CQF Module 6, Session 7: PDEs and Martingales

Which, when and why?

CQF

In this lecture...

...

- quantitative finance, computational finance, financial mathematics or mathematical finance?
- who does what in QuantLand?
- the link between PDEs and probabilities;
- problems, methods and models;
- managing model risk;

1. Quantitative Finance, Computational Finance, Financial Mathematics or Mathematical Finance?

The Quantitative Finance Triangle

There are three main perspectives from which one may join the field of quantitative finance:

- mathematical sciences;
- financial economics;
- computing;

Figure: The quantitative finance triangle

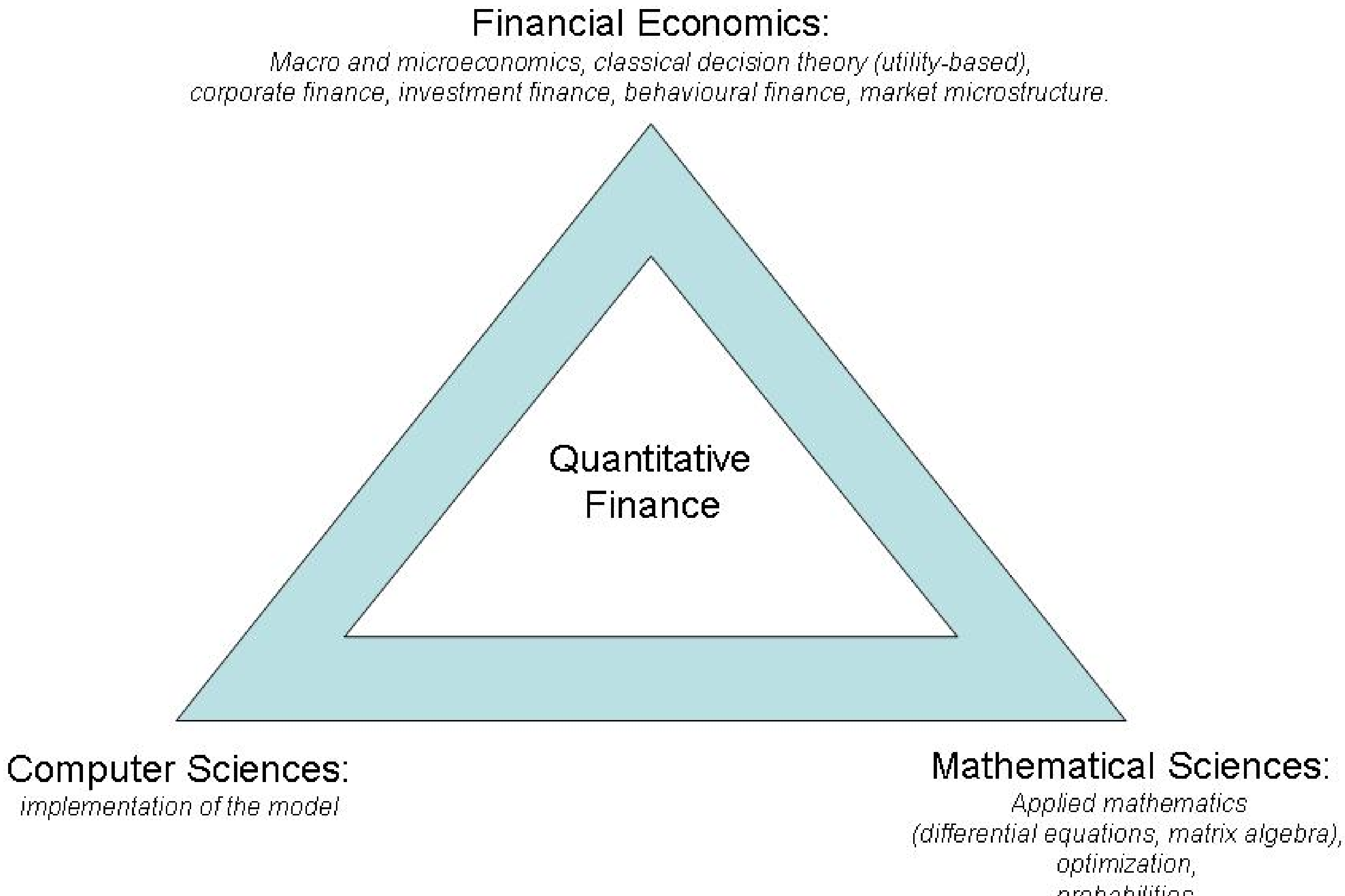


Figure: Outside forces

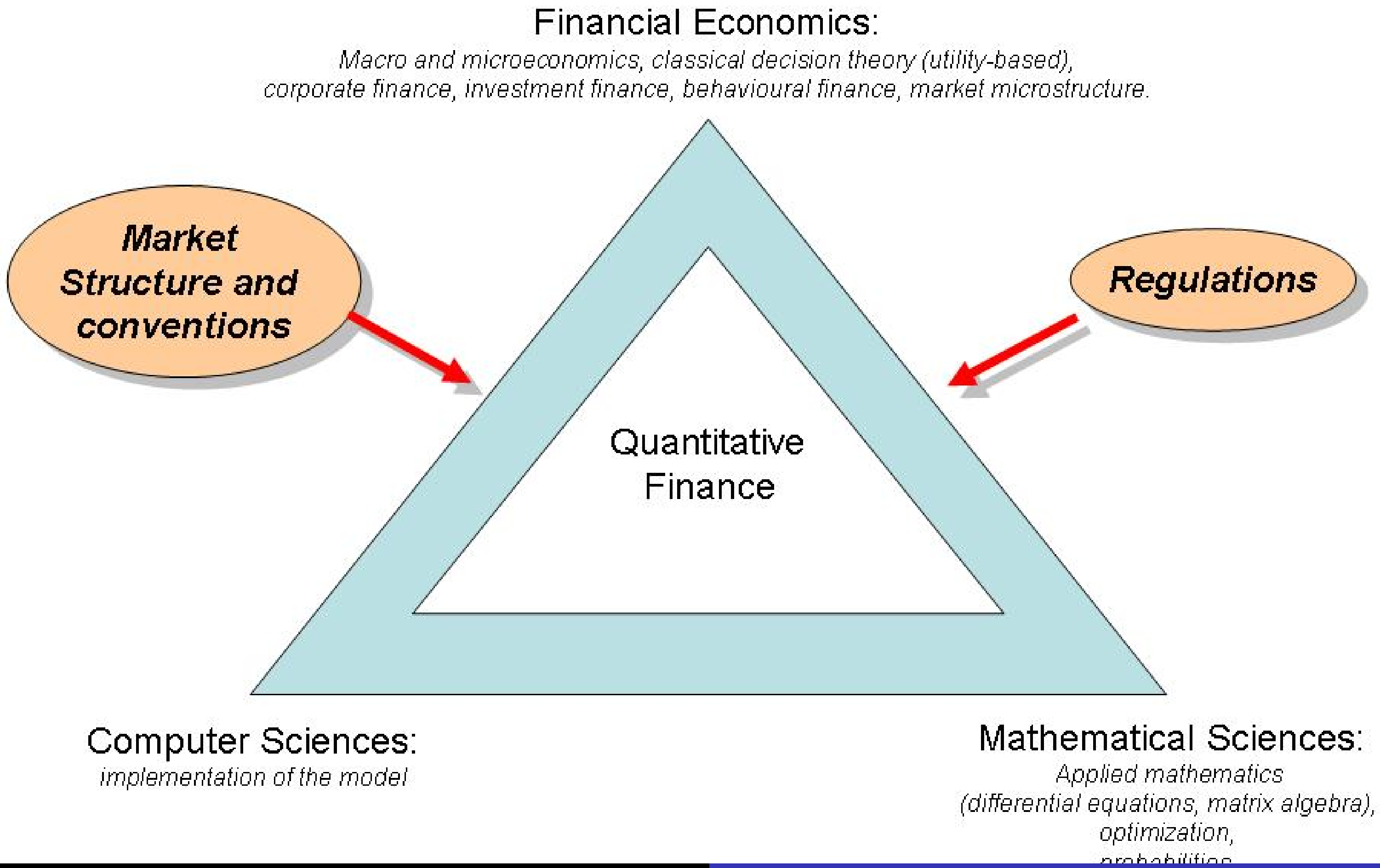
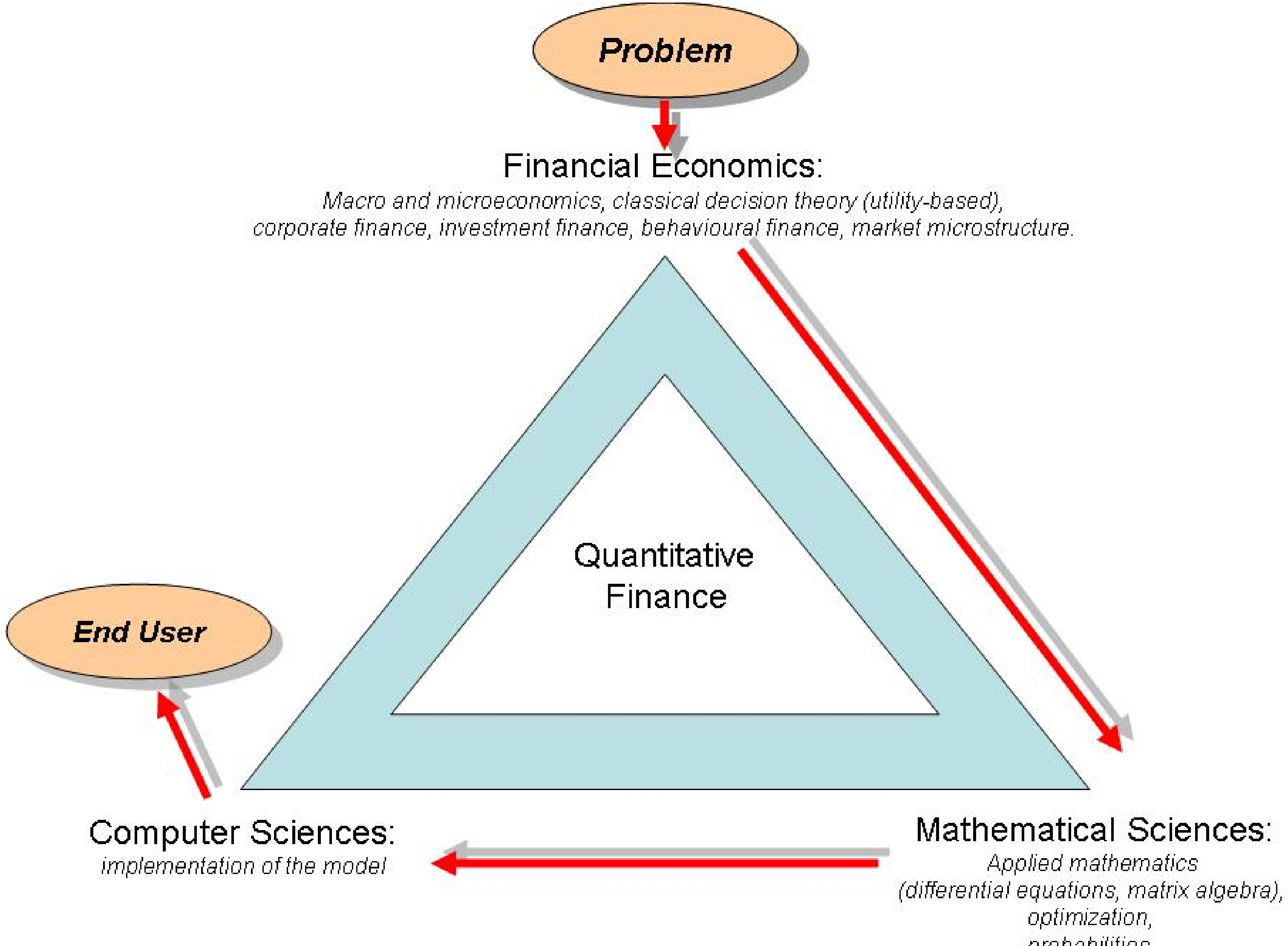


Figure: From problem to solution



2. Who Does What in Quant Land?

Some traditional backgrounds:

- the physicist and the mathematician;
- the computer scientist and the engineer;
- the financial economist and the economist;
- the econometrist and the statistician;

2.1. The Physicist and The Mathematician

Specialities:

- (mathematical) physics;
- probabilities and pure mathematics;
- applied mathematics (PDE);
