note the (extra) normalizing factor $p(t,T)$in the denominator .

In particular, if the short rate r is deterministic we have

$$f(t,T) = \mathbb{E}^{\mathbb{Q}}[X \mid \mathcal{F}_t],$$

and note that for the forward price there is no discounting.

Proof

The net cash flow at maturity is $X - f(t,T)$. If at time t this equals zero (as in the definition of the forward contract), we obtain $\Pi_t[X] = \Pi_t[f(t,T)]$. We compute the latter. For any X one has

$$\Pi_t[X] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s\,\mathrm{d}s}\cdot X \;\middle|\; \mathcal{F}_t\right].$$

Take now $X = f(t,T)$. Since $f(t,T)$ is known at t, we obviously have (why?) from the definition of $p(t,T)$ on slide 273

$$\Pi_t[f(t,T)] = \mathbb{E}^{\mathbb{Q}}\left[f(t,T)e^{-\int_t^T r_s\,\mathrm{d}s}\middle|\; \mathcal{F}_t\right] = f(t,T)p(t,T),$$

but then also $\Pi_t[X] = f(t,T)p(t,T)$, which implies the main result. If r is a deterministic constant then $p(t,T) = e^{-r(T-t)}$ which gives us the second formula.

$\longrightarrow$ End of Lecture 7b $\longleftarrow$

Lecture 8

Futures Contracts

Futures Contracts, definition

A futures contract on a 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 $F(t; T, X) \in \mathcal{F}_t$, 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, both $\mathcal{F}_T$-measurable. Note $F(T; T, X) = X$.

(iii) During an arbitrary time interval $(s, t]$ the holder of the contract receives the amount $F(t; T, X) - F(s; T, X)$. (The cash flow $F(t; T, X)$ thus looks similar to a dividend.

(iv) The *spot price*, at any time t prior to delivery, for buying or selling the futures contract, is by definition equal to zero.

A quote

Björk writes in his book (Arbitrage Theory in Continuous Time, 3rd edition, page 455) the following text, to which I agree to a large extent.

- If the reader thinks that a futures contract conceptually is a somewhat complicated object, then the author is inclined to agree.

Futures Price Formula (see Section 29.2)

From the definition (in particular (iii)) it is clear that a futures contract is a considered *price-dividend pair* (S, D) with

$$S \equiv 0,\ D_t = F(t, T).$$

From general theory (slide 267), the normalized gains process

$$G_t^Z = \frac{S_t}{B_t} + \int_0^t \frac{1}{B_s} \mathrm{d}D_s$$

is a $\mathbb{Q}$-martingale. Since now $S \equiv 0$ and $\mathrm{d}D_t = \mathrm{d}F(t, T)$ this implies that

$$\mathrm{d}M_t := \frac{1}{B_t} \mathrm{d}F(t, T) \text{ (often of the type } h_t\, \mathrm{d}W_t^{\mathbb{Q}})$$

is a martingale increment, which implies (why?) that $\mathrm{d}F(t, T) = B_t\, \mathrm{d}M_t$ (often $B_t h_t\, \mathrm{d}W_t^{\mathbb{Q}}$) is a martingale increment.

Futures price

Thus $F(t,T)$ is a $\mathbb{Q}$-martingale in t and we have (see (iii) again)

$$F(t,T) = \mathbb{E}^{\mathbb{Q}}[F(T,T) \mid \mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[X \mid \mathcal{F}_t].$$

This is summarized in the following result.

Theorem. The futures price process is given by

$$F(t,T) = \mathbb{E}^{\mathbb{Q}}[X \mid \mathcal{F}_t].$$

Note that there is no discounting in this price formula.

Corollary. If the short rate is deterministic, then the futures and forward prices coincide, compare to the formula on slide 274.

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 in the notation and write F_t, i.e., $F_t = F(t, T)$. *Don't confuse this notation with e.g. a partial derivative of F.*

Definition. A European futures call option X, with strike price K and exercise date T, on a futures contract with delivery date T_1 will, *if exercised* at $T < T_1$, pay to the holder

- the amount $F(T, T_1) - K$ in cash, so $X = (F(T, T_1) - K)^+$;
- a long position in the underlying futures contract.

The long position above can immediately be closed at no cost (the spot price of futures is zero), so the focus will be on $(F(T, T_1) - K)^+$.

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, tons of potatoes, hundreds of pigs, etc.).
- The market place for futures and futures options is often the same. This facilitates hedging etc.

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.

Pricing futures options, set up

We consider a futures contract with delivery date T_1 (fixed) and use the notation $F_t = F(t, T_1)$. We assume the following dynamics for F.

$$\mathrm{d}F_t = \mu F_t \,\mathrm{d}t + \sigma F_t \,\mathrm{d}W_t.$$

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" (given previously developed knowledge).

Pricing futures options, martingale result

From risk neutral valuation we know that the price process $\Pi_t[\Phi]$ is of the form (by analogy with pricing of stock derivatives)

$$\Pi_t[\Phi] = f(t, F_t),$$

where f is given by, using the $F_t = F$, the F-value at t,

$$f(t, F) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,F}[\Phi(F_T)] := e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\Phi(F_T) \mid F_t = F].$$

We therefore need the $\mathbb{Q}$-dynamics for F. Let's quote from slide 280.

"The futures price process F_t is a $\mathbb{Q}$-martingale."

Thus the $\mathbb{Q}$-dynamics of F are given by (no $\mathrm{d}t$ terms as in the $\mathbb{P}$-dynamics on the previous slide)

$$\mathrm{d}F_t = \sigma F_t \, \mathrm{d}W_t^{\mathbb{Q}}.$$

Pricing futures options, comparing to dividends

Summarizing, we thus have

$$f(t,F) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,F}[\Phi(F_T)],$$
$$\mathrm{d}F_t = \sigma F_t \,\mathrm{d}W^{\mathbb{Q}}_t.$$

Now recall the formula for a stock with continuous dividend yield q, see slide 269,

$$f(t,s) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,s}[\Phi(S_T)]$$

with $\mathbb{Q}$-dynamics

$$\mathrm{d}S_t = (r-q)S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W^{\mathbb{Q}}_t.$$

Observation. If we set $q = r$ and replace S_t with F_t 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 (book Proposition 7.13). With the notation as above we have

$$f(t, F) = f_0(t, Fe^{-r(T-t)}).$$

The proof of this proposition is Exercise 16.5.

Moral: Reset today's futures price F to $Fe^{-r(T-t)}$ and use your formulas for stock options.

Compare to slide 261 for the dividends case, similar story; replace q there with r to arrive at the formula for the futures case.

Black 1976 Formula

The price of a futures option with exercise date T and exercise price K is given by (F is the futures price at time t)

$$c = e^{-r(T-t)}\{FN[d_1] - KN[d_2]\},$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\log\left(\frac{F}{K}\right) + \frac{1}{2}\sigma^2(T-t)\right],$$
$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

[*To prove this, see the formula on slide 263 with $q = r$ and $s = F$.*]

Currency Derivatives

Foreign e**X**change market, **FX**

Chapter 17

Pure Currency Contracts

Consider two markets, domestic (Portugal) and foreign (USA), with

r^d = domestic short rate,

r^f = foreign short rate,

X = exchange rate.

Note that r^d and r^f are usually different!

The *exchange rate* X is quoted as

$$\frac{\text{units of the domestic currency}}{\text{unit of the foreign currency}}.$$

Example: If €1=\$1.09, then $X = \frac{1}{1.09}$.

Simple Model for the exhange rate (Garman-Kohlhagen)

The $\mathbb{P}$-dynamics are given as (note the different bank accounts)

$$\begin{aligned}
\mathrm{d}X_t &= \alpha X_t \,\mathrm{d}t + \sigma X_t \,\mathrm{d}W_t, \\
\mathrm{d}B_t^d &= r^d B_t^d \,\mathrm{d}t, \\
\mathrm{d}B_t^f &= r^f B_t^f \,\mathrm{d}t.
\end{aligned}$$

Main Problem. Find an arbitrage free price (in domestic currency) for the currency derivative, Z, of the form

$$Z = \Phi(X_T).$$

Typical example. European Call on the exchange rate X,

$$Z = \max\{X_T - K, 0\}.$$

Naive idea to solve the problem

For the European Call, use the standard Black-Scholes formula, with S replaced by X and r replaced by r^d.

Is this OK?

"This is again a suspicious question",
you can guess the answer!

NO!

WHY?

$\longrightarrow$ End of Lecture 8a $\longleftarrow$

Main Idea

- Notice first a difference. A stock is a financial product that is traded on the market. An exchange rate X is *not* a traded product. This is a first argument why analogies with a stock market break down.
- When you buy stock you just keep the asset, although asset prices keep on fluctuating, until you sell it. There is no interest on asset prices.
- When you buy dollars, these are put into a bank account, giving the interest r^f. Although the exchange rate may fluctuate, this does *not* affect what you have on your bank account in the foreign currency.

Moral. Buying a currency is like (there are many similarities) buying a *dividend-paying* stock with dividend yield $q = r^f$, as we shall see on slide 301.

Transformation technique

- Transform all objects into **domestically traded** asset prices. This yields a model of the transformed market.
- Use standard techniques on the transformed model.

Transformed Market

- Investing foreign currency in the foreign bank gives value dynamics **in foreign currency** according to

$$\mathrm{d}B_t^f = r^f B_t^f \,\mathrm{d}t.$$

- B^f units of the foreign currency is worth $B^f \cdot X$ in the domestic currency ($B_t^f \cdot X_t$ at time t).
- Trading in the foreign currency is equivalent to trading in a domestic market with the *domestic* price process (this is the transformation)

$$\tilde{B}_t^f = B_t^f \cdot X_t.$$

- Study the domestic market consisting of the two assets

$$\tilde{B}^f,\ B^d.$$

Market dynamics

Summary (all in domestic prices).

$$\begin{aligned} \mathrm{d}X_t &= \alpha X_t \,\mathrm{d}t + \sigma X_t \,\mathrm{d}W, \\ \tilde{B}^f_t &= B^f_t \cdot X_t. \end{aligned}$$

Using Itô we have domestic market dynamics

$$\begin{aligned} \mathrm{d}\tilde{B}^f_t &= (\alpha + r^f)\tilde{B}^f_t \,\mathrm{d}t + \sigma \tilde{B}^f_t \,\mathrm{d}W_t, \\ \mathrm{d}B^d_t &= r^d B^d_t \,\mathrm{d}t. \end{aligned}$$

Let $\mathbb{Q}^d$ be an EMM for the domestic market and W^d a Wiener process under $\mathbb{Q}^d$. Standard results for "Girsanov" give us $\mathbb{Q}^d$-dynamics for domestically traded asset prices (note the domestic interest rate r^d for B^d),

$$\begin{aligned} \mathrm{d}\tilde{B}^f_t &= r^d \tilde{B}^f_t \,\mathrm{d}t + \sigma \tilde{B}^f_t \,\mathrm{d}W^d_t. \\ \mathrm{d}B^d_t &= r^d B^d_t \,\mathrm{d}t. \end{aligned}$$

Risk neutral Valuation of a currency derivative

Itô gives us $\mathbb{Q}^d$-dynamics for the quotient $X_t = \tilde{B}_t^f / B_t^f$ (check!),

$$\mathrm{d}X_t = (r^d - r^f)X_t\,\mathrm{d}t + \sigma X_t\,\mathrm{d}W_t^d.$$

Theorem. The arbitrage free price $\Pi_t[\Phi]$ is given by

$$\Pi_t[\Phi] = F(t, X_t),$$

where

$$F(t,x) = e^{-r^d(T-t)}\mathbb{E}_{t,x}^{\mathbb{Q}}[\Phi(X_T)].$$

Pricing PDE

The Feynman-Kac formula yields the following result (recall again the $\mathbb{Q}^d$-dynamics of X).

Theorem. The pricing function F solves the boundary value problem

$$\frac{\partial F}{\partial t} + x(r^d - r^f)\frac{\partial F}{\partial x} + \frac{1}{2}x^2\sigma^2\frac{\partial^2 F}{\partial x^2} - r^d F = 0,$$
$$F(T, x) = \Phi(x).$$

Currency vs Equity Derivatives

Proposition. Introduce the following pricing functions.

- $F^0(t,x)$, the pricing function for the claim $Z = \Phi(X_T)$, 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 relation holds

$$F(t,x) = F^0\left(t, xe^{-r^f(T-t)}\right).$$

This is very similar to the result on slide 261 for the dividend case, where we had q instead of r^f.

Currency Option Formula

For a European currency call there exists an explicit formula for the pricing function F. The price of such a call is given by

$$F(t,x) = xe^{-r^f(T-t)}N[d_1] - e^{-r^d(T-t)}KN[d_2],$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left\{\log\left(\frac{x}{K}\right) + \left(r^d - r^f + \frac{1}{2}\sigma_X^2\right)(T-t)\right\},$$
$$d_2 = d_1(t,x) - \sigma\sqrt{T-t}.$$

Upon renaming the constants, this formula is the same as the one on slide 263 where we dealt with dividends.