- operations research (optimal control, optimization);

2.1.1. The (Mathematical) Physicist

Examples: Emanuel Derman (Columbia U); Jim Gatherall (NYU and Merrill Lynch), Lane Hughston (Imperial College), Alex Lipton (Merrill Lynch and Imperial College);

Depending on their orientation, mathematical physicists may be either close to applied mathematicians, using differential equations to model markets and products in the same way physical phenomenons are modelled (Jim Gatheral), or closer to pure mathematicians and probabilists (Lane Hughston).

2.1.2. The Probabilist and the Pure Mathematician

Examples: Jean Jacod (Paris VI), David Heath (Carnegie Mellon U.), Paul Malliavin (Paris VI), Philip Protter (Cornell U.), Walter Schachermayer (TU Wien), Albert Shiriaev (Russian Academy of Science), Steven Shreve (Carnegie Mellon U.), Marc Yor (Paris VI).

They are generally more interested the intrinsic properties of the mathematical world. They consider problems for the elegance of the mathematics required to solving them.

2.1.3. The Applied Mathematician (PDE)

Examples: Riaz Ahmad, Marco Avallaneda (NYU), Paul Wilmott;

They are interested in the dynamics of the problem, and use DEs (mostly PDEs and ODEs) to model the physical properties of the problem.