Probability measures for a currency market

Martingale Analysis

We need the risk neutral measures for each of the markets,

$$\mathbb{Q}^d = \text{domestic martingale measure},$$
$$\mathbb{Q}^f = \text{foreign martingale measure},$$

and the corresponding likelihood processes (all R-N derivatives defined on $\mathcal{F}_t$),

$$L_t = \frac{\mathrm{d}\mathbb{Q}^f}{\mathrm{d}\mathbb{Q}^d}, \quad L_t^d = \frac{\mathrm{d}\mathbb{Q}^d}{\mathrm{d}\mathbb{P}}, \quad L_t^f = \frac{\mathrm{d}\mathbb{Q}^f}{\mathrm{d}\mathbb{P}}.$$

The $\mathbb{P}$-dynamics of X are

$$\mathrm{d}X_t = \alpha_t X_t \,\mathrm{d}t + \sigma_t X_t \,\mathrm{d}W_t,$$

where α and σ are arbitrary adapted processes and W is $\mathbb{P}$-Wiener.

Problem. How are $\mathbb{Q}^d$ and $\mathbb{Q}^f$ related?

Of course through L^d and L^f, but how exactly?

Main Idea

Fix an arbitrary foreign T-claim Z.

- Compute foreign price and *change to domestic currency*. The foreign price at $t = 0$ will be

$$\Pi_0[Z] = X_0 \mathbb{E}^{\mathbb{Q}^f} \left[e^{-\int_0^T r_s^f \, \mathrm{d}s} Z \right].$$

This can be written as (in domestic currency)

$$\Pi_0[Z] = X_0 \mathbb{E}^{\mathbb{Q}^d} \left[L_T e^{-\int_0^T r_s^f \, \mathrm{d}s} Z \right].$$

OR

- Change into domestic currency at T ($Z \to ZX_T$) and then compute domestic arbitrage free price. This gives us

$$\Pi_0[Z] = \mathbb{E}^{\mathbb{Q}^d} \left[e^{-\int_0^T r_s^d \, \mathrm{d}s} X_T \cdot Z \right].$$

- These expressions must be equal for all choices of $Z \in \mathcal{F}_T$.

Intermezzo

Consider random variables X, Y, Z (general case, nothing to do with currencies) for which the expectations below exist.

Lemma. If

$$\mathbb{E}(XZ) = \mathbb{E}(YZ) \text{ for all } Z,$$

then $X = Y$ a.s.

Proof. Rewrite the equality as $\mathbb{E}(X - Y)Z = 0$ and take the special choice $Z = \mathbf{1}\{X > Y\}$. Then $(X - Y)Z \geq 0$ and (still) $\mathbb{E}(X - Y)Z = 0$. It follows from basic measure theory that then necessarily $(X - Y)Z = 0$ a.s., $\mathbb{P}((X - Y)\mathbf{1}\{X > Y\} = 0) = 1$. It follows that $\mathbb{P}(X - Y = 0) = 1$.

Back to currencies

From slide 305 we obtain

$$\mathbb{E}^{\mathbb{Q}^d}\left[e^{-\int_0^T r_s^d\,\mathrm{d}s}X_T\cdot Z\right]=X_0\mathbb{E}^{\mathbb{Q}^d}\left[L_T e^{-\int_0^T r_s^f\,\mathrm{d}s}Z\right]$$

for all T-claims Z. Change from T to t. By the lemma this implies the following result.

Theorem. The exchange rate X is given by

$$X_t = X_0 e^{\int_0^t (r_s^d - r_s^f)\,\mathrm{d}s} L_t,$$

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 the definitions on slide 241,

$$D_t^f = \frac{L_t^f}{B_t^f},\ D_t^d = \frac{L_t^d}{B_t^d},$$

and then (on $\mathcal{F}_t$)

$$L_t = \frac{\mathrm{d}\mathbb{Q}^f}{\mathrm{d}\mathbb{Q}^d} = \frac{\frac{\mathrm{d}\mathbb{Q}^f}{\mathrm{d}\mathbb{P}}}{\frac{\mathrm{d}\mathbb{Q}^d}{\mathrm{d}\mathbb{P}}} = \frac{L_t^f}{L_t^d} = \frac{D_t^f B_t^f}{D_t^d B_t^d} = \frac{D_t^f}{D_t^d}\frac{B_t^f}{B_t^d},$$

together with

$$e^{\int_0^t (r_s^d - r_s^f)\,\mathrm{d}s} = \frac{B_t^d}{B_t^f}.$$

Put the displayed results together.

$\mathbb{Q}^d$-Dynamics of X, first steps

In particular, since L is a positive $\mathbb{Q}^d$-martingale the $\mathbb{Q}^d$ dynamics of L are of the form

$$\text{(a)} \qquad \mathrm{d}L_t = L_t \phi_t \, \mathrm{d}W_t^d,$$

where W^d is $\mathbb{Q}^d$-Wiener. See the general theory on slide 176 which you apply with $\mathbb{Q}^d$ instead of $\mathbb{P}$ and $\mathbb{Q}^f$ replacing $\mathbb{Q}$.

From the Theorem on the previous page and

$$\text{(b)} \qquad X_t = X_0 e^{\int_0^t (r_s^d - r_s^f)\,\mathrm{d}s} L_t,$$

the $\mathbb{Q}^d$-dynamics of X follows by (a), (b), Itô and a "bit of work" (for you to do!) as

$$\mathrm{d}X_t = (r_t^d - r_t^f) X_t \, \mathrm{d}t + \phi_t X_t \, \mathrm{d}W_t^d.$$

Note the ϕ_t in front of $\mathrm{d}W_t^d$, compare to the $\mathbb{P}$-dynamics on slide 304.

$\mathbb{Q}^d$-Dynamics of X, result

It follows that the Girsanov kernel ϕ equals the exchange rate volatility σ and we have the general $\mathbb{Q}^d$ dynamics as in the following Theorem.

Theorem. The $\mathbb{Q}^d$ dynamics of X are of the form

$$\mathrm{d}X_t = (r_t^d - r_t^f)X_t \,\mathrm{d}t + X_t\sigma_t \,\mathrm{d}W_t^d,$$

already known from slide 299.

⟶ End of Lecture 8b ⟵

Lecture 9

Probability measures for a currency market continued

Market Prices of Risk

Recall the domestic and foreign stochastic discount factors,

$$D_t^d = e^{-\int_0^t r_s^d \,\mathrm{d}s} L_t^d = \frac{L_t^d}{B_t^d},\ D_t^f = \frac{L_t^f}{B_t^f}.$$

We also have representations under $\mathbb{P}$ like

$$\mathrm{d}L_t^d = L_t^d \phi_t^d \,\mathrm{d}W_t,$$

where $-\phi_t^d = \lambda_t^d$ is the domestic *market price of risk*[1] and similar for ϕ_t^f etc. Apply now Itô calculus, this is an exercise, to deduce

$$X_t = X_0 \frac{D_t^f}{D_t^d}.$$

[1] This is new terminology and will be explained in detail later, around slide 356; just accept the name for now.

The exchange rate volatility

One then "easily" obtains for some process α

$$\mathrm{d}X_t = \alpha_t X_t \,\mathrm{d}t + (\lambda_t^d - \lambda_t^f) X_t \,\mathrm{d}W_t,$$

where we do not care about the exact shape of α (we don't specify it). We conclude as follows.

Theorem: The exchange rate volatility is given by

$$\sigma_t = \lambda_t^d - \lambda_t^f,$$

a relation between market prices of risk and volatility.

Siegel's Paradox

Assume that the domestic and the foreign markets are *both* **risk neutral** and assume constant short rates. We now have the following surprising (?) argument, based on scenarios A and B.

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 (recall that now $\mathbb{Q}^d = \mathbb{P}$ by the risk neutrality assumption)

$$e^{-r^d T}\mathbb{E}^{\mathbb{P}}[X_T].$$

We thus have

$$X_0 e^{-r^f T} = e^{-r^d T}\mathbb{E}^{\mathbb{P}}[X_T].$$

Siegel's Paradox continued

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 (at $t = 0$) is

$$\frac{1}{X_0} e^{-r^d T}.$$

The €1 claim is identical to a T-claim of $\frac{1}{X_T}$ dollars. So, by the foreign risk neutrality $\mathbb{Q}^f = \mathbb{P}$, we obtain the dollar price as

$$e^{-r^f T} \mathbb{E}^{\mathbb{P}} \left[\frac{1}{X_T} \right],$$

which gives us by equating the two obtained dollar prices

$$\frac{1}{X_0} e^{-r^d T} = e^{-r^f T} \mathbb{E}^{\mathbb{P}} \left[\frac{1}{X_T} \right].$$

Intermezzo, Jensen's inequalities

Let X be a random variable, with values in an interval $I \subset \mathbb{R}$, of which the relevant expectations below are finite. Let ϕ be a *convex* function on I. Then Jensen's inequality is

$$\mathbb{E}\phi(X) \geq \phi(\mathbb{E}X).$$

Moreover, if ϕ is strictly convex and if X is not a constant, then one even has the strict version

$$\mathbb{E}\phi(X) > \phi(\mathbb{E}X).$$

An example of such a ϕ is $\phi(x) = \frac{1}{x}$, when $I = (0, \infty)$. Hence, for positive X, not constant, it holds that $\mathbb{E}[\frac{1}{X}] > \frac{1}{\mathbb{E}[X]}$.

Back to Siegel's Paradox

Recall our results from scenarios A and B, slides 314 and 315,

$$X_0 e^{-r^f T} = e^{-r^d T} \mathbb{E}^{\mathbb{P}}[X_T],$$
$$\frac{1}{X_0} e^{-r^d T} = e^{-r^f T} \mathbb{E}^{\mathbb{P}}\left[\frac{1}{X_T}\right].$$

Combining (multiply for instance) these equations gives us

$$\mathbb{E}^{\mathbb{P}}\left[\frac{1}{X_T}\right] = \frac{1}{\mathbb{E}^{\mathbb{P}}[X_T]}.$$

This is, by Jensen's inequality, 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?

Formal analysis of Siegel's Paradox

Question. Can we assume that both the domestic and the foreign markets are risk neutral?

Answer. Generally no (because of Jensen).

Argument. The assumption would be equivalent to assuming the $\mathbb{P} = \mathbb{Q}^d = \mathbb{Q}^f$ i.e., $\lambda_t^d = \lambda_t^f = 0$, as we must have $L_t^d = L_t^f \equiv 1$.

However, we know (see slide 313) 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 for X because of

$$\mathrm{d}X_t = \alpha_t X_t \,\mathrm{d}t + 0 \times X_t \,\mathrm{d}W_t = X_t \alpha_t \,\mathrm{d}t,$$

which is not realistic.

Moral, solving (convincingly?) the paradox

The previous slide gave us the mathematical result, but the intuitive question remains why Americans cannot be risk neutral at the same time as Europeans. The answer is roughly as follows.

- Risk neutrality (or risk aversion) is always **defined in terms of a given numéraire**, but here we have B^d and B^f.
- It is **not** an attitude towards **risk as such**, but rather refers to a specific market.
- You can therefore **not** be risk neutral w.r.t. two different numéraires at the same time unless the ratio between them is deterministic. In particular we cannot have risk neutrality w.r.t. Dollars and Euros at the same time.
- If you are risk neutral in one market, you cannot be so in the other one because of the *stochastic* exchange rate, an additional risk.

$\longrightarrow$ **End of Lecture 9a** $\longleftarrow$

Change of Numéraire

Chapter 26

Recap of General Theory

Consider a market with asset prices

$$S_t^0, S_t^1, \ldots, S_t^N.$$

Theorem (FTAP 1). The market is arbitrage free **iff** there exists an EMM, i.e., a measure $\mathbb{Q}$ such that

- $\mathbb{Q}$ and $\mathbb{P}$ are equivalent, i.e.,

  $$\mathbb{Q} \sim \mathbb{P}, \text{ and}$$

- 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 $\mathbb{Q}$-martingales.

Recap continued

Recall the normalized market

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

- 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*. Indeed, if $\mathrm{d}Z_t^0 = r_t Z_t \,\mathrm{d}t$, $Z_0^0 = 1$, then $Z_t^0 = \exp(\int_0^t r_s \,\mathrm{d}s)$.

  Hence, if $Z_t^0 \equiv 1$, then $r_t \equiv 0$.

Dependence on numéraire

- The EMM $\mathbb{Q}$ will obviously depend on the *choice of numéraire*, so we should really write $\mathbb{Q}^0$ to emphasize that we are using S^0 as numéraire.
- So far we have only considered the case when the numéraire asset is the bank account, i.e., when

$$S_t^0 = B_t.$$

  In this case, the martingale measure $\mathbb{Q}^B$ is referred to as "the risk neutral martingale measure".
- Henceforth the notation $\mathbb{Q}$ (without upper index) will only be used for the risk neutral martingale measure, i.e., $\mathbb{Q} = \mathbb{Q}^B$.
- We will now consider the case of a general numéraire.

General change of numéraire.

- Consider a financial market, including a bank account B.
- Assume that the market is using a fixed risk neutral measure $\mathbb{Q}$ as pricing measure, so $\frac{S_t^i}{B_t}$ is a $\mathbb{Q}$-martingale if $\mathbb{Q} = \mathbb{Q}^B$ for each i.
- Choose, *alternatively*, a fixed asset S as a numéraire, and denote the corresponding martingale measure by $\mathbb{Q}^S$, so that

$$\frac{S_t^i}{S_t} \text{ is a martingale under } \mathbb{Q}^S.$$

Problems.