2.1.4. The Operation Research Scientist

Examples: Mark Davis (Imperial College), Michael Dempster (Cambridge U.), Paul Glasserman (Columbia University), Stanley Pliska (U. Illinois), Mete Soner (Koch U.), Agnes Sulem (INRIA), Nizar Touzi (Ecole Polytechnique), William Ziemba (UBC);

Operations research scientists are the intellectual heirs to the interdisciplinary approach to science pioneered in the UK and the USA during World War II. Broadly speaking, their emphasis is on the mathematics of decision making (optimization) and of complex "real life" systems (aeronautics and space program).

Most of them have worked or currently work on portfolio selection and investment management.

2.1.5. Concluding on the mathematicians

The difference between “pure” and “applied” mathematicians is more tenuous than most people would like to admit.

To solve a concrete problem, an applied mathematician may well derive a new mathematical result or even initiate the study of a new branch of mathematics.

In the derivation of new results, a pure mathematician often spearheads the discovery of new, more general methods that can then be applied to solve ever more complicated real-life problems.

In fact, some of the greatest mathematicians have/had a mind for both abstraction and application (the names of Archimedes, Euler, Gauss, Fourier, Poincaré, von Neumann and Mandelbrot spring to mind).

2.2. The Computer Scientist and the Engineer

Examples: Bruno Dupire (Bloomberg and NYU);

Computer scientists and engineers tend to be doers who make things work. They do not publish results[1] although they are often happy to show their latest tricks.

[1] hence the lack of examples!

2.3. The Financial Economist and The Economist

Examples: Peter Carr (Bloomberg and NYU), Darrel Duffie (Stanford), John Hull (U. of Toronto), Jonathan Ingersoll (Yale U.), Robert Jarrow (Conell U.), Andrew Lo (MIT), Harry Markowitz (UC San Diego), Robert Merton (Harvard U.), Stephen Ross (MIT), Paul Samuelson (MIT), William Sharpe (Stanford U.), Alan White (U. of Toronto);

Economists have led the development of both financial economists and quantitative finance for the past 60 years. Many of the concepts quants use daily stem directly from the foundations set in the 1950s by Arrow, Debreu, Markowitz, Modigliani and Miller.

Ignoring this simple reality means ignoring the fact that finance is as much (if not more) a social science as a physical science.

2.4. The Econometrist and the Statistician

Examples: Christian Gourieroux (U. of Toronto), Jean-Paul Laurent (U. of Lyon I and BNP Paribas), Kenneth Singleton (Stanford U.);

Econometrics and statistics play two important roles in the story of quantitative financial economics.

- **econometrics** is traditionally used to test empirically the validity of economic theories (such as the CAPM or the APT);
- **statistics** is not only the mathematical foundation of econometrics, but it is also widely used to analyse the vast quantities of data generated by both the economy and the financial markets.

Who Else?

Behavioural finance and **neuroeconomics** have been on the rise, especially since Daniel Kahneman was awarded the 2002 Nobel Memorial Prize in Economics.

The main critic of behavioural finance and neuroeconomics toward classical economics is that standard utility theory does not capture the decision-making process of individuals. First, individuals do not make optimal decisions: they suffer from a number of biases. Second, individuals are more loss averse than assumed by economic theory.

3. The Link between Probability and PDEs

3.1. The Kolmogorov Equation and the Generator of a Stochastic Process

3.1.1. The Kolmogorov Equation

Take a process $X(t)$ with dynamics

$$X(t) \quad = \quad a(t, X(t))dt + b(t, X(t))dW_t$$

where $W(t)$ is a standard Brownian motion.

The drift and diffusion coefficients are the limits

$$\begin{aligned} a(s,x) &= \lim_{t \downarrow s} \frac{1}{t-s} \mathbf{E}\left[X(t) - X(s) | X(s) = x\right] \\ b^2(s,x) &= \lim_{t \downarrow s} \frac{1}{t-s} \mathbf{E}\left[(X(t) - X(s))^2 | X(s) = x\right] \end{aligned}$$

The **transition probability** $p(s, x; t, y)$ for the process $X(t)$ is defined as

$$P\left[a < x < b \text{ at time } t | x \text{ at time } s\right] = \int_a^b p(s, x; t, y) dy$$

The transition probability $p(s,x;t,y)$ satisfies

- the **Kolmogorov forward equation** (or **Fokker Planck equation**):

$$\frac{\partial p}{\partial t} + \frac{\partial}{\partial y}\left(a(t,y)p\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(b^2(t,y)p\right) = 0$$

  for (s,x) fixed;

- the **Kolmogorov backward equation**:

$$\frac{\partial p}{\partial s} + \frac{\partial}{\partial x}\left(a(s,x)p\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(b^2(s,x)p\right) = 0$$

  for (t,y) fixed.

When $X(t)$ is a standard Brownian motion (i.e. $a(x,t) \equiv 0$ and $b(x,t) \equiv 0$), the transition probability $p(s,x;t,y)$ is a Normal PDF:

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

Itô, generator and Kolmogorov equation

Consider a process $Y(t) = f(t, X(t))$ for some "nice" function $f(s, x)$. By Itô,

$$\begin{aligned} dY(t) &= \left(\frac{\partial f}{\partial s}(t, X(t)) + \frac{\partial f}{\partial x}(t, X(t)) a(t, X(t)) \right. \\ &\quad \left. + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \left(b^2(t, X(t)) p \right) \right) dt \\ &\quad + \frac{\partial f}{\partial x}(t, X(t)) b(t, X(t)) dW(t) \end{aligned}$$

The drift coefficient of $Y(t)$ defined as the limit

$$\mathcal{A}f(s,x) := \lim_{t\downarrow s} \frac{1}{t-s}\mathbf{E}\left[f(X(t)) - f(X(s))|X(s)=x\right]$$

is called the **generator** of $Y(t)$.

The diffusion coefficient of $Y(t)$ is defined as the limit

$$\mathcal{C}f(s,x) := \lim_{t\downarrow s} \frac{1}{t-s}\mathbf{E}\left[(f(X(t)) - f(X(s)))^2|X(s)=x\right]$$

Hence, we could write the dynamics of $Y(t)$ as

$$dY(t) \;=\; \mathcal{A}f(s,x)dt + \mathcal{C}f(s,x)dW(t)$$

The **transition probability** $p_f(s,x;t,y)$ for the process $Y(t)$ satisfies

- the **Kolmogorov forward equation** (or **Fokker Planck equation**):

$$\frac{\partial q}{\partial t} + \frac{\partial}{\partial y}\left(\mathcal{A}f(t,y)p_f\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\mathcal{C}^2 f(t,y)p_f\right) = 0$$

  for (s,x) fixed;

- the **Kolmogorov backward equation**:

$$\frac{\partial q}{\partial s} + \frac{\partial}{\partial x}\left(\mathcal{A}f(s,x)p_f\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\mathcal{C}^2 f(s,x)p_f\right) = 0$$

  for (t,y) fixed.

3.2. The Feynman-Kač Formula

The Feynman-Kač formula is very useful in finance to go from a parabolic PDE to an expectation, and vice versa.

Its main application is to go from the fundamental asset pricing formula to the Black-Scholes PDE.