- Determine $\mathbb{Q}^S$, i.e., determine

$$L_t = \frac{\mathrm{d}\mathbb{Q}^S}{\mathrm{d}\mathbb{Q}} \text{ on } \mathcal{F}_t.$$

- Develop pricing formulas for contingent claims using $\mathbb{Q}^S$ instead of $\mathbb{Q}$. We will write $\mathbb{E}^S$ for expectation under $\mathbb{Q}^S$.

Constructing $\mathbb{Q}^S$

Fix a T-claim X. From general theory we know that

$$\Pi_0[X] = \mathbb{E}^{\mathbb{Q}}\left[\frac{X}{B_T}\right], \text{ as } \frac{\Pi_t(X)}{B_t} \text{ is a } \mathbb{Q}^B\text{-martingale.}$$

Since $\mathbb{Q}^S$ is a martingale measure for the numéraire S, the normalized process

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

is a $\mathbb{Q}^S$-martingale. We thus have with $L_T = \frac{\mathrm{d}\mathbb{Q}^S}{\mathrm{d}\mathbb{Q}}$ on $\mathcal{F}_T$,

$$\frac{\Pi_0[X]}{S_0} = \mathbb{E}^S\left[\frac{\Pi_T[X]}{S_T}\right] = \mathbb{E}^S\left[\frac{X}{S_T}\right] = \mathbb{E}^{\mathbb{Q}}\left[L_T\frac{X}{S_T}\right].$$

From this we obtain (use that S_0 is nonrandom)

$$\Pi_0[X] = \mathbb{E}^{\mathbb{Q}}\left[L_T\frac{X\cdot S_0}{S_T}\right].$$

Finding L_t

For all $X \in \mathcal{F}_T$ we thus have

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{X}{B_T}\right] = \mathbb{E}^{\mathbb{Q}}\left[L_T \frac{X \cdot S_0}{S_T}\right].$$

Recall slide 306. If, in short, on a $(\Omega, \mathcal{F}, \mathbb{P})$ one has $\mathbb{E}[Y \cdot X] = \mathbb{E}[Z \cdot X]$, for all Z, then we have $Y = Z$, $\mathbb{P}$-a.s.

From this result we conclude that $\mathbb{Q}$-a.s.

$$\frac{1}{B_T} = L_T \frac{S_0}{S_T}.$$

By equivalence of the measures, this is also valid $\mathbb{P}$-a.s. and similarly

$$\frac{1}{B_t} = L_t \frac{S_0}{S_t}, \ \mathbb{P}\text{-a.s.}$$

Main result

Proposition. The likelihood process

$$L_t = \frac{\mathrm{d}\mathbb{Q}^S}{\mathrm{d}\mathbb{Q}} \text{ on } \mathcal{F}_t$$

is given by

$$L_t = \frac{S_t}{B_t} \cdot \frac{1}{S_0} = \frac{S_t/S_0}{B_t/B_0}.$$

Moreover, application of general theory, see slide 176 yields that L is a $\mathbb{Q}$-martingale.

Easy exercises

1. Convince yourself that indeed L is a $\mathbb{Q}$-martingale.
2. Assume that a process A_t has the property that A_t/B_t is a $\mathbb{Q}$-martingale. Show that this implies that A_t/S_t is a $\mathbb{Q}^S$-martingale. Interpret the result.

   Hint: One possibility is to show this using the Abstract Bayes formula. See also the similar Exercise 3 of the *Additional Exercises*.

Pricing

Theorem. For every T-claim X we have the pricing formula

$$\Pi_t[X] = S_t \mathbb{E}^S \left[\frac{X}{S_T} \mid \mathcal{F}_t \right].$$

Proof. Follows directly from the $\mathbb{Q}^S$-martingale property of $\Pi_t[X]/S_t$, parallel to the usual property under $\mathbb{Q}$.

Note. We observe S_t directly on the market.

The pricing formula above is particularly useful when X is of the form

$$X = S_T \cdot Y.$$

In this case we obtain straight from the theorem

$$\Pi_t[X] = \Pi_t[S_T Y] = S_t \mathbb{E}^S[Y \mid \mathcal{F}_t],$$

with only one random variable, Y, instead of the ratio $\frac{X}{S_T}$ inside the conditional expectation.

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 $\mathbb{Q}^0$ we obtain from the homogeneity

$$\begin{aligned}\Pi_t[X] &= S_t^0 \mathbb{E}^0\left[\frac{\Phi\left(S_T^0, S_T^1\right)}{S_T^0} \mid \mathcal{F}_t\right] \\ &= S_t^0 \mathbb{E}^0\left[\Phi\left(1, \frac{S_T^1}{S_T^0}\right) \mid \mathcal{F}_t\right].\end{aligned}$$

Important example continued

Proposition. For a claim of the form

$$X = \Phi\left(S_T^0, S_T^1\right),$$

where Φ is homogeneous, we have

$$\Pi_t[X] = S_t^0 \mathbb{E}^0[\varphi(Z_T) \mid \mathcal{F}_t],$$

where

$$\varphi(z) = \Phi(1, z),$$
$$Z_T = \frac{S_T^1}{S_T^0}.$$

Note that Z_T is the normalized version of S_T^1 when using the numéraire process S^0.

Exchange option

Consider an exchange option (the name has *nothing* to do with exhange rates!), i.e., a claim X given by

$$X = \max\{S_T^1 - S_T^0, 0\}.$$

Since $\Phi(x, y) = \max\{x - y, 0\}$ is homogeneous (check!), we obtain

$$\Pi_t[X] = S_t^0 \mathbb{E}^0[\max\{Z_T - 1, 0\} \mid \mathcal{F}_t].$$

- We deal with a European Call on Z with strike price $K = 1$.
- Zero interest rate in the normalized market.
- If S^0 and S^1 are both geometric Brownian motions (GBMs), then so is Z (needs proof, see slide 333), and the price will be given by a *version* of the Black-Scholes formula. Piece of cake!

Intermezzo, some remarks

That Z on the previous page is a GBM, is not immediately obvious. On the background is the following statement.

$$\textit{If } X \textit{ and } Y \textit{ are normal, so are } X + Y \textit{ and } X - Y\,.$$

But for this to hold true, one needs an *extra assumption*, namely that the vector (X, Y) *is bivariate normal*, or (special case) that X and Y are independent.

A GBM X is such that $\log X_t$ has a normal distribution, similar for another GBM Y. Hence, it follows from the above claim that $\log X_t - \log Y_t$ is normal under an independence assumption, in which case we have $\log \frac{X_t}{Y_t}$ is normal. This suggests that $\frac{X_t}{Y_t}$ is a GBM as well.

The way to do this properly is to write down the dynamics for X and Y and play a bit with Itô for $\frac{X_t}{Y_t}$.

Identifying the Girsanov Transformation

Assume the $\mathbb{Q}$-dynamics of S ($S = S^1$) are known as

$$\mathrm{d}S_t = r_t S_t \,\mathrm{d}t + S_t v_t \,\mathrm{d}W_t^{\mathbb{Q}}.$$

From slide 327 we know

$$L_t = \frac{S_t}{S_0 B_t},\ L_t = \frac{\mathrm{d}\mathbb{Q}^S}{\mathrm{d}\mathbb{Q}} \text{ on } \mathcal{F}_t.$$

From this we immediately have, assuming $\mathrm{d}B_t = r_t B_t \,\mathrm{d}t$,

$$\mathrm{d}L_t = L_t v_t \,\mathrm{d}W_t^{\mathbb{Q}}.$$

and we can summarize.

Theorem. The Girsanov kernel is given by the numéraire volatility v_t, i.e.,

$$\mathrm{d}L_t = L_t v_t \,\mathrm{d}W_t^{\mathbb{Q}}.$$

Application: zero coupon bonds

[*ZCBs already briefly seen on slide 273.*]

A *zero coupon T-bond* is a contract which gives you the claim

$$X \equiv 1 \text{ at time } T.$$

The price process $\Pi_t[1]$ is denoted by $p(t, T)$.

Allowing a stochastic, but adapted, short rate process r_t we have

$$\mathrm{d}B_t = r_t B_t \,\mathrm{d}t.$$

This gives us $B_t = e^{\int_0^t r_s \,\mathrm{d}s}$, and using standard risk neutral valuation we have

$$p(t, T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s \,\mathrm{d}s} \mid \mathcal{F}_t\right] = B_t \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{B_T} \mid \mathcal{F}_t\right].$$

Note. $p(T, T) = 1$.

The forward measure $\mathbb{Q}^T$, and a special numéraire

Here is the setting.

- Consider a fixed T.
- Choose the *bond price* process $p(t,T)$ as a numéraire.
- The corresponding martingale measure is denoted by $\mathbb{Q}^T$ and referred to as "the *T-forward measure*".

For any T-claim X, from the general theorem on slide 329,

$$\Pi_t[X] = p(t,T)\mathbb{E}^{\mathbb{Q}^T}\left[\frac{\Pi_T[X]}{p(T,T)} \mid \mathcal{F}_t\right],$$

$$\Pi_T[X] = X,\ p(T,T) = 1.$$

Theorem. For any T-claim X we have

$$\Pi_t[X] = p(t,T)\mathbb{E}^{\mathbb{Q}^T}[X \mid \mathcal{F}_t].$$

An application to a option pricing

Note. The latter formula for $\Pi_t[X]$ is (sometimes) more useful than the standard formule $\Pi_t[X] = B_t \mathbb{E}^{\mathbb{Q}}[\frac{X}{B_T} \mid \mathcal{F}_t]$, because of the expectation (now under $\mathbb{Q}^T$) of the single random variable X.

Application. Consider the European call $X = \max\{S_T - K, 0\}$.

We will give a formula for $\Pi_0[X]$ involving two numéraires, as well as the measure $\mathbb{Q}^S$ and the forward measure $\mathbb{Q}^T$. Write X as

$$X = (S_T - K) \cdot I_{\{S_T \geq K\}} = S_T I_{\{S_T \geq K\}} - K I_{\{S_T \geq K\}}.$$

Use $\mathbb{Q}^S$ on the first term and $\mathbb{Q}^T$ on the second to obtain

$$\Pi_0[X] = S_0 \cdot \mathbb{Q}^S(S_T \geq K) - K \cdot p(0,T) \cdot \mathbb{Q}^T(S_T \geq K).$$

This is not the final answer, as one would like to compute the two probabilities, for which the dynamics of S are needed.

$\longrightarrow$ **End of Lecture 9b** $\longleftarrow$

Lecture 10

Incomplete Markets

Chapter 15

Recall (in)completeness

- In a complete market every claim can be hedged, and thus has a unique price, the value of the replicating portfolio.
- The meta-theorem says that in case $N < R$, not sufficiently many assets (for hedging), the market is incomplete.
- What else can lead to incompleteness?

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

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

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 (catastrophe) to take place.

Problems

Weather derivatives.

The temperature is *not* the price of a traded asset.

Electricity derivatives.

Electric energy can*not* easily be stored.

CAT-bonds.

Natural disasters are *not* traded assets.

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

Typical Factor Model Setup

Given is

- an underlying factor process X, which is **not** the price process of a traded asset, with dynamics under the objective probability measure $\mathbb{P}$ as

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t;$$

- a risk free asset with dynamics

$$\mathrm{d}B_t = rB_t\,\mathrm{d}t.$$

Problem. Find arbitrage free price $\Pi_t[Z]$ of a derivative of the form

$$Z = \Phi(X_T).$$

Note the similarities *and* differences with the earlier set up, X is *not* the price of a *traded asset*.

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 $\Pi_t[\Phi]$, also denoted $\Pi_t(Z)$, uniquely determined by the $\mathbb{P}$-dynamics of X, and the requirement of an arbitrage free derivatives market?

Again a question that should make you suspicious.

Answer

NO!!

WHY?

Are there differences between the current and previous set ups?

Stock Price Model vs Factor Model

The two models looks similar.

Black-Scholes Model.

$$\begin{aligned} \mathrm{d}S_t &= \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \\ \mathrm{d}B_t &= rB_t \,\mathrm{d}t. \end{aligned}$$

Factor Model.

$$\begin{aligned} \mathrm{d}X_t &= \mu(t, X_t) \,\mathrm{d}t + \sigma(t, X_t) \,\mathrm{d}W_t, \\ \mathrm{d}B_t &= rB_t \,\mathrm{d}t. \end{aligned}$$

Question: What is the difference?

Answer

- X is **not the price of a traded asset!**
- We can not form a portfolio based on X, hence X is useless for hedging purposes.

The meta-theorem applied

- **Rule of thumb.**

  $N = 0$ (no risky asset),

  $R = 1$ (one source of randomness, W).

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

- **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.

Consequences

- There is **not** a unique price for a **particular** derivative; this is typical for incomplete markets.
- Still, 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. In a way, the benchmark replaces the underlying stock in the previous setting.

We consider two given claims $\Phi(X_T)$ and $\Gamma(X_T)$. We assume they are traded with prices, under the same underlying factor

$$\Pi_t[\Phi] = f(t, X_t),$$
$$\Pi_t[\Gamma] = g(t, X_t).$$

How are these prices related?

Set up, a program as before

- Form a self financing portfolio with value process V based on Φ and Γ. Use Itô on f and g to get portfolio dynamics. The self financing condition reads

$$\mathrm{d}V = V\left[u^f \frac{\mathrm{d}f}{f} + u^g \frac{dg}{g}\right],$$

  where u^f and u^g are the relative portfolio weights.
- Choose the portfolio weights such that the $\mathrm{d}W$-term vanishes. Then we have

$$\mathrm{d}V = V \cdot k\,\mathrm{d}t,$$

  a "synthetic bank" with k as the short rate.
- Absence of arbitrage implies

$$k = r.$$

- Read off the relation $k = r$!

The program in action I

Recall $\Pi_t[\Phi] = f(t, X_t)$. From Itô one obtains

$$\mathrm{d}f = f\mu_f \,\mathrm{d}t + f\sigma_f \,\mathrm{d}W,$$

where

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

Recall $\Pi_t[\Gamma] = g(t, X_t)$. Of course, there is then a similar expression for $\mathrm{d}g$. Write it down, as well as analogous expressions for μ^g and σ^g.

Recall the portfolio dynamics

$$\mathrm{d}V = V\left[u^f \frac{\mathrm{d}f}{f} + u^g \frac{dg}{g}\right].$$

Reshuffling terms gives us

$$\mathrm{d}V = V \cdot \left[u^f \mu_f + u^g \mu_g\right] \mathrm{d}t + V \cdot \left[u^f \sigma_f + u^g \sigma_g\right] \mathrm{d}W.$$

Our next aim is to annihilate the $\mathrm{d}W$-term.

The program in action II

Let the portfolio weights solve the system

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

The solution is (provided $\sigma_f \neq \sigma_g$, the claims should not coincide)

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

The portfolio dynamics then become

$$\begin{aligned} \mathrm{d}V &= V \cdot \left[u^f \mu_f + u^g \mu_g \right] \mathrm{d}t \\ &= V \cdot \left[\frac{\mu_g \sigma_f - \mu_f \sigma_g}{\sigma_f - \sigma_g} \right] \mathrm{d}t. \end{aligned}$$

The program in action III

Absence of arbitrage requires

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

which can be written as

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

Important observation. The quotients do **not** depend on the particular choice of contract, are the same for every f (and g).

This should be considered as the internal consistency relation alluded to on slide 351.

Of course, such a relation should also hold when pricing derivatives of a *traded asset* S, instead of X. Check with previous results whether this is true.

Market price of risk

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

Note. It is the same λ for all choices of f.

Alternative terminology for λ is

- Risk premium per unit of volatility,
- Sharpe Ratio.

Interpretation

The above result is concisely formulated as follows.

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

Observation. The relation

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

is actually a PDE! This can be seen by inspecting the expressions for μ_f and σ_f, that both involve partial derivatives of f.

Pricing Equation

The pricing function f solves the boundary problem (PDE)

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

Here the coefficients are related to the $\mathbb{P}$-dynamics of X,

$$\mathrm{d}X = \mu(t, X)\,\mathrm{d}t + \sigma(t, X)\,\mathrm{d}W.$$

Crucial question: can we solve this PDE?

$\longrightarrow$ End of Lecture 10a $\longleftarrow$

Answer

No!!

Why??

The answer explained

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. But

- λ is not given exogenously,
- λ is not determined endogenously.

This looks hopeless! And it leads to another question.

Question

Who determines λ?

Answer

THE MARKET!

Interpreting λ

Recall $\Pi_t[\Phi] = f(t, X_t)$ with f-dynamics $\mathrm{d}f = f\mu_f \,\mathrm{d}t + f\sigma_f \,\mathrm{d}W_t$, and that λ is defined as

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

Qualitative properties of λ are as follows.

- 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 neutral**.
- If you make an assumption about λ, then you implicitly make an assumption about the aggregate risk aversion of the market. And vice versa, you may learn about λ from market (agents) behaviour, or perhaps by analyzing a particular derivative.

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.
  - **Partly** by the requirement of an arbitrage free derivative market. **All** pricing functions satisfies the **same** PDE, depending on λ, of course with different boundary conditions depending on the derivative at hand.
  - **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 that is (implicitly) used by the market.

Risk Neutral Valuation

We recall from slide 358 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 theory we obtain a risk neutral valuation formula. Indeed, f is given by

$$f(t, x) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}_{t,x}[\Phi(X_T)],$$

coupled to the $\mathbb{Q}$-dynamics of X,

$$\mathrm{d}X_t = (\mu - \lambda\sigma)\,\mathrm{d}t + \sigma\,\mathrm{d}W^{\mathbb{Q}}_t.$$

Some technical additions

Think of a Girsanov transformation that relates $\mathbb{Q}$ to $\mathbb{P}$. The likelihood process L, $L_t = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}$ on $\mathcal{F}_t$, satisfies

$$\mathrm{d}L_t = L_t \phi_t \, \mathrm{d}W_t, \text{ and } \mathrm{d}W_t = \phi_t \, \mathrm{d}t + \mathrm{d}W_t^{\mathbb{Q}}.$$

If X satisfies

$$\mathrm{d}X_t = \mu_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}W,$$

then we also have

$$\mathrm{d}X_t = (\mu_t + \sigma_t \phi_t) \, \mathrm{d}t + \sigma_t \, \mathrm{d}W_t^{\mathbb{Q}}.$$

So, we conclude from the previous slide $\lambda = -\phi$ as on slide 356. See also slides 371 and 372 ahead.

Remarks and interpretation

As before, one has the following observations.

- The price is the expected discounted value of future payments.
- This expectation should **not** be taken under the "objective" probabilities $\mathbb{P}$, but under "risk adjusted" probabilities $\mathbb{Q}$.
- 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**, i.e., product dependent, not a uniform market property, and **diversifiable**, then one can argue that the factor should *not* be priced by the market. Compare with Arbitrage Pricing Theory (APT), e.g. CAPM.
- Mathematically this means that $\lambda = 0$, i.e., $\mathbb{P} = \mathbb{Q}$, see slide 366. Hence, **the risk neutral distribution coincides with the objective distribution**.
- We thus have what is called the **actuarial pricing formula**

$$f(t,x) = e^{-r(T-t)}\mathbb{E}^{\mathbb{P}}_{t,x}[\Phi(X_T)],$$

where we use the objective probability measure $\mathbb{P}$.

Modeling Issues I

Temperatur.

A standard model is given by the *Ornstein-Uhlenbeck* process

$$\mathrm{d}X_t = (m(t) - bX_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t,$$

where m is the mean temperature capturing seasonal variations. This often works reasonably well. As a result of this model, X_t has a normal distribution.

Electricity.

A (naive) model for the spot electricity price is

$$\mathrm{d}S_t = S_t(m(t) - a\log S_t)\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t$$

Using the Itô formula, one can show that $\log S_t$ behaves as an Ornstein-Uhlenbeck process, this implies lognormal prices. Electricity prices are however in reality very far from lognormal, because of "spikes" in the prices. More complicated models are needed, and used.

Modeling Issues II

CAT bonds.

Here we have to use the theory of point processes and the theory of extremal statistics to model natural disasters. The simplest point process is the *Poisson process*, also popular in queuing theory. This is outside the scope of this course.

Martingale Analysis

Model. Under $\mathbb{P}$ we have

$$\begin{aligned} \mathrm{d}X_t &= \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t, \\ \mathrm{d}B_t &= rB_t\,\mathrm{d}t, \end{aligned}$$

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

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

is a $\mathbb{Q}$-martingale, which is always the case.

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

Girsanov transformations

Compare this part also to slide 366. The usual Girsanov set up departs from

$$\mathrm{d}L_t = L_t \phi_t \, \mathrm{d}W_t.$$

We have the collective $\mathbb{P}$-dynamics

$$\begin{aligned} \mathrm{d}X_t &= \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t, \\ \mathrm{d}L_t &= L_t \phi_t \, \mathrm{d}W_t, \\ \mathrm{d}\mathbb{Q} &= L_t \mathrm{d}\mathbb{P} \text{ on } \mathcal{F}_t, \end{aligned}$$

and the relation

$$\mathrm{d}W_t = \phi_t \, \mathrm{d}t + \mathrm{d}W_t^{\mathbb{Q}}.$$

Result

Martingale pricing, and the Feynman-Kac formula, says that the pricing function F of a T-claim is given by

$$F(t,x) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,x}[Z],$$

using the $\mathbb{Q}$-dynamics of X, which are given by

$$\mathrm{d}X_t = (\mu(t,X_t) + \sigma(t,X_t)\phi_t)\,\mathrm{d}t + \sigma(t,X_t)\,\mathrm{d}W^{\mathbb{Q}}_t.$$

Result. We have $\lambda_t = -\phi_t$, i.e., the Girsanov kernel ϕ equals minus the market price of risk λ, as we have already seen on slide 365.

Several Risk Factors, an outlook

We recall the dynamics of the pricing function f (slide 353), $\mathrm{d}f = f\mu_f\,\mathrm{d}t + f\sigma_f\,\mathrm{d}W_t$, and the Market Price of Risk relation

$$\mu_f - r = \lambda\sigma_f.$$

In a *multifactor* model of the type (the W^i are independent Wiener processes)

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sum_{i=1}^{n} \sigma_i(t, X_t)\,\mathrm{d}W_t^i,$$

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

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

Compare with CAPM (*if you know what this is*).

$\longrightarrow$ **End of Lecture 10b** $\longleftarrow$

Stochastic Control Theory

Chapter 19

Contents, two main topics

- Dynamic programming, theory, mathematical background.
- Investment theory, economic application.

Lecture 11

Dynamic Programming

- The basic idea.
- Deriving the HJB equation.
- The verification Theorem.
- The linear quadratic (Gaussian) regulator, a classic topic in systems and control theory.

Optimal control problem, the formulation

The goal is to solve the maximization problem

$$\max_u \mathbb{E}\left[\int_0^T F(t, X_t, u_t)\,\mathrm{d}t + \Phi(X_T)\right],$$

subject to

$$\begin{aligned}
\mathrm{d}X_t &= \mu(t, X_t, u_t)\,\mathrm{d}t + \sigma(t, X_t, u_t)\,\mathrm{d}W_t,\\
X_0 &= x_0,\\
u_t &\in U(t, X_t),\ \forall t \in [0, T].
\end{aligned}$$

Remarks.

- The process X depends on the u_t, and is therefore also often denoted X^u.
- Often the set U is the same for all (t, x).
- We will only consider **feedback control laws**, i.e. controls of the form $u_t = \mathbf{u}(t, X_t)$, in which case X becomes a Markov process.

Terminology

Note. In this set up, there are no state space constraints, like e.g. $X_t \geq 0$. But they will appear later.

Terminology.

$$X = \text{state variable, } X_t \in \mathbb{R}^n,$$
$$u = \text{control variable, } u_t \in \mathbb{R}^k,$$
$$U = \text{control constraint, a subset of } \mathbb{R}^k.$$

The problem is called an *optimal control* problem.

Main idea

- Embed the problem above in a family of problems indexed by starting point in time and space. See slide 385.
- Tie all these problems together by a PDE, the Hamilton-Jacobi-Bellman (HJB) equation. See slide 398.
- The control problem is thus reduced to the problem of solving the deterministic HJB equation. This is in general rather hard, but at least it is "easier" than solving the original problem directly.

Some notation

This all looks rather messy, but try to see through!

- For any fixed vector $u \in \mathbb{R}^k$, the functions μ^u, σ^u and C^u (of three variables) are defined by (prime denotes transposition)

$$\begin{aligned}\mu^u(t,x) &= \mu(t,x,u),\\ \sigma^u(t,x) &= \sigma(t,x,u),\\ C^u(t,x) &= \sigma(t,x,u)\sigma(t,x,u)'.\end{aligned}$$

- For any *control law* $\mathbf{u}$ (recall $u_t = \mathbf{u}(t, X_t)$), the functions $\mu^{\mathbf{u}}$, $\sigma^{\mathbf{u}}$, $C^{\mathbf{u}}(t,x)$ and $F^{\mathbf{u}}(t,x)$ are defined by

$$\begin{aligned}\mu^{\mathbf{u}}(t,x) &= \mu(t,x,\mathbf{u}(t,x)),\\ \sigma^{\mathbf{u}}(t,x) &= \sigma(t,x,\mathbf{u}(t,x)),\\ C^{\mathbf{u}}(t,x) &= \sigma(t,x,\mathbf{u}(t,x))\sigma(t,x,\mathbf{u}(t,x))',\\ F^{\mathbf{u}}(t,x) &= F(t,x,\mathbf{u}(t,x)).\end{aligned}$$

Note that these functions are now functions of two variables.