A similar reasoning can also be seen in stochastic control in the derivation of the Hamilton-Jacobi-Bellman PDE.

Key Fact (The Feynman-Kač Formula)

Assume that $V(t, S)$ solves the boundary value problem

$$\begin{aligned}\frac{\partial V}{\partial t}(t,s) + \mu(t,s)\frac{\partial V}{\partial s}(t,s) + \frac{1}{2}\sigma^2(t,s)\frac{\partial^2 V}{\partial s^2}(t,s) - rV(t,s) &= 0 \\ V(T,s) = G(s) \end{aligned} \quad (1)$$

and that the process S follows the dynamics

$$dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dX(t)$$

where $X(t)$ is a Brownian motion. Then, the function V has the representation

$$V(t, S_t) = e^{-r(T-t)}\mathbf{E}\left[G(S_T)|\mathcal{F}_t\right] \quad (2)$$

4. Which? A Guide to Problems and Methods

1. European style derivatives;
2. Derivatives with embedded decisions (American & Bermudean);
3. "Exotic" derivatives;
4. Transaction costs;
5. Risk management;
6. Jumps-diffusion problems;
7. Dynamic portfolio selection;
8. Incomplete markets and super-replications;

4.1. European Style Derivatives

The two most widespread methodologies to price European style derivatives are the no-arbitrage (PDE) method and the (probabilistic) equivalent martingale methods.

With the no-arbitrage method, it is very easy to derive a Black-Scholes style PDE. From there, we obtain directly the sensitivities (Delta, Gamma...). Further work yield a Black-Scholes style formula. Alternatively, the PDE can be evaluated numerically via FDM.

With the equivalent martingale method, we obtain directly a Black-Scholes style formula. The advantage of this method is that it can be generalized into a "fundamental pricing formula." This pricing formula will work (with minor modifications) on most problems and forms the basis for the use of Monte-Carlo methods

4.2. Derivatives with embedded decisions (American & Bermudean)

In an American call option, the holder has the right to exercise at any time $0 < t \leq T$ to get the option payoff $\max(S_t - K, 0)$. The holder will try to maximize the value of the option by selecting an optimal exercise strategy.

Because of this feature, pricing a American option requires jointly solving a PDE and an embedded optimization problem. This results in a so-called **free boundary problem**.

Such problem can be solved numerically by applying the **dynamic programming principle** within a FDM scheme.

Consider a diffusion PDE:

$$\frac{\partial V}{\partial t} + a(S,t)\frac{\partial^2 V}{\partial S^2} + b(S,t)\frac{\partial V}{\partial S} + c(S,t)V = 0$$

and a finite difference grid with steps

$$\delta S = \frac{S_{max}}{I}$$
$$\delta t = \frac{T-t}{K}$$

and points

$$S = i\delta S$$
$$t = T - k\delta t$$

with $0 \leq i \leq I$ and $0 \leq k \leq K$.

The explicit FDM approximation for the PDE is

$$V_i^{k+1} = A_i^k V_{i-1}^k + \left(1 + B_i^k\right) V_i^k + C_i^k V_{i+1}^k$$

where

$$\begin{aligned} A_i^k &= \frac{\delta t}{\delta S^2} a_i^k - \frac{1}{2}\frac{\delta t}{\delta S} b_i^k \\ B_i^k &= -2\frac{\delta t}{\delta S^2} a_i^k - \delta t c_i^k \\ C_i^k &= \frac{\delta t}{\delta S^2} a_i^k + \frac{1}{2}\frac{\delta t}{\delta S} b_i^k \end{aligned}$$

At each node the option holder must decide whether he/she exercises the option. The holder will only exercise the option if the payoff at that node is greater than V_i^{k+1}, the value of keeping the option alive over the next time step.

The value of the option at node (k, i) is the maximum of

- the payoff $\max\left(S_i^k - K, 0\right)$;
- the continuation value V_i^{k+1}.

Probabilistic Approach

Probabilistic methods yield few insights and no analytical solutions.

Furthermore, Monte-Carlo methods are inefficient when it comes to modelling embedded decisions. The reason is that the decision-making process (as encapsulated in the dynamic programming principle) runs backward in time while simulations run forward in time.

4.3. “Exotic” Derivatives

The category of “exotic” derivatives commonly includes

- forward-start;
- chooser;
- compound;
- binary;
- barrier;
- lookback;
- Asian;
- basket;
- passport.

In the Black-Scholes world, most exotic options can be priced using either a PDE-based or a probabilistic approach.

As is the case for European derivatives,

- The PDE approach makes for fast implementation using FDM;
- The probabilistic approach produces analytical pricing formulae and elicits the properties of the exotic derivative.

4.4. Transaction costs

One of the key assumptions made by Black and Scholes in their original work[2] is the absence of transaction costs.

This assumption is unrealistic from a practical point of view. Even for large financial institutions, buying and selling securities costs money.

As a result of these transaction costs, traders have an incentive not to rebalance their hedge continuously, contrary to what standard models (such as Black-Scholes) would imply.

[2] and a standard assumption in the field of financial economics

4.4.1.How do we model transaction costs?

The simplest way to model transaction costs is as a **proportion** of the market value of the trade. This corresponds to the case of a 0.5% fee charged on the value of the trade.

For example, denote by κ the transaction fee and by ν_t the number of shares traded[3] at time t. The transaction cost (in £) is equal to

$$\kappa|\nu_t|S_t$$

[3] $\nu_t > 0$ indicates a purchase of $|\nu_t|$ shares while $\nu_t < 0$ represents a sale of $|\nu_t|$

4.4.2. Transaction cost and derivative pricing

The most effective techniques to address the question of transaction cost in derivative pricing are all PDE-based.

How do we value an option in a Black-Scholes world but with transaction cost?

Because trading the underlying asset is costly, we will not hedge continuously. Rather, we will hedge in discrete time. The appropriate model for the behaviour of the underlying is therefore the discretized GBM:

$$\delta S = \mu S \delta t + \sigma S (\delta t)^{1/2} Z$$

where $Z \sim N(0, 1)$.

Next, we construct a hedged portfolio Π by selling Δ units of the underlying asset:

$$\Pi = V - \Delta S$$

The change in value of the portfolio $\delta\Pi$ over a short time step δt is equal to

$$\begin{aligned}\delta\Pi &= \sigma S\left(\frac{\partial V}{\partial S} - \Delta\right)(\delta t)^{1/2} Z \\ &+ \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} Z^2 + \mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} - \mu \Delta S\right)\delta t - \kappa|\nu|S\end{aligned}$$

The number of units of underlying assets held in the hedging portfolio at time t is still[4]

$$\Delta_t = \frac{\partial V}{\partial S}$$

[4] You can see this by either 1). choosing the Delta which minimizes the variance of the change in portfolio value, or 2). conducting the analysis in continuous time, only switching to discrete time in the later stages.