More notation I

- For any fixed vector $u \in \mathbb{R}^k$, the partial differential operator $\mathcal{A}^u$ is defined by

$$\mathcal{A}^u = \sum_{i=1}^{n} \mu_i^u(t,x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} C_{ij}^u(t,x) \frac{\partial^2}{\partial x_i \partial x_j}.$$

  This operator is also called the *generator* of the Markov process X satisfying the SDE on slide 379. If f is a sufficiently differentiable function and $Y_t = f(t, X_t)$, then the Itô formula can be given in terms of the generator,

$$\mathrm{d}Y_t = (f_t(t,X_t) + \mathcal{A}^u f(t,X_t))\,\mathrm{d}t + f_x(t,X_t)\sigma^u(t,X_t)\,\mathrm{d}W_t.$$

- For any control law $\mathbf{u}$, the partial differential operator $\mathcal{A}^{\mathbf{u}}$ is defined by

$$\mathcal{A}^{\mathbf{u}} = \sum_{i=1}^{n} \mu_i^{\mathbf{u}}(t,x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} C_{ij}^{\mathbf{u}}(t,x) \frac{\partial^2}{\partial x_i \partial x_j}.$$

More notation II

- For any control law $\mathbf{u}$, the process $X^{\mathbf{u}}$ is the solution to the SDE

$$\mathrm{d}X_t^{\mathbf{u}} = \mu(t, X_t^{\mathbf{u}}, \mathbf{u}_t)\,\mathrm{d}t + \sigma(t, X_t^{\mathbf{u}}, \mathbf{u}_t)\,\mathrm{d}W_t,$$

where

$$\mathbf{u}_t = \mathbf{u}(t, X_t^{\mathbf{u}}).$$

This equation can be rewritten as

$$\mathrm{d}X_t^{\mathbf{u}} = \mu^{\mathbf{u}}(t, X_t^{\mathbf{u}})\,\mathrm{d}t + \sigma^{\mathbf{u}}(t, X_t^{\mathbf{u}})\,\mathrm{d}W_t,$$

which implies (under some technical conditions) that $X^{\mathbf{u}}$ is a Markov process. This property is a prerequisite to apply the *Dynamic Programming* technique that we will use the solve the optimal control problem.

Embedding the problem

We embed the optimal control problem of slide 379 into a family of problems $\mathcal{P}_{t,x}$. These are defined now.

For every fixed (t, x) the control problem $\mathcal{P}_{t,x}$ is defined as the problem to maximize, starting from the initial time t and initial state $X_t^{\mathbf{u}} = x$, the expectation

$$\mathbb{E}_{t,x}\left[\int_t^T F(s, X_s^{\mathbf{u}}, \mathbf{u}_s)\,\mathrm{d}s + \Phi(X_T^{\mathbf{u}})\right],$$

given the dynamics

$$\begin{aligned}\mathrm{d}X_s^{\mathbf{u}} &= \mu(s, X_s^{\mathbf{u}}, \mathbf{u}_s)\,\mathrm{d}s + \sigma(s, X_s^{\mathbf{u}}, \mathbf{u}_s)\,\mathrm{d}W_s,\ s \geq t,\\ X_t &= x,\end{aligned}$$

and the constraints

$$\mathbf{u}(s, y) \in U, \quad \forall (s, y) \in [t, T] \times \mathbb{R}^n.$$

The original problem was $\mathcal{P}_{0,x_0}$, i.e., $\mathcal{P}_{t,x}$ with $t = 0$, $x = x_0$,

The optimal value function

- The **value function** (the functions $\mathbf{u}$ belong to some set $\mathcal{U}$)

  $$\mathcal{J} : \mathbb{R}_+ \times \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}$$

  is defined by (recall that $x = X_t^{\mathbf{u}}$)

  $$\mathcal{J}(t, x, \mathbf{u}) = \mathbb{E}_{t,x}\left[\int_t^T F(s, X_s^{\mathbf{u}}, \mathbf{u}_s)\,\mathrm{d}s + \Phi(X_T^{\mathbf{u}})\right],$$

  given the dynamics above. The $X_s^{\mathbf{u}}$ will also depend on x, so one should write $X_s^{\mathbf{u},x}$, but this becomes notationally too heavy.
- Recall that our problem is a maximization problem. The **optimal value function**

  $$V : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$$

  is defined by

  $$V(t, x) = \sup_{\mathbf{u} \in \mathcal{U}} \mathcal{J}(t, x, \mathbf{u}),$$

  and note that $V(T, x) = \Phi(x)$.

Aim and assumptions

In the solution process, our aim is to derive a PDE for V, the HJB equation. To that end we assume the following.

- There exists an optimal control law $\hat{\mathbf{u}}$. This means that the supremum on slide 386 is attained for some $\hat{\mathbf{u}} \in \mathcal{U}$, which turns the supremum into a maximum.
- 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. We will make big steps, bypassing and ignoring mathematical delicacies that would require a finer analysis for a complete treatment.

Bellman's 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.

The principle can be paraphrased as "If the shortest route from Lisbon to Porto passes through Coimbra, then the shortest route from Coimbra to Porto is exactly that part of the first route." Try to understand.

Basic strategy

To derive the PDE do as follows.

- Fix $(t, x) \in (0, T) \times \mathbb{R}^n$.
- Choose a real number h (interpreted as a "small" time increment).
- Choose an arbitrary control law $\mathbf{u}$ on the time interval $[t, t + h]$.

Now define the control law $\mathbf{u}^\star$ by

$$\mathbf{u}^\star(s, y) = \begin{cases} \mathbf{u}(s, y), & (s, y) \in [t, t + h] \times \mathbb{R}^n \\ \hat{\mathbf{u}}(s, y), & (s, y) \in (t + h, T] \times \mathbb{R}^n. \end{cases}$$

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 $\hat{\mathbf{u}}$ law during the rest of the time period.

Note that $\mathbf{u}^\star$ is (by definition) "worse" than the optimal $\hat{\mathbf{u}}$ on the whole time interval $[t, T]$

Basic idea

The whole idea of Dynamic Programming 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 on slide 389.
- Compute (formally) the expected utilities obtained by the respective strategies.
- Using the *obvious* fact that $\hat{\mathbf{u}}$ is at least as good as $\mathbf{u}^\star$, and letting h tend to zero, we obtain, by heuristic reasoning, our fundamental PDE.

Intermezzo: a bit more on conditional expectations I

We have to optimize an expression which is a conditional expectation $\mathbb{E}_{t,x}$ of some functional, now *abbreviated* $G(t, T)$, which depends on values of a Markov process X in the future after time t, given that $X_t = x$.

We can also compute $\mathbb{E}[G(t, T)|\mathcal{F}_t]$. By the Markov property this becomes $\mathbb{E}[G(t, T)|X_t]$, which is a function of t and X_t, $g(t, X_t)$ say.

On the event $\{X_t = x\}$ the conditional expectation thus becomes $g(t, x)$. For this reason, we also use the notation $\mathbb{E}_{t,x}[G(t, T)]$ for $g(t, x)$.

Intermezzo: a bit more on conditional expectations II

Suppose we first compute $\mathbb{E}[G(t+h, T)|\mathcal{F}_{t+h}]$. This then becomes $g(t+h, X_{t+h})$, and then we compute $\mathbb{E}[G(t+h, T)|\mathcal{F}_t]$. By the law of iterated conditional expectations, we have

$$\mathbb{E}[G(t+h, T)|\mathcal{F}_t] = \mathbb{E}[\mathbb{E}[G(t+h, T)|\mathcal{F}_{t+h}]|\mathcal{F}_t] =$$

$$\mathbb{E}[g(t+h, X_{t+h})|\mathcal{F}_t] = \mathbb{E}[g(t+h, X_{t+h})|X_t].$$

When $X_t = x$ we thus have

$$\mathbb{E}_{t,x}[G(t+h, T)] = \mathbb{E}_{t,x}[g(t+h, X_{t+h})].$$

We apply this principle to compute the expected utility on the next slide, where G is a value function.

$\longrightarrow$ End of Lecture 11a $\longleftarrow$

Strategy values

I. Expected utility for $\hat{\mathbf{u}}$. From the definition of V, slide 386,

$$\mathcal{J}(t, x, \hat{\mathbf{u}}) = V(t, x).$$

II. Expected utility for $\mathbf{u}^\star$. Split the time interval $[t, T]$ as $[t, t+h) \cup [t+h, T]$ and compute partial expected utilities.

- The expected utility for $[t, t+h)$ is given by (no terminal utility!)

$$\mathbb{E}_{t,x}\left[\int_t^{t+h} F(s, X_s^{\mathbf{u}}, \mathbf{u}_s)\,\mathrm{d}s\right].$$

- The conditional expected utility over $[t+h, T]$ given (t, x) is

$$\mathbb{E}_{t,x}\left[V(t+h, X_{t+h}^{\mathbf{u}})\right],$$

upon realizing that application of a control $\mathbf{u}$ on $[t, t+h)$ results in a state $X_{t+h}^{\mathbf{u}}$ at time $t+h$. This is in fact an *iterated* conditional expectation, see slide 392.

Comparing strategies

The total expected utility for Strategy II thus becomes

$$V_{\mathrm{II}} := \mathbb{E}_{t,x}\left[\int_t^{t+h} F(s, X_s^{\mathbf{u}}, \mathbf{u}_s)\,\mathrm{d}s + V(t+h, X_{t+h}^{\mathbf{u}})\right].$$

We have trivially (the first strategy uses the optimal control)

$$V(t,x) \geq V_{\mathrm{II}} = \mathbb{E}_{t,x}\left[\int_t^{t+h} F(s, X_s^{\mathbf{u}}, \mathbf{u}_s)\,\mathrm{d}s + V(t+h, X_{t+h}^{\mathbf{u}})\right].$$

Remark. We have equality above if and only if the control law $\mathbf{u}$ is the optimal law $\hat{\mathbf{u}}$.

Towards HJB

Now use Itô to obtain

$$V(t+h, X^{\mathbf{u}}_{t+h}) = V(t,x) + \int_t^{t+h} \left(\frac{\partial V}{\partial t}(s, X^{\mathbf{u}}_s) + \mathcal{A}^{\mathbf{u}} V(s, X^{\mathbf{u}}_s) \right) \mathrm{d}s + \int_t^{t+h} V_x(s, X^{\mathbf{u}}_s) \sigma^{\mathbf{u}} \, \mathrm{d}W_s,$$

and plug into this the formula for V_{II} above.

Observe that the Itô integral has (under conditions, here we are not precise) conditional expectation given $\mathcal{F}_t$ equal to zero.

More heuristic reasoning

Use the inequality $V(t,x) \geq V_{\mathrm{II}}$ of slide 394 to obtain

$$\mathbb{E}_{t,x}\left[\int_t^{t+h}\left(F(s,X_s^{\mathbf{u}},\mathbf{u}_s)+\frac{\partial V}{\partial t}(s,X_s^{\mathbf{u}})+\mathcal{A}^{\mathbf{u}}V(s,X_s^{\mathbf{u}})\right)\,\mathrm{d}s\right]\leq 0.$$

Go to the limit. Divide by h, move h within the expectation and let h tend to zero. We thus get, using $X_t^{\mathbf{u}}=x$, the initial state for any $\mathbf{u}$, and that the expectation is not needed anymore,

$$F(t,x,u)+\frac{\partial V}{\partial t}(t,x)+\mathcal{A}^{u}V(t,x)\leq 0.$$

On the background, all this is partially justified by the fact

$$\frac{1}{h}\int_t^{t+h} g(s)\,\mathrm{d}s \to g(t) \text{ for continuous } g.$$

Last steps towards the HJB equation

Recall the inequality from the previous slide

$$F(t,x,u) + \frac{\partial V}{\partial t}(t,x) + \mathcal{A}^u V(t,x) \leq 0,$$

This holds for all $u = \mathbf{u}(t,x)$, with equality if and only if $u = \hat{\mathbf{u}}(t,x)$, which follows from the remark on slide 394.

We thus arrive at the **HJB equation**

$$\frac{\partial V}{\partial t}(t,x) + \sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u V(t,x)\} = 0.$$

The HJB equation

Theorem. Under suitable regularity assumptions the following hold.

I. V satisfies the Hamilton–Jacobi–Bellman equation

$$\frac{\partial V}{\partial t}(t,x) + \sup_{u \in U}\{F(t,x,u) + \mathcal{A}^u V(t,x)\} = 0,$$

$$V(T,x) = \Phi(x).$$

II. For each $(t,x) \in [0,T] \times \mathbb{R}^n$ the supremum in the HJB equation above is attained by $u = \hat{\mathbf{u}}(t,x)$, i.e., by the optimal control.

Remark. A full, mathematically sound, proof is rather hard. Not given!

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 equation is thus derived as a **necessary** condition, and requires strong *ad hoc* regularity assumptions, alternatively the use of viscosity solutions techniques (this is advanced math).

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** below.

Consequence. Application of the Verification Theorem gives you the solution to the optimal control problem, at least in principle.

The Verification Theorem, assumptions

Theorem. Suppose that we have two functions $H(t,x)$ and $g(t,x)$, such that the following requirements are met.

- H is sufficiently integrable, and solves the HJB equation

$$\begin{cases} \dfrac{\partial H}{\partial t}(t,x) + \sup\limits_{u \in U}\{F(t,x,u) + \mathcal{A}^u H(t,x)\} = 0, \\ H(T,x) = \Phi(x). \end{cases}$$

- Consider the following static setting. 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 $u = g(t,x)$.

The Verification Theorem, assertions

Under the stipulated assumptions the following hold.

- The optimal value function V to the original optimal control problem is given by

$$V(t,x) = H(t,x),$$

  with H the assumed solution to the HJB equation.
- There exists an optimal control law $\hat{\mathbf{u}}$, and in fact

$$\hat{\mathbf{u}}(t,x) = g(t,x).$$

Proof, relatively simple, in a separate document, and in the book on pages 291–293.

Handling the HJB equation I; book Section 19.4

Here are a few steps to solve the HJB equation for V.