The number of units of the underlying asset traded between time t and time $t + \delta t$ is equal to the difference between the number of units held at time $t + \delta t$ and the number of units held at time t:

$$\begin{aligned} \nu_t &= \Delta_{t+\delta t} - \Delta_t \\ &= \frac{\partial V}{\partial S}(t + \delta t, S + \delta S) - \frac{\partial V}{\partial S}(t, S) \end{aligned}$$

Apply Taylor to expand the $\frac{\partial V}{\partial S}(t+\delta t, S+\delta S)$ term into

$$\frac{\partial V}{\partial S}(t+\delta t, S+\delta S) = \frac{\partial V}{\partial S}(t,S) + \frac{\partial^2 V}{\partial^2 S}(t,S)\delta S + \frac{\partial^2 V}{\partial S \partial t}(t,S)\delta t + \dots$$

Substituting $\delta S = \mu S \delta t + \sigma S (\delta t)^{1/2} Z$, this expansion becomes

$$\begin{aligned} \frac{\partial V}{\partial S}(t+\delta t, S+\delta S) &= \frac{\partial V}{\partial S}(t,S) + \frac{\partial^2 V}{\partial^2 S}(t,S)\sigma S(\delta t)^{1/2} Z \\ &+ \left(\frac{\partial^2 V}{\partial^2 S}(t,S)\mu S + \frac{\partial^2 V}{\partial S \partial t}(t,S) \right) \delta t + \dots \end{aligned}$$

The number of units of the underlying asset traded over the interval $[t, t + \delta t]$ is, to leading order $(o((\delta t)^{1/2}))$,

$$\nu \approx \frac{\partial^2 V}{\partial^2 S}(t, S)\sigma S(\delta t)^{1/2} Z$$

This quantity is a random variable (due to the presence of Z). As a result, the transaction cost we will have to pay over the period $[t, t + \delta t]$ is also a random variable. The expected value of this cost is equal to

$$\mathbf{E}\left[\kappa|\nu|S\right] \approx \kappa\sigma S^2 \sqrt{\frac{2}{\pi}} \left|\frac{\partial^2 V}{\partial S^2}\right| (\delta t)^{1/2}$$

because $\mathbf{E}\left[|Z|\right] = \sqrt{\frac{2}{\pi}}$.

The expected change in the portfolio value, $\mathbf{E}[\delta\Pi]$, is equal to

$$\mathbf{E}[\delta\Pi] = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \kappa\sigma S^2 \sqrt{\frac{2}{\pi\delta t}} \left| \frac{\partial^2 V}{\partial S^2} \right| \right) \delta t$$

But the portfolio is fully hedged, and it is therefore risk-free. To prevent arbitrage, the expected change in the portfolio value should be equal to the change in an equivalent investment at the risk-free rate, i.e.

$$\mathbf{E}[\delta\Pi] = r\left(V - S\frac{\partial V}{\partial S} \right) \delta t$$

Rearranging and dividing by δt, we obtain the Black-Scholes equation with transactions costs:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \kappa \sigma S^2 \sqrt{\frac{2}{\pi \delta t}} \left| \frac{\partial^2 V}{\partial S^2} \right| + rS \frac{\partial V}{\partial S} - rV = 0$$

Note that the new term $\kappa\sigma S^2\sqrt{\frac{2}{\pi\delta t}}\left|\frac{\partial^2 V}{\partial S^2}\right|$:

- transforms the Black-Scholes PDE into a nonlinear parabolic PDE;
- is a function of $\Gamma = \frac{\partial^2 V}{\partial S^2}$, which measures the degree of mishedging;

- can be written as $\kappa\sigma S^2\sqrt{\frac{2}{\pi\delta t}}\frac{\partial^2 V}{\partial S^2}$ **in the case of a vanilla call**, implying that the PDE can be interpreted as the classical Black-Scholes PDE but with adjusted variance

$$\sigma_c^2 = \sigma^2 - 2\kappa\sigma\sqrt{\frac{2}{\pi\delta t}}$$

- can be written as $-\kappa\sigma S^2\sqrt{\frac{2}{\pi\delta t}}\frac{\partial^2 V}{\partial S^2}$ **in the case of a vanilla put**, implying that the PDE can be interpreted as the classical Black-Scholes PDE but with adjusted variance

$$\sigma_p^2 = \sigma^2 + 2\kappa\sigma\sqrt{\frac{2}{\pi\delta t}}$$

4.4.3. What about probabilistic approaches?

Recently, probabilitic methods have been used to address a problem related to transaction costs.

From a market perspective, transaction costs are but a specific manifestation of a larger phenomenon: **market liquidity**.[5]

Broadly speaking, market liquidity is an indication of how easy it is to trade securities on that market. The more liquid a market is, the lower the bid-ask spread or transaction fee charged to execute a trade.

[5] The type of transaction costs we have seen so far only truly applies to equity markets. In fixed income and FX markets, the "cost" of trading is measured by the bid-ask spread.

Modelling market liquidity is considerably more complex than modelling simple transaction costs.

Indeed, while we have modelled transaction costs as exhogeneous to the market and the underlying asset, we need to model market liquidity endogenously.

This is generally done by modelling both a **supply curve** and a **demand curve** for the security, in a manner consistent with market microstructure models.

This approach, which combines financial economics, market microstructure and stochastic analysis is still in its infancy. The key article is Cetin, Jarrow and Protter [1].

4.5. Risk Management

PDEs and probabilities give two different and complementary levels of information useful for risk management.

4.5.1 PDEs in risk management

In the PDE approach, risk measurement is related to the computation of the Greeks: the sensitivities to a small change in the variables (time, stock price) and parameters (volatility, interest rates) of the pricing problem.

This leads to the reinterpretation of the Black-Scholes PDE as the “risk management” equation

$$\Theta + \frac{1}{2}\sigma^2 S^2 \Gamma + rS\Delta - rV = 0$$

with similar interpretation for other pricing PDEs.

In the PDE approach, risk management is **local** and concentrated on a **single underlying instrument**.

- **Local**: a PDE and its partial derivatives are only relevant for small changes in time and state of order dt and dS. As a result, Greeks are not useful to estimate potential losses at a 1-day, 1-week or 1-month horizon.
- **Single underlying instrument**: because PDEs and Greeks only apply to the instrument they price and to its underlying security, they are not useful to get a consolidated view of risk across an entire trading book.

4.5.2. Probabilities in risk management

With their focus on the probability distribution of the underlying and its derivative, probabilistic methods are better suited for risk measurement and management at the portfolio level and over a longer period of time (days, weeks, months...).

In this approach, Monte Carlo simulations are used to compute the probability distribution of losses on the portfolio / trading book. This distribution is then used to compute a measure of risk such as standard deviation, VaR or CVaR.

Probabilistic methods further point to specific risk measurement difficulties linked to the nature of no-arbitrage pricing as both the distribution obtained by simulation and the pricing PDE are derived in the pricing measure, but not in the real world measure. Yet, risk management takes place in the real world, not in the pricing world.

The implication of this discrepancy is that the mean of the (pricing world) probability distribution derived through simulation will be lower than the mean of the actual (real world) probability distribution. Over a short-term horizon of a few days the difference between the two distributions will be small. But this difference will increase with the time horizon, and could be significant over a long period of time (a year or more).