- Fix $(t, x) \in [0, T] \times \mathbb{R}^n$ and solve the static optimization problem of slide 400 in the Verification Theorem, assuming that a maximizer exists,

$$\max_{u \in U} \{F(t, x, u) + \mathcal{A}^u V(t, x)\} z.$$

  Here u is the only variable, whereas t and x are fixed parameters. The functions F, μ, σ and V are considered as given.

- 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

$$\hat{\mathbf{u}} = \hat{\mathbf{u}}(t, x; V). \tag{1}$$

Handling the HJB equation II

- 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 (!) PDE

$$\frac{\partial V}{\partial t}(t,x) + F^{\hat{\mathbf{u}}}(t,x) + \mathcal{A}^{\hat{\mathbf{u}}}(t,x)V(t,x) = 0,$$
$$V(T,x) = \Phi(x).$$

- Now we solve the PDE above, hoping that we are able to do so! Then we put the solution V into expression (1) for $\hat{\mathbf{u}}$. Using the verification theorem we can identify V as the optimal value function, and $\hat{\mathbf{u}}$ as the optimal control law.

A remaining question. These steps outline a nice guideline, but can you really perform them in concrete situations?

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, based on experience, one usually tries to **guess** a solution. I.e., we typically make a parametrized **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. This is the experienced.
- Most of the *known* solved control problems have, to some extent, been "rigged" in order to be analytically solvable.

The Linear Quadratic Regulator, Section 19.5

We now consider a classical problem in systems and control theory. All processes below are $\mathbb{R}^1$-valued, but there exist straightforward extension to multidimensional situations. The problem is a minimization problem, as opposed to the previous maximization problem. It is the following, also known as the LQG problem.

$$\min_{u_t \in \mathbb{R}} \mathbb{E}\left[\int_0^T (QX_t^2 + Ru_t^2)\,\mathrm{d}t + HX_T^2\right],$$

with dynamics

$$\mathrm{d}X_t = (AX_t + Bu_t)\,\mathrm{d}t + C\,\mathrm{d}W_t.$$

Engineering interpretation. Control a vehicle in such a way that it stays close to the origin (small Qx^2 and Hx^2) while at the same time keeping the for this purpose needed "energy" Ru^2 small.

The real numbers Q, R, H, A, B and C are assumed to be known. We assume that R is strictly positive.

Intermezzo, some generators

Suppose that a process X satisfies the SDE of the type

$$\mathrm{d}X_t = \mu^u(t, X_t)\,\mathrm{d}t + \sigma^u(t, X_t)\,\mathrm{d}W_t.$$

The generator for this SDE acting on a function $f(t, x)$ is described by

$$\mathcal{A}^u f(t, x) = f_x(t, x)\mu^u(t, x) + \tfrac{1}{2} f_{xx}(t, x)\sigma^u(t, x)^2,$$

where f_x denotes partial derivative w.r.t. x, etc.

If the SDE is linear,

$$\mathrm{d}X_t = (AX_t + Bu_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t,$$

then

$$\mathcal{A}^u f(t, x) = f_x(t, x)(Ax + Bu) + \tfrac{1}{2} f_{xx}(t, x)\sigma^2.$$

Handling the Problem

The HJB equation becomes, use the general formula on slide 398, the expression for the generator on slide 406 and not the inf instead of sup,

$$\frac{\partial V}{\partial t}(t,x) + \inf_{u\in\mathbb{R}} \left\{ Qx^2 + Ru^2 + V_x(t,x)(Ax+Bu) \right\} + \frac{1}{2}\frac{\partial^2 V}{\partial x^2}(t,x)C^2 = 0,$$

$$V(T,x) = Hx^2.$$

For each fixed choice of (t,x) we now have to solve the static unconstrained optimization problem, the minimization in u of

$$p(u) := Qx^2 + Ru^2 + V_x(t,x)(Ax+Bu).$$

Note that $p(u)$ is quadratic in u and that a minimum exists since $R > 0$.

Finding the optimal u

To find the minimum we put $p'(u) = 0$. This gives

$$2Ru + V_x B = 0.$$

This gives us the optimal $\hat{u}$ as

$$\hat{u} = -\frac{1}{2}\frac{B}{R}V_x.$$

Note. This is our candidate of optimal control law, but it (still) 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 time dependent, but not depending on x. Note that $V(T, x) = Hx^2$, which yields the boundary conditions

$$P(T) = H,\ q(T) = 0.$$

Working with the guess for V

With this trial solution for V we have (dots denote derivatives w.r.t. t),

$$\begin{aligned}\frac{\partial V}{\partial t}(t,x) &= \dot{P}(t)x^2 + \dot{q}(t),\\ V_x(t,x) &= 2P(t)x,\\ V_{xx}(t,x) &= 2P(t).\end{aligned}$$

This gives

$$\hat{u}(t,x) = -\frac{B}{R}P(t)x.$$

Inserting these expressions into the HJB equation we get

$$x^2\big(\dot{P} + Q - \frac{B^2}{R}P^2 + 2AP\big) + \dot{q} + PC^2 = 0.$$

This equation has to hold for all t and *for all* x.

Therefore the part with x^2 has to be zero, as well as the part without x. This leads to two ODEs with boundary conditions,

$$\dot{P} = \frac{B^2}{R} P^2 - 2AP - Q, \; P(T) = H,$$
$$\dot{q} = -C^2 P, \; q(T) = 0.$$

The ODE for P is a **Riccati equation**, for which it is quite some work to write down the explicit solution. The equation for q can then be integrated directly. This gives $q(t) = C^2 \int_t^T P(s)\,\mathrm{d}s$.

Final Result for LQG.

$$V(t,x) = P(t)x^2 + \int_t^T C^2 P(s)\,\mathrm{d}s,$$
$$\hat{\mathbf{u}}(t,x) = -\frac{B}{R} P(t)x, \text{ a } \textit{linear (in } x\textit{) feedback law.}$$

$\longrightarrow$ **End of Lecture 11b** $\longleftarrow$

Lecture 12

Back to finance: Investment Theory, Section 19.6

- Problem formulation.
- An extension of HJB.
- The simplest consumption-investment problem.
- The Merton fund separation results, illustrations.

Recap of basic facts and a new concept

We consider a market with n assets.

S_t^i = price of asset No. i,

h_t^i = units of asset No. i in portfolio,

w_t^i = portfolio weight on asset No. i (previously u_t^i),

X_t = portfolio value, (previously V_t).

In addition to these familiar quantities there is also *consumption*, a new concept, as introduced now.

$C_t = \int_0^t c_s \, \mathrm{d}s$, *cumulative* consumption over the interval $[0, t]$,

c_t = the consumption *rate*, assumed to be positive.

Furthermore, we have the familiar relations

$$X_t = \sum_{i=0}^{n} h_t^i S_t^i, \qquad w_t^i = \frac{h_t^i S_t^i}{X_t}, \qquad \sum_{i=0}^{n} w_t^i = 1.$$

Self financing in presence of consumption

The dynamics of a self financing portfolio in terms of relative weights, and when *consumption* takes place, are *defined* by

$$\mathrm{d}X_t = X_t \sum_{i=0}^{n} w_t^i \frac{\mathrm{d}S_t^i}{S_t^i} - c_t \,\mathrm{d}t.$$

Note the *minus* in front of the c_t, as opposed to a plus sign when dealing with dividends. Recall from slide 250, partly in the notation used for the value process there, that in the dividend case

$$\mathrm{d}V_t = V_t \sum_{i=0}^{n} w_t^i \frac{\mathrm{d}S_t^i}{S_t^i} + V_t \sum_{i=0}^{n} w_t^i \frac{\mathrm{d}D_t^i}{S_t^i},$$

where D^i stood for dividend. The dividend flow contributed positively to the wealth derived from the portfolio, whereas in the current setting, consumption decreases wealth as one spends income on it.

Simplest model, $n = 1$

Assume a scalar risky asset and a constant short rate r.

$$\begin{aligned} \mathrm{d}S_t &= \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \\ \mathrm{d}B_t &= rB_t \,\mathrm{d}t. \end{aligned}$$

Problem. Maximize expected utility of consumption over time, a maximization problem of the form

$$\max_{w^0, w^1, c} \mathbb{E}\left[\int_0^T F(t, c_t)\,\mathrm{d}t\right].$$

The quantities w^0, w^1, c (time varying) are now viewed as *control variables*. The aim is to apply Dynamic Programming. The self financing dynamics of X can be recast as (verify!)

$$\mathrm{d}X_t = (w_t^0 r + w_t^1 \alpha)X_t \,\mathrm{d}t - c_t \,\mathrm{d}t + w_t^1 \sigma X_t \,\mathrm{d}W_t.$$

Question, related to the constraints

The constraints on the variables are

$$c_t \geq 0, \quad \forall t \geq 0,$$
$$w_t^0 + w_t^1 = 1, \quad \forall t \geq 0.$$

The following question arises.

Is the maximization problem a sensible one?

Again a suspicious question!

Answer

Nonsense!

Some issues

What are the problems?

- Without further restrictions we can obtain unlimited utility by simply consuming arbitrary large amounts.
- The wealth may go negative, and there is nothing in the problem formulation which prohibits this.
- We would like to impose a constraint of type $X_t \geq 0$ but this is a **state constraint** and Dynamic Programming (leading to the HJB equation) does not allow us to do this, see slide 380.

Good News. Dynamic Programming can be generalized to handle (some) problems of this kind.

The use of *stopping times* helps us here.

Intermezzo, stopping times

Stopping times are random times, so random variables. So, more formally, they are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, that is endowed with a filtration $\{\mathcal{F}_t, t \geq 0\}$.

A stopping time is a random variable $\tau : \Omega \to [0, \infty]$ with the property that $\{\tau \leq t\} \in \mathcal{F}_t$, for all $t \geq 0$.

The interpretation is that for all t one can decide whether $\tau \leq t$ has occurred on the basis of the available information at time t. We will hardly use this formalism, only a bit as a vehicle to derive certain results.

The most relevant example is that τ is the first moment that a certain stochastic process passes a threshold, for instance the moment that bankruptcy occurs. In this case the event $\{\tau = \infty\}$ means that bankruptcy never occurs.

Generalized optimal control problem

Let D be a "nice" *open subset* of $[0, T] \times \mathbb{R}^n$ and consider the following problem. Note the **stopping time** τ in the formulation of the problem, which acts as a horizon.

$$\max_{u \in U} \mathbb{E}\left[\int_0^\tau F(s, X_s^{\mathbf{u}}, \mathbf{u}_s) ds + \Phi(\tau, X_\tau^{\mathbf{u}})\right].$$

The dynamics of X are (as before)

$$\begin{aligned} \mathrm{d}X_t &= \mu(t, X_t, u_t)\,\mathrm{d}t + \sigma(t, X_t, u_t)\,\mathrm{d}W_t, \\ X_0 &= x_0 \in D. \end{aligned}$$

The stopping time τ is defined by

$$\tau = \inf\{t \geq 0 \mid (t, X_t) \in \partial D\} \wedge T.$$

Here ∂D denotes the *boundary* of D (see examples later on), and the wedge $\wedge$ denotes minimum, $a \wedge b = \min\{a, b\}$. Note that one has $\tau \leq T$.

Generalized HJB equation

Theorem. Given enough regularity the following hold.

1. The optimal value function satisfies

$$\begin{cases} \dfrac{\partial V}{\partial t}(t,x) + \sup\limits_{u\in U}\{F(t,x,u) + \mathcal{A}^u V(t,x)\} = 0, & \forall (t,x) \in D, \\ V(t,x) = \Phi(t,x), & \forall (t,x) \in \partial D. \end{cases}$$

   Note the boundary condition, it is the counterpart of $V(T,x) = \Phi(x)$ encountered before.

2. We have an "*obvious*" verification theorem, that can be obtained from the one we have seen by obvious adjustments. For instance, on slide 400 one replaces $H(T,x) = \phi(x)$ with $H(t,x) = \Phi(t,x)$, for all $(t,x) \in \partial D$.

Reformulated problem

We now consider the problem

$$\max_{c\geq 0, w\in\mathbb{R}} \mathbb{E}\left[\int_0^\tau F(t,c_t)\,\mathrm{d}t + \Phi(X_T)\right],$$

where the "ruin time" of "bankruptcy time" τ is defined by

$$\tau = \inf\{t \geq 0 \mid X_t = 0\} \wedge T.$$

This corresponds to $D = [0,T]\times(0,\infty)$. The boundary of D consists of the (t,x) such that $t = T$ or $x = 0$, so $\partial D = \{T\}\times[0,\infty)\cup[0,T]\times\{0\}$.

Furthermore we used the adjusted notation $w_t^1 = w_t, w_t^0 = 1 - w_t$. Thus there is no constraint on w_t anymore.

The dynamics of X on slide 415 become, with c_t and w_t as control variables,

$$\mathrm{d}X_t = w_t(\alpha - r)X_t\,\mathrm{d}t + (rX_t - c_t)\,\mathrm{d}t + w_t\sigma X_t\,\mathrm{d}W_t.$$

HJB Equation

For simplicity we assume that $\Phi = 0$, which gives the boundary conditions below.

$$\frac{\partial V}{\partial t} + \sup_{c \geq 0, w \in \mathbb{R}} \left\{ (F(t,c) + wx(\alpha - r)\frac{\partial V}{\partial x} + (rx - c)\frac{\partial V}{\partial x} + \frac{1}{2}x^2 w^2 \sigma^2 \frac{\partial^2 V}{\partial x^2} \right\} = 0,$$

$$V(T, x) = 0,$$

$$V(t, 0) = 0.$$

For an illustrative example we now specialize to

$$F(t,c) = e^{-\delta t} c^\gamma.$$

So we have to maximize, w.r.t. $c \geq 0$ and $w \in \mathbb{R}$,

$$G(c,w) := e^{-\delta t} c^\gamma + wx(\alpha - r)\frac{\partial V}{\partial x} + (rx - c)\frac{\partial V}{\partial x} + \frac{1}{2}x^2 w^2 \sigma^2 \frac{\partial^2 V}{\partial x^2}.$$

Analysis of the HJB Equation I

The first order conditions for this problem, $\frac{\partial G}{\partial c} = 0$, $\frac{\partial G}{\partial w} = 0$, yield

$$\gamma c^{\gamma-1} = e^{\delta t} V_x,$$
$$w = -\frac{V_x}{x \cdot V_{xx}} \cdot \frac{\alpha - r}{\sigma^2}.$$

We make the **Ansatz**, resembling $F(t, c)$,

$$V(t, x) = e^{-\delta t} h(t) x^{\gamma}.$$

Because of the boundary conditions, recall $\Phi = 0$, we must demand that

$$h(T) = 0. \tag{5}$$

Analysis of the HJB Equation II

Given a V of this form we have (using the dot to denote the time derivative),

$$\begin{aligned} V_t(t,x) &= e^{-\delta t}\dot{h}(t)x^\gamma - \delta e^{-\delta t}h(t)x^\gamma, \\ V_x(t,x) &= \gamma e^{-\delta t}h(t)x^{\gamma-1}, \\ V_{xx}(t,x) &= \gamma(\gamma-1)e^{-\delta t}h(t)x^{\gamma-2}. \end{aligned}$$

This gives us from the first order conditions on slide 424,

$$\begin{aligned} \hat{w}(t,x) &= \frac{\alpha - r}{\sigma^2(1-\gamma)}, \text{ a constant!}, \\ \hat{c}(t,x) &= xh(t)^{-1/(1-\gamma)}, \text{ linear in } x. \end{aligned}$$

Plug all this into HJB on slide 424 and try to solve.

Analysis of the HJB Equation III

After rearrangements (tedious work!) 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 (more tedious work)

$$A = \frac{\gamma(\alpha - r)^2}{2\sigma^2(1-\gamma)} + r\gamma - \delta, \; B = 1 - \gamma.$$

The equation above is to hold for all x, t, so h must solve the ODE

$$\dot{h}(t) + Ah(t) + Bh(t)^{-\gamma/(1-\gamma)} = 0, \; h(T) = 0.$$

An equation of this kind is known as a **Bernoulli equation**, and it can be solved explicitly. See Exercises 19.2, 19.3. This gives $h(t)$ and then also $V(t, x)$.

$\longrightarrow$ **End of Lecture 12a** $\longleftarrow$

Merton's Mutual Fund Theorems

Section 19.7

1. The case without a risk free asset.
2. The case with one risk free asset.

The case without a risk free asset, set up

We consider n risky assets with dynamics

$$\mathrm{d}S_i = S_i\alpha_i\,\mathrm{d}t + S_i\sigma_i\,\mathrm{d}W, \quad i = 1, \ldots, n,$$

where W is Wiener in $\mathbb{R}^k$, $\sigma_i \in \mathbb{R}^{1\times k}$. In vector form,

$$\mathrm{d}S = D(S)\alpha\,\mathrm{d}t + D(S)\sigma\,\mathrm{d}W.$$

Here we used

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{pmatrix} \in \mathbb{R}^{n\times k},$$

and $D(S) = \mathrm{diag}(S_1, \ldots, S_n)$, a diagonal matrix.

Formal problem

The formal problem, here the final reward at τ is taken zero, is

$$\max_{c,w} \mathbb{E}\left[\int_0^\tau F(t, c_t)\,\mathrm{d}t\right],$$

given the dynamics, recall the self financing condition on slide 414,

$$\mathrm{d}X = Xw'\alpha\,\mathrm{d}t - c\,\mathrm{d}t + Xw'\sigma\,\mathrm{d}W,$$

in addition to the dynamics of the S^i. The constraints are

$$e'w_t = 1, \text{ with } e' = (1, \ldots, 1),\ w_t' = (w_t^1, \ldots, w_t^n),\ c_t \geq 0.$$

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. This implies an arbitrage free and complete market when $n = k$.

Intermezzo: Multiple Wiener processes

Let W be a k-dimensional standard Wiener process. By definition this means that every component W^i is a one-dimensional Wiener process and that the components are independent. Consider the column vector of differentials $\mathrm{d}W_t$ and its transpose $\mathrm{d}W_t'$. One can show that, with I_k the k-dimensional identity matrix,

$$\mathrm{d}W_t \mathrm{d}W_t' = I_k \, \mathrm{d}t.$$

Let now a one-dimensional process X satisfy, for some a_t,

$$\mathrm{d}X_t = a_t \, \mathrm{d}t + b(X_t)\sigma \, \mathrm{d}W_t.$$

Then, with $\Sigma = \sigma\sigma'$,

$$\begin{aligned}(\mathrm{d}X_t)^2 &= b(X_t)\sigma \mathrm{d}W_t \mathrm{d}W_t' \sigma' b(X_t)' \\ &= b(X_t)\sigma I_k \sigma' b(X_t)' \mathrm{d}t \\ &= b(X_t)\Sigma b(X_t)' \mathrm{d}t.\end{aligned}$$

Solution I

Recall from slide 383 the generator $\mathcal{A}^u$ for a general control u. Here we have $\mathcal{A}^{c,w}$ and the HJB equation is now

$$\begin{cases} V_t(t,x) + \sup_{e'w=1, c\geq 0} \{F(t,c) + \mathcal{A}^{c,w} V(t,x)\} = 0, \\ V(T,x) = 0, \\ V(t,0) = 0. \end{cases}$$

where, see slide 430 for more explanation,

$$\mathcal{A}^{c,w} V = xw'\alpha V_x - cV_x + \frac{1}{2}x^2 w'\Sigma w V_{xx} \text{ with } \Sigma = \sigma\sigma'.$$

Of course, the optimal c, w will depend on t, x, V.

Solution II

The HJB equation is

$$\begin{cases} V_t + \sup_{w'e=1, c\geq 0} \left[F(t,c) + (xw'\alpha - c)V_x + \frac{1}{2}x^2 w'\Sigma w V_{xx} \right] = 0, \\ V(T,x) = 0, \\ V(t,0) = 0. \end{cases}$$

We relax the constraint $w'e = 1$ for the static optimization

problem, by applying a standard technique, the use of a Lagrange function $\mathcal{L}$. This function is given by

$$\mathcal{L}(c,w,\lambda) = F(t,c) + (xw'\alpha - c)V_x + \frac{1}{2}x^2 w'\Sigma w V_{xx} + \lambda(1 - w'e).$$

Solution III

We maximize $\mathcal{L}$ w.r.t. its three variables.

The first order condition for c is

$$F_c = V_x.$$

The first order condition for the vector w is

$$x\alpha V_x + x^2 V_{xx}\Sigma w = \lambda e,$$

where we used that the gradient of $w'\Sigma w$ is $2\Sigma w$.

We can solve for w and obtain the column vector

$$\hat{w} = \Sigma^{-1}\left(\frac{\lambda}{x^2 V_{xx}}e - \frac{xV_x}{x^2 V_{xx}}\alpha\right).$$

The relation $e'\hat{w} = 1$ gives λ as

$$\lambda = \frac{x^2 V_{xx} + xV_x e'\Sigma^{-1}\alpha}{e'\Sigma^{-1}e},$$

Solution IV

Inserting λ in the expression for $\hat{w}$ gives us, after "some manipulation",

$$\hat{w} = \frac{1}{e'\Sigma^{-1}e}\Sigma^{-1}e + \frac{V_x}{xV_{xx}}\Sigma^{-1}\left(\frac{e'\Sigma^{-1}\alpha}{e'\Sigma^{-1}e}e - \alpha\right).$$

The ratio $\frac{V_x}{xV_{xx}}$ depends on t. Hence the expression for $\hat{w}$ is of the form

$$\hat{w}(t) = g + Y(t)h,$$

where the fixed vectors g and h are given by

$$g = \frac{1}{e'\Sigma^{-1}e}\Sigma^{-1}e,$$

$$h = \Sigma^{-1}\left(\frac{e'\Sigma^{-1}\alpha}{e'\Sigma^{-1}e}e - \alpha\right),$$

whereas $Y(t)$ is given by the time dependent random expression

$$Y(t) = \frac{V_x(t, X(t))}{X(t)V_{xx}(t, X(t))}.$$

Solution V

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 "portfolio hyperplane" $\{w \in \mathbb{R}^n \mid e'w = 1\}$.

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. So, 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

Theorem. There exists a family of *mutual funds*, given by $w^s = g + sh$, $\in \mathbb{R}$, such that

(i) for each fixed s the portfolio w^s stays fixed over time;

(ii) for fixed a, b with $a \neq b$ the optimal portfolio $\hat{w}(t)$ is, obtained by allocating all resources between the fixed funds w^a and w^b, i.e.,

$$\hat{w}(t) = \mu^a(t)w^a + \mu^b(t)w^b,$$
$$\mu^a(t) = \frac{Y(t) - b}{b - a},\ \mu^b(t) = 1 - \mu^a(t).$$

Remark. To obtain the *geometric interpretation*, one doesn't need to precisely know the utility function F, nor the optimal value function V. Surprising!

The case with a risk free asset, set up

Again we consider the model of slide 428,

$$\mathrm{d}S = D(S)\alpha\,\mathrm{d}t + D(S)\sigma\,\mathrm{d}W(t),$$

and also assume the risk free asset B with dynamics

$$\mathrm{d}B = rB\,\mathrm{d}t.$$

We denote $B = S^0$ and consider portfolio weights $w^0, w^1, \ldots, w^n$, where $\sum_{i=0}^{n} w^i = 1$. We eliminate w^0 by the relation

$$w^0 = 1 - \sum_{i=1}^{n} w^i,$$

and denote

$$w = (w^1, \ldots, w^n)'.$$

Note. We now have $w \in \mathbb{R}^n$ without constraints. Hence, no Laplace function is needed.

HJB equation

We obtain from the self financing condition the dynamics

$$\mathrm{d}X = Xw'(\alpha - re)\,\mathrm{d}t + (rX - c)\,\mathrm{d}t + Xw'\sigma\,\mathrm{d}W,$$

where $e = (1, 1, \ldots, 1)'$.

The HJB equation now becomes

$$\begin{cases} V_t(t,x) + \sup\limits_{c\geq 0, w\in\mathbb{R}^n} \{F(t,c) + \mathcal{A}^{c,w}V(t,x)\} = 0, \\ V(T,x) = 0, \\ V(t,0) = 0, \end{cases}$$

where, use again that $\sigma\sigma' = \Sigma$,

$$\begin{aligned} \mathcal{A}^{c,w}V = {} & xw'(\alpha - re)V_x(t,x) + (rx - c)V_x(t,x) \\ & + \frac{1}{2}x^2 w'\Sigma w V_{xx}(t,x). \end{aligned}$$

First order conditions

We maximize

$$F(t,c) + xw'(\alpha - re)V_x + (rx - c)V_x + \frac{1}{2}x^2 w'\Sigma w V_{xx}$$

w.r.t. $c \geq 0$ and $w \in \mathbb{R}^n$.

The first order conditions for the variables c and W lead to

$$F_c = V_x,$$
$$\hat{w} = -\frac{V_x}{xV_{xx}}\Sigma^{-1}(\alpha - re),$$

with a geometrically obvious economic interpretation of the kind as on slide 435.

Mutual Fund Separation Theorem

Theorem. The optimal portfolio satisfies the following properties.

(i) The optimal portfolio consists of an allocation between two fixed mutual funds w^0 and w^f.

(ii) The fund w^0 consists *only* of the risk free asset.

(iii) The fund w^f consists *only* of the risky assets, and is given by

$$w^f = \Sigma^{-1}(\alpha - re).$$

The relative allocations of wealth are

$$\mu_t^f = -\frac{V_x(t, X_t)}{X_t V_{xx}(t, X_t)}, \ \mu_t^0 = 1 - \mu_t^f.$$

Remark. These results don't follow from those for the case without riskless asset. See the book.

$\longrightarrow$ **End of Lecture 12b** $\longleftarrow$

Lecture 13

The Martingale Approach to Optimal Investment Theory

Chapter 20

Contents

- Completeness of the market, an essential assumption.
- Decoupling the wealth profile from the portfolio choice.
- Lagrange relaxation.
- Solving the general wealth problem.
- Example: Log utility.
- Example: The numéraire portfolio.

Problem Formulation

We allow a time dependent version of the model on slide 437,

$$\begin{aligned} \mathrm{d}S_t &= D(S_t)\alpha_t\,\mathrm{d}t + D(S_t)\sigma_t\,\mathrm{d}W_t, \\ \mathrm{d}B_t &= rB_t\,\mathrm{d}t. \end{aligned}$$

Assumptions.