4.6. Jumps Diffusion Problems

Jump-diffusion processes have become increasingly popular in recent years. They have been applied to

- introduce non-normality in the distribution of asset returns;
- model asset price discontinuities;
- model corporate bankruptcy and default (in intensity-based credit risk models).

In the CQF, we have encountered jump-diffusion processes of the form

$$dS(t) = \mu(t, S(t))dt + \sigma(t, S(t))dX(t) + \gamma(t, S(t^-))dJ(t)$$

where

$$dJ(t) = \begin{cases} 0 & \text{with probability } (1 - \lambda dt) \\ 1 & \text{with probability } (\lambda dt) \end{cases}$$

is the differential of a **Poisson process** with intensity λ.

In this model, at each time t, a jump could occur. The size of this jump is equal to:

$$\Delta S(t) = S(t) - S(t^{-})$$

4.6.1. Lévy processes

The type of jump-diffusion process we have encountered so far is just one simple example of a larger class of jump diffusion processes called **Lévy (-driven) processes** of the form

$$\begin{aligned} dS(t) &= \mu(t, S(t))dt + \sigma(t, S(t))dX(t) \\ &+ \int_{|z|\leq R} \gamma(t, S(t^-), z)\tilde{N}(dt, dz) \\ &+ \int_{|z|>R} \gamma(t, S(t^-), z)N(dt, dz) \end{aligned}$$

The term $N(dt, dz)$ is the differential of the **Poisson measure** $N(t, U)$.

The Poisson measure represents the number of jump of size $\Delta S(s) \in U$ which have occurred between time 0 and time t. Mathematically,

$$N(t, U) = \sum_{s:0<s\leq t} \mathbf{1}_{\{\Delta S(s)\in U\}}$$

The Poisson measure produces a non-decreasing pure jump process. As a result, it has a positive drift and is not a martingale.

The "expected drift" of this process is called the compensator. It is a measure called the **Lévy measure** $\nu(U)$ and is computed as

$$\nu(U) = \mathbf{E}\left[N(1, U)\right]$$

Hence, the **compensated Poisson measure**

$$\tilde{N}(t, U) = N(t, U) - \nu(U)$$

is a martingale.

The two integrals

- $\int_{|z|\leq R}\gamma(t, S(t^-), z)\tilde{N}(dt, dz)$,
- $\int_{|z|>R}\gamma(t, S(t^-), z)N(dt, dz)$.

respectively represent the

- small jumps (arbitrarily cut at a level R) which occur very frequently. Because small jumps are so frequent, their expected drift tends to become large. To prevent them from having a potentially explosive behaviour, it is customary to "compensate" them. Once compensated, the behaviour of the small jumps approximate the behaviour of a Brownian motion.
- large jumps (arbitrarily cut at a level R) which occur infrequently.

As the type of jump-diffusion process becomes more complex, we will no longer be dealing with the familiar PDEs but with much more complex Partial Integro-Differential Equation (PIDE).

The end result is that dealing with jump-diffusion processes and Lévy processes in general will require some type of probabilistic reasoning which can be developed around two complementary approaches:

1. using the characteristic function of the Lévy process;
2. using stochastic analysis;

4.6.2. The characteristic function of Lévy processes

Before Itô calculus and the thorough study of martingales, the analysis of stochastic processes revolved around either their Kolmogorov equation or the characteristic function of their distribution.

The characteristic function is the Fourier transform of the distribution function. It is therefore both useful and easy to manipulate.

Lévy processes have a particularly elegant characteristic function given by the celebrated **Lévy Kintchine formula**.

A vast literature on option pricing with Lévy processes exploits the properties of the characteristic function (see Schoutens [6] for an overview).

4.6.3. Stochastic analysis

The most recent approach to jump-diffusion processes is purely stochastic.

It makes use of jump-diffusion versions of the classical results such as Itô, Grisanov, Feynman Kǎc, the fundamental asset pricing formula etc... and of more abstract results related to the theory of semi-martingales.

4.7. Dynamic Portfolio Selection

Most of the questions we have explored during the CQF are related to the pricing of derivative securities. However, derivatives pricing is just one aspect of quantitative finance.

Another important class of problem is the **portfolio selection problem**: if you had £1 million today, how would you invest it to retire comfortably?

We have seen (in Lecture 2.1 and Lecture 2.2) the original static approach developed by Markowitz and Sharpe to solve the portfolio selection problem.

One of the main drawbacks of the Markowitz approach is precisely that it is static. Several dynamic approaches have been developed to address this weakness:

1. The most established approach calls for the application of an important technique called stochastic control, which stands at the crossroad of optimization, stochastic analysis and PDE theory.
2. A stochastic programming approach is often used to solve concrete long-term investment problems insurance companies and pension funds face.
3. More recently, a continuous time martingale approach has also been developed.

4.7.1. The stochastic control approach

In the late 1960s, Merton [4] (see also [5]) proposed a formulation of the investment problem as a stochastic control problem where the objective is to maximize the investor's utility of wealth.

The Merton model constitutes one of the few nonlinear stochastic control problems which can be solved analytically.

For clarity, we will only sketch the case in which an investor has the choice between investing in a risky asset and a risk-free asset.

Traditionally, the risky asset S_t is modelled using a geometric Brownian motion,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW(t)$$

where μ and σ are two constants and $W(t)$ is a Brownian motion.

The risk-free asset $B(t)$ follows the dynamics given by

$$\frac{dB_t}{B_t} = rdt$$

where r is the risk-free rate.

The investor has an horizon of T years.

The main objective of the investor is therefore to determine at time $t < T$, the proportion w_t of his/her wealth invested in the risky security (and by extension the proportion $1 - w_t$ of his/her wealth invested in the risk-free security) so as to maximize some measure of his/her terminal wealth $X(T)$.

The time t dynamics of the investor's wealth X_t with a proportion w_t invested in the risky asset and a proportion $(1 - w_t)$ invested in the risk-free asset is given by

$$dX_t = w_t(\mu - r)X_t dt + rX_t dt + w_t \sigma X_t dW_t, \qquad X(0) = x$$

Let's now add one extra level of complexity. Let's say that the investor is not just interested in the total wealth accumulated by time T, but that he/she would also like to draw some money out of the portfolio to finance their consumption.

This requirement can be modelled as a dividend drawn from the wealth at a yet to-be-determined level c_t. The equation for the investor's wealth becomes:

$$dX_t = w_t(\mu - r)X_t dt + (rX_t - c_t)dt + w_t \sigma X_t dW_t, \qquad X(0) = x$$

Standard economic theory tells us that wealth is generally not a good indicator of "economic satisfaction."

Instead, we need to consider a function of wealth, called a **utility function**, which tracks the satisfaction the investor derives from his/her wealth based on his/her level of risk tolerance.

In our problem, we need two utility functions:

1. a bequest function $\Phi(X_T)$ which tracks the satisfaction derived from reaching a given level of wealth $X(T)$ at the end of the investment horizon T;
2. an (instantaneous) utility function $U(t, c_t)$ which indicates the satisfaction derived from being able to draw an amount c_t out of the portfolio at time t.

Broadly speaking, the investor's objective is to find an optimal asset allocation w_t and an optimal consumption c_t to maximize both

- the total utility from consumption, equal to $\int_0^T U(t, c_t)dt$, and;
- the utility from terminal wealth $\Phi(X_T)$;

But this optimization cannot be performed directly, because the wealth $X(t)$ is stochastic.

Instead, we will do the usual thing and maximize the **expected value** of the utility derived from current consumption and terminal wealth:

$$\mathbf{E}\left[\int_0^T U(t, c_t)dt + \Phi(X_T)\right]$$

We must address one last constraint: **bankruptcy**. Realistically, we cannot allow the investor's wealth to become negative. A non positive wealth should automatically result in bankruptcy, terminating our investment problem.

To model bankruptcy, we will introduce random variable τ called a a **stopping time**. The stopping time τ is defined as

$$\tau := \inf \{t > 0 | X_t = 0\} \wedge T$$

and it represents the earliest of either

1. the time at which the investor goes bankrupt; or
2. end of the investment horizon.

To sum things up, we are therefore trying to maximize the expectation

$$\mathbf{E}\left[\int_0^\tau U(t, c_t)dt + \Phi(\tau, X_\tau))\right]$$

with

$$\begin{aligned}\Phi(T, X_T) &= \phi(X_T) \\ \Phi(\tau, X_\tau) &= 0 \qquad \text{if } \tau < T\end{aligned}$$

Mathematically, we now finish formalizing our problem by defining the value function $V(t, X_T)$ as the solution to the optimization

$$V(t, X_T) := \max_{w \in \mathbb{R}, c \in \mathbb{R}^+} \mathbf{E}\left[\int_0^\tau U(t, c_t)dt + \Phi(X_\tau)\right]$$

Note that this problem is set in the physical measure $\mathbb{P}$. At no point do we need to change measure to solve this problem.

The associated **Hamilton-Jacobi-Bellman PDE** can be expressed as

$$\frac{\partial V}{\partial t} + \max_{w \in \mathbb{R}, c \in \mathbb{R}^+} \left\{ U(t, c) + wx(\mu - r)\frac{\partial V}{\partial x} + (rx - c)\frac{\partial V}{\partial x} + \frac{1}{2}x^2 w^2 \sigma^2 \frac{\partial^2 V}{\partial x^2} \right\} = 0$$

with boundary condition

$$V(\tau, x) = \Phi(\tau, x)$$

This PDE is parabolic and nonlinear (due to the maximization).

Observe that the following optimization problem is embedded in the PDE

$$\max_{w\in\mathbb{R},c\in\mathbb{R}^+} H(t,x,w,c)$$

where

$$H(t,x,w,c) := U(t,c) + wx(\mu - r)\frac{\partial V}{\partial x} + (rx - c)\frac{\partial V}{\partial x} + \frac{1}{2}x^2w^2\sigma^2\frac{\partial^2 V}{\partial x^2}$$

The first order condition

$$\frac{\partial H(t,x,w,c)}{\partial w} = 0$$
$$\frac{\partial H(t,x,w,c)}{\partial c} = 0$$

gives us the candidate optimal asset allocation w^*

$$w^* = -\frac{\mu - r}{\sigma^2} \frac{\frac{\partial V}{\partial x}}{x \frac{\partial^2 V}{\partial x^2}}$$

and the condition

$$\frac{\partial U}{\partial c} = \frac{\partial V}{\partial x}$$

When we make the simplifying assumption that $\Phi = 0$, an analytical solution exists for several choices of function U. For example, with the choice

$$U(t, c) = e^{-\delta t} c^{\gamma}$$

where $0 < \gamma < 1$ is the risk-aversion coefficient and $\delta > 0$ is a discount rate, the value function is of the form

$$V(t, c) = e^{-\delta t} h(t) x^{\gamma}$$

The function $h(t)$ solves the ODE

$$\begin{aligned} \dot{h}(t) + Ah(t) + Bh(t)^{-\frac{\gamma}{1-\gamma}} \\ h(T) = 0 \end{aligned}$$

with

$$\begin{aligned} A = \frac{1}{2}\frac{\gamma(\mu - r)^2}{\sigma^2(1-\gamma)} + r\gamma - \delta \\ B = 1 - \gamma \end{aligned}$$

The optimal weight and optimal consumption are respectively given by

$$\begin{aligned} w^* &= \frac{\mu - r}{\sigma^2(1-\gamma)} \\ c^* &= xh(t)^{-\frac{1}{1-\gamma}} \end{aligned}$$

What is the practicality of all this? Unfortunately, limited!

This model contains many important ideas and is useful to give an idea of how investors should behave (normative model). However, it is also too simplistic and too restrictive to be practically relevant in the investment management industry.

4.7.2. The stochastic programming approach

The stochastic programming approach involves the design of a multi-period multi-scenario stochastic programming models. This method has the important advantage of acknowledging

- That financial markets are dynamic in nature, and;
- That it is generally more important to avoid financial disaster in difficult times than generating considerable returns in good times. Thus the choice of scenario is important in modelling accurately the left tail of the return distribution.

The stochastic programming approach has produced good results in the investment industry. However, it is generally regarded by economists as an "ad-hoc" model because it is not explicitly utility-based.

A highly readable and insightful overview of stochastic programming and its applications to investment management can

4.7.3. The martingale approach

A probabilistic approach based on a duality principle well-known in (convex) optimization theory also exists.

This duality approach has the advantage of being naturally consistent with economic theory and utility theory[6]. In fact, obtaining the expected utility of wealth is relatively straightforward (in simple cases).

Another advantage of this approach is its power when dealing with problems set in incomplete markets (more on this later).

---

[6] By contrast, the stochastic control approach outlined earlier is only consistent with utility theory because of the two utility functions U and Φ plugged in the problem *en lieu* of the usual cost functions

However, the duality approach also suffers from some major drawbacks:

- the approach is mathematically-demanding and does not have the intuitive appeal of stochastic control;
- obtaining the optimal asset allocation is difficult and time consuming: essentially, the optimal asset allocation is only derived once the entire problem has been solved and all the loose ends have been tied up. This is particularly unpleasant because the optimal asset allocation is by far the most important information in the entire problem!

The only readable overview of the subject can be found in Chapter 3 of Delbaen and Schachermayer [2] (and the optimal asset allocation is no even derived!).

For more information on this approach, the two classic works are Karatzas and Shreve [7] and Kramkov and Schachermayer [3].

4.8. Incomplete Markets and Super-Replication

The original Black-Scholes argument was based on the idea of **market completeness**. Because the cash market (underlying stock plus bank account) is already complete, meaning that you do not have more sources of risk than you can trade, the option is redundant. A dynamically hedged portfolio can therefore be constructed by buying the option, selling an appropriate position in the underlying stock and investing any remaining amount in a bank account.

Market completeness is a simple and reassuring idea.

But it crumbles fast. As soon as we consider interest rates markets, add jumps or stochastic volatility to our underlying stocks, or more generally consider factor models in which the dynamics of our underlying assets depend on a number of economic or statistical factors, we do have more sources of risk than we can directly trade.