- Drift and diffusion terms are adapted processes.
- The market is **complete**, i.e., $R = N$ in the meta-theorem.
- We have a given initial wealth x_0.

Problem. Maximize *terminal wealth*,

$$\max_{h\in\mathcal{H}} \mathbb{E}^{\mathbb{P}}[\Phi(X_T^h)],$$

where X_T^h is the terminal wealth arising from a portfolio process h, and $\mathcal{H}$ is the set of self financing portfolios.

Some observations

- In a complete market, there is a unique martingale measure $\mathbb{Q}$.
- Every claim Z satisfying the budget constraint

$$e^{-rT}\mathbb{E}^{\mathbb{Q}}[Z] = x_0,$$

  is attainable by a self financing $h \in \mathcal{H}$ and vice versa.

To see the latter, argue as follows. As Z is attainable, we have the pricing rule $\Pi_0[Z] = e^{-rT}\mathbb{E}^{\mathbb{Q}}[X_T^h]$, where X^h is the value process (wealth process) associated with the replicating portfolio h.

But, as the discounted version of X^h is a $\mathbb{Q}$-martingale, $\mathbb{E}^{\mathbb{Q}}[e^{-rT}X_T^h] = X_0^h$, which proves the claim if one chooses $x_0 = X_0^h$.

Reformulation of the problem

- Given that hedging is possible, we can thus write our problem as

$$\max_{Z} \mathbb{E}^{\mathbb{P}}[\Phi(Z)]$$

subject to the constraint

$$e^{-rT}\mathbb{E}^{\mathbb{Q}}[Z] = x_0.$$

- We can forget the wealth dynamics! At least for the time being, see Step 2 below.

Note that the maximization problem is under the real world measure $\mathbb{P}$ and that the budget constraint is described under $\mathbb{Q}$.

Basic Ideas

Recall our problem,

$$\max_{Z} \mathbb{E}^{\mathbb{P}}[\Phi(Z)] \text{ subject to } e^{-rT}\mathbb{E}^{\mathbb{Q}}[Z] = x_0.$$

Idea I. We can **decouple** the optimal portfolio problem into steps.

1. Finding the optimal wealth profile $\hat{Z}$.
2. Given $\hat{Z}$, find the replicating portfolio. Here the dynamics will come in.

Idea II.

- Rewrite the constraint under the measure $\mathbb{P}$ instead of $\mathbb{Q}$.
- Use *Lagrangian techniques* to relax the constraint.

The budget constraint

First we rewrite the budget constraint $e^{-rT}\mathbb{E}^{\mathbb{Q}}[Z] = x_0$ as an expectation under $\mathbb{P}$ by using the R-N derivative $L_T = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}$ on $\mathcal{F}_T$. As we have

$$\mathbb{E}^{\mathbb{Q}}[Z] = \mathbb{E}^{\mathbb{P}}[L_T Z],$$

our problem becomes

$$\max_Z \mathbb{E}^{\mathbb{P}}[\Phi(Z)] \text{ subject to } e^{-rT}\mathbb{E}^{\mathbb{P}}[L_T Z] = x_0.$$

Recall the likelihood process L. It satisfies $L_t = \mathbb{E}^{\mathbb{P}}[L_T|\mathcal{F}_t]$ and L_t is the R-N derivative $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}$ on $\mathcal{F}_t$, $0 \leq t \leq T$.

Lagrange formulation

The Lagrangian of the problem is

$$\mathcal{L} = \mathbb{E}^{\mathbb{P}}[\Phi(Z)] + \lambda \left[x_0 - e^{-rT} \mathbb{E}^{\mathbb{P}}[L_T Z] \right],$$

and note that both expectations are under $\mathbb{P}$. We rewrite the Lagrangian as a single expectation,

$$\mathcal{L} = \mathbb{E}^{\mathbb{P}} \left[\Phi(Z) - \lambda e^{-rT} L_T Z \right] + \lambda x_0.$$

Here the additional variable λ is called a *dual variable*.

The optimal wealth profile I

Given enough convexity and regularity we now expect, given λ, to find the optimal Z by maximizing

$$\mathcal{L} = \mathbb{E}^{\mathbb{P}}\left[\Phi(Z) - \lambda e^{-rT} L_T Z\right] + \lambda x_0$$

over unconstrained Z. Writing the expectation as a Lebesgue integral, we aim to maximize

$$\int_\Omega \left[\Phi(Z(\omega)) - \lambda e^{-rT} L_T(\omega) Z(\omega)\right] \mathrm{d}\mathbb{P}(\omega).$$

This is a trivial problem!

Indeed, we can simply maximize *under the integral*. With $z := Z(\omega)$ for each ω separately, we treat

$$\max_z \left[\Phi(z) - \lambda e^{-rT} L_T(\omega) z\right] =: \max_z \ell(z).$$

The optimal wealth profile II

Our problem was $\max_z \ell(z) = \max_z [\Phi(z) - \lambda e^{-rT} L_T(\omega) z]$.

The first order condition obtained from $\ell'(z) = 0$ gives

$$\Phi'(z) = \lambda e^{-rT} L_T(\omega).$$

The optimal random variable $\hat{Z}$ is thus given by

$$\hat{Z} = \hat{Z}_\lambda := G\big(\lambda e^{-rT} L_T\big),$$

where G is the inverse function of Φ', $G(y) = [\Phi']^{-1}(y)$.

The dual variable λ is then determined by the budget contraint

$$e^{-rT} \mathbb{E}^{\mathbb{P}} \left[L_T \hat{Z}_\lambda \right] = x_0,$$

from which one can hopefully find λ.

$\longrightarrow$ **End of Lecture 13a** $\longleftarrow$

Example – log utility, explicit results

Assume that $\Phi(x) = \log x$. Then $\Phi'(x) = \frac{1}{x}$ and hence the inverse function of Φ' is $G(y) = \frac{1}{y}$ (put $y = \frac{1}{x}$ and solve for x). Thus

$$\hat{Z} = G\left(\lambda e^{-rT} L_T\right) = \frac{1}{\lambda L_T} e^{rT}.$$

Finally λ is determined by the budget contraint, see slide 451,

$$e^{-rT}\mathbb{E}^{\mathbb{P}}\left[L_T \hat{Z}\right] = x_0, \text{ i.e., } e^{-rT}\mathbb{E}^{\mathbb{P}}\left[L_T \frac{1}{\lambda L_T} e^{rT}\right] = x_0.$$

So $\lambda = \frac{1}{x_0}$ and the optimal wealth at T given the budget x_0 is

$$\hat{Z} = x_0 e^{rT} \frac{1}{L_T}.$$

The optimal wealth process

We have computed the optimal **terminal** wealth profile

$$\hat{Z} = \hat{X}_T = x_0 e^{rT} \frac{1}{L_T}.$$

Question. What does the optimal wealth **process** $\hat{X}_t$ look like?

As discounted prices of traded assets are $\mathbb{Q}$-martingales, one has

$$e^{-rt}\hat{X}_t = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\hat{X}_T \,\middle|\, \mathcal{F}_t\right].$$

Hence, using the above formula for $\hat{X}_T$, we obtain

$$\hat{X}_t = x_0 e^{rt}\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{L_T} \,\middle|\, \mathcal{F}_t\right].$$

Now we observe that the inverses $\frac{1}{L_t}$ are R-N derivatives in the opposite direction, $\frac{1}{L_t} = \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}$ on $\mathcal{F}_t$. Hence they form a martingale under $\mathbb{Q}$! This gives us

$$\hat{X}_t = x_0 e^{rt} \frac{1}{L_t}.$$

The Optimal Portfolio, a recap

We encounter two problems.

- We have computed the optimal wealth process, the $\hat{X}_t$, for which we have to know $\frac{1}{L_t}$.
- How do we compute the optimal portfolio?

Recall that $\hat{X}_t = x_0 e^{rt} \frac{1}{L_t}$, more conveniently written as

$$\hat{X}_t = x_0 e^{rt} L_t^{-1}.$$

With the Itô product formula we find

$$\mathrm{d}\hat{X}_t = r\hat{X}_t \,\mathrm{d}t + x_0 e^{rt} \,\mathrm{d}L_t^{-1}.$$

Recall that L_t^{-1} is the R-N derivative $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}$ on $\mathcal{F}_t$ and that $\mathrm{d}W_t^{\mathbb{Q}} = \phi_t \,\mathrm{d}t + \mathrm{d}W_t$, when $\mathrm{d}L_t = \phi_t L_t \,\mathrm{d}W_t$.

About L^{-1}

The process L^{-1} is a $\mathbb{Q}$-martingale. This implies that we should have for some ψ_t the equation

$$\begin{aligned} \mathrm{d}L_t^{-1} &= \psi_t L_t^{-1}\,\mathrm{d}W_t^{\mathbb{Q}} \\ &= (\ldots)\,\mathrm{d}t + \psi_t L_t^{-1}\,\mathrm{d}W_t, \end{aligned}$$

where we used that $\mathrm{d}W_t^{\mathbb{Q}} = \phi_t\,\mathrm{d}t + \mathrm{d}W_t$ and where it is not relevant for us to know what is on the dots.

Using the expression for $\mathrm{d}L_t^{-1}$ we get

$$\mathrm{d}\hat{X}_t = (\ldots)\,\mathrm{d}t + \psi_t \hat{X}_t\,\mathrm{d}W_t.$$

General theory for the measure change from $\mathbb{Q}$ to $\mathbb{P}$ gives $\mathrm{d}W_t = \psi_t\,\mathrm{d}t + \mathrm{d}W_t^{\mathbb{Q}}$. Together with $\mathrm{d}W_t^{\mathbb{Q}} = \phi_t\,\mathrm{d}t + \mathrm{d}W_t$ this implies that $\psi_t = -\phi_t$.

The conclusion is that

$$\mathrm{d}\hat{X}_t = (\ldots)\,\mathrm{d}t - \phi_t \hat{X}_t\,\mathrm{d}W_t.$$

where it is still not relevant for us to know what is on the dots.

Basic program, some further steps

1. Recall

$$\mathrm{d}\hat{X}_t = (\dots)\,\mathrm{d}t - \phi_t \hat{X}_t\,\mathrm{d}W_t.$$

2. Assume for simplicity a standard Black-Scholes model,

$$\mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t,$$
$$\mathrm{d}B_t = rB_t\,\mathrm{d}t.$$

Let $\hat{u}_t$ be the relative portfolio investment in the stock. One then has the self financing condition

$$\mathrm{d}\hat{X}_t = \hat{X}_t\left[(1-\hat{u}_t)\frac{\mathrm{d}B_t}{B_t} + \hat{u}_t\frac{\mathrm{d}S_t}{S_t}\right],$$

which we now write as

$$\mathrm{d}\hat{X}_t = \hat{X}_t(\dots)\,\mathrm{d}t + \hat{X}_t\hat{u}_t\sigma\,\mathrm{d}W_t.$$

The optimal $\hat{u}_t$ in the Black-Scholes model

By comparing the $\mathrm{d}W_t$ terms in the two equations for $\mathrm{d}\hat{X}_t$, we can now identify $\hat{u}$ as

$$\hat{u}_t = -\frac{\phi_t}{\sigma}.$$

But, in the context of the Black-Scholes model, see slide 215, we found $\phi_t = \frac{r-\mu}{\sigma}$.

Result. The optimal portfolio is determined by

$$\hat{u}_t = \frac{\mu - r}{\sigma^2}$$

Note that $\hat{u}$ yields a "myopic" portfolio in the sense that it does not depend on the time horizon T. Note also that $\hat{u}$ is constant.

A final digression: the numéraire portfolio

The **standard approach** to establish a measure-numéraire pair is the following.

- Choose first a fixed candidate numéraire (portfolio) N.
- Find the corresponding martingale measure, i.e., find $\mathbb{Q}^N$ s.t. $\frac{B}{N}$ and $\frac{S}{N}$ are $\mathbb{Q}^N$-martingales.

The **alternative approach** is as follows.

- Choose first a fixed measure $\mathbb{Q} \sim \mathbb{P}$.
- Find numéraire N such that $\mathbb{Q} = \mathbb{Q}^N$.

Special case.

- Make the special choice $\mathbb{Q} = \mathbb{P}$.
- Find numéraire N such that $\mathbb{Q}^N = \mathbb{P}$ i.e., such that $\frac{B}{N}$ and $\frac{S}{N}$ are martingales under the **objective** measure $\mathbb{P}$.
- This numéraire N is called the **numéraire portfolio**.

Log utility and the numéraire portfolio

Theorem. The portfolio $\hat{X}$, also named *growth optimal portfolio*, is the numéraire portfolio.

Proof. We have to show that the process

$$Y_t = \frac{S_t}{\hat{X}_t}$$

is a $\mathbb{P}$-martingale, as well as $\frac{B_t}{\hat{X}_t}$. We know from slide 453 that $\frac{S_t}{\hat{X}_t} = x_0^{-1} e^{-rt} S_t L_t$. This is indeed a $\mathbb{P}$-martingale, as desired, since $x_0^{-1} e^{-rt} S_t$ is a (very familiar) $\mathbb{Q}$-martingale, as follows from Additional Exercise 3, or Exercise C.9 in the book.

The story for $\frac{B_t}{\hat{X}_t}$ is almost a triviality. Check!

⟶ End of Lecture 13b ⟵

This is the end of the course.

Thank you for your attention and good luck with the exam!