Unless we find some trick to “complete” the market[7] or wave away any extra risk[8] we will have to deal with **incomplete markets**.

Incomplete markets are therefore the norm rather than the exception.

---

[7]For example using a second bond to hedge the interest rate risk of the bond we try to price.

[8]Think about Merton’s argument for jump-diffusion option pricing models.

4.8.1. Incomplete markets and PDE methods

The use of PDE methods in incomplete market is limited by the fact that to derive a pricing PDE we generally need to have eliminated all the risk (i.e. the randomness) by

- "completing the market" and using a no-arbitrage argument;
- taking the expectation at some point in the derivation;

so that the dynamics of the hedging portfolio is purely "deterministic".

These two strategies may or may not be possible, and even when they are possible, they may or may not be sensible.

4.8.2. Incomplete markets and probability methods

Probabilistic methods fare relatively better than PDE methods in incomplete markets. The reason is that while PDE methods try to get rid of the randomness as early as possible, probabilistic methods embrace randomness.

The fundamental asset pricing formula can be extended to solve a wide range of pricing problems in incomplete market.

The key difference with the result developed in Lecture 3.3 is that in incomplete markets we will generally find more than one equivalent martingale measure[9]. The main question is therefore how to choose the most appropriate martingale measure for the problem at hand.

---

[9] We had already made this observation in the interest rate market in Lecture 4.3.

4.8.2. Super-replication

The incomplete of market raises the prospects that some financial instruments simply cannot be replicated because they are exposed to a higher number of risks than can be traded on financial markets.

In this case, the conventional no-arbitrage approach fails and we are left with the question of finding other ways of pricing these instruments.

One recent development in this domain is the idea of **super-replication**.

The intuition is the following. Since we cannot hedge the cash-flows of the instrument, can we at least find a portfolio of securities which will deliver at least the same cash-flows as the instrument in all states of the world?

The price today of setting-up the super-replicating portfolio will provide an approximate price for the instrument we are attempting to price.

Super replication effectively transforms a pricing problem into a dynamic portfolio selection problem which is generally solved through either:

- stochastic control[10];
- martingale methods[11];

Implementing either of these techniques requires an intensive use of numerous stochastic analysis techniques and results.

[10] See works by Touzi and Soner

[11] See works by Karatzas, Kramkov, Schachermayer and Shreve

Note that an important implication of the super-replication idea is that the degree of risk we are willing to take in the super replicating portfolio will also be a factor of our risk tolerance and therefore of our utility. Consequently, prices obtained through super-replication will vary from one investor to another.

4.9. To summarize...

Probabilities and PDEs are two complementary approach of the same problem but with different emphases;

Structure and nature of the problem should dictate the choice of method;

Don't hesitate to switch halfway through if it makes sense!

1. **European style derivatives**: tie;
2. **Derivatives with embedded decisions (American & Bermudean)**: PDE (numerical), tie (analytical);
3. **Exotic derivatives**: advantage PDE;
4. **Transaction costs**: advantage PDE;
5. **Risk management**: PDE (localized), probabilities (global);
6. **Jumps-diffusion problems**: advantage probability;
7. **Dynamic portfolio selection**: the perfect blend?;
8. **Incomplete markets and super-replications**: advantage probabilities;

5. Managing Model Risk

A model is an approximation of reality, not reality itself.

Every model requires assumptions. Every tool or approach has strength and weaknesses. So here is a rough guide to managing model risk for risk managers, model validation and financial engineers.

5.1. Understanding the Product/Trade

No mathematics required at this stage.

Follow the cash flow! Draw diagrams of the product or trades to understand where the payments to all counterparties go and when, and in what circumstances, these payments occur.

Once you are able to talk someone through the product/trade, you are ready to move to the next step.

5.2. Writing the Problem

The next step is to clearly formulate the pricing or risk management problem you need to solve.

From the previous step, you know how the product works and therefore which variable or factor has an impact on the product's price and risks.

List these variables and factors from the most important to the least important. This will give you a rough idea of where the degree of complexity should be in your approach. Which factors can you afford to treat as constant in your initial approach? Which factors do you have to model as stochastic?

Write also the relation between the various factors and variables.

- How accurately should you model that dependence?
- Can you afford to assume that minor factors are independent from the more important factors?

5.3. Bringing Problem and Methodologies Together

- Are products similar to the one you are working on already traded? How are they priced?
- In the literature, look for methodologies developed for the type of problem you are working on. How do they work? What are the benefits? What are the pitfalls? Are they appropriate in your case?
- Is there any other tool, technique that might be more appropriate to solve your problem? How do they work? What are the benefits and pitfalls?

5.4. Implementation

Get a rough first version of your model working (at least numerically) as soon as possible to give you

1. something to look at to check if your intuition is correct;
2. more time for a substantial mathematical analysis of the properties of your problem and solution.

DO NOT skip the mathematical analysis! Intuition and numerical methods can be misleading, but solid mathematics generally points you in the right direction.

The next step is to improve your model. Go back to the drawing board, reconsidering your problem. What can you do to make it more realistic?

5.5. Further Tips

- **Diagnose your model**: check that your model works by using it to price simpler securities for which an analytical solution exists;
- **Use more than one approach**: build a second model using a different set of techniques or assumptions. This second model does not need to be as sophisticated but it should be useful in diagnosing what goes right or could go wrong with your model;
- **Be weary of "all-in-one" systems**: if your company uses one such system, always build a little toy model in your favourite software or in Excel to check that the "all-in-one" gives the expected result;
- **Stress test and crash test**: look for what set of conditions make your model tick, tumble and capsize.

In this lecture, we have seen...

- quantitative finance, computational finance, financial mathematics or mathematical finance;
- who does what in QuantLand;
- the link between PDEs and probabilities;
- problems, methods and models;
- steps to manage model risk;

U. Cetin, R. Jarrow, and P. Protter.
Liquidity risk and arbitrage pricing theory.
Finance and Stochastics, 8(3):311–341, 2004.

F. Delbaen and W. Schachermayer.
The Mathematics of Arbitrage.
Springer Finance. Springer, 2006.

D. Kramkov and W. Schachermayer.
Necessary and sufficient conditions in the problem of optimal investment in incomplete markets.
Annals of Applied Probability, 13(4):1504–1516, 2003.

R.C. Merton.
Lifetime portfolio selection under uncertainty: The continuous time case.
Review of Economics and Statitsics, 51:247–257, 1969.

R.C. Merton.

Optimal consumption and portfolio rules in a continuous-time model.
Journal of Economic Theory, 3:373–4113, 1971.

W. Schoutens.
Levy Processes in Finance: Pricing Financial Derivatives.
2003.

S.E. Shreve and I. Karatzas.
Methods of Mathematical Finance.
Stochastic Modeling and Applied Probability. Springer-Verlag, New-York, 2 edition, 1998.

W.T. Ziemba.
The Stochastic Programming Approach to Asset, Liability, and Wealth Management.
Research Foundation Publication. CFA Institute, 2003.