The Black–Scholes Model

1973 $\longrightarrow$ 1997

In this lecture. . .

- the assumptions that go into the Black–Scholes model
- foundations of options theory: delta hedging and no arbitrage
- the Black–Scholes partial differential equation
- the Black–Scholes formulæ for calls, puts and simple digitals
- the meaning and importance of the 'greeks,' delta, gamma, theta.

By the end of this lecture you will be able to

- derive the Black–Scholes partial differential equation

- quote formulæ for simple contracts

- understand the meaning of the common greeks

Introduction

The Black–Scholes equation was the biggest breakthrough in the pricing of options.

The theory is quite straightforward, using just the ideas from stochastic calculus that we have already seen.

The end result is a diffusion-type partial differential equation which can be used for the pricing of many different derivatives.

What determines the value of an option?

The value of an option is a function of the stock price S and time t.

The value of the option is also a function of parameters in the contract, such as the strike price E and the time to expiry $T-t$, T is the date of expiry.

The value will also depend on properties of the asset, such as its drift and its volatility, as well as the risk-free rate of interest:

$$V(S, t; \sigma, \mu; E, T; r).$$

Semi-colons separate different types of variables and parameters.

- S and t are variables;

- σ and μ are parameters associated with the asset price;

- E and T are parameters associated with the particular contract;

- r is a parameter associated with the currency.

For the moment just use $V(S, t)$ to denote the option value.

The Black–Scholes assumptions

- The underlying follows a lognormal random walk with known volatility
- The risk-free interest rate is a known function of time
- There are no dividends on the underlying
- Delta hedging is done continuously
- There are no transaction costs on the underlying
- There are no arbitrage opportunities + more

And more. . .

A very special portfolio

We assume that the asset evolves according to

$$dS = \mu S\, dt + \sigma S\, dX.$$

Then we imagine constructing a special portfolio.

Use Π to denote the value of a portfolio of one long option position and a short position in some quantity Δ, **delta**, of the underlying:

$$\Pi = V(S,t) - \Delta S. \tag{1}$$

$t \to t + dt$

Intuition: Think of moves in S and accompanying move in V, for a call.

How does the value of the portfolio change?

The change in the portfolio value is due partly to the change in the option value and partly to the change in the underlying:

$$d\Pi = dV - \Delta\ dS. \quad (1a)$$

holding Δ fixed across $[t, t+dt]$

- Notice that Δ has not changed during the time step.

$V = V(S,t)$ $\therefore$ dV is obtained by doing Itô on $V(S,t)$

From Itô we have

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt.$$

$$\left(\frac{\partial V}{\partial S} - \Delta \right) dS$$

Thus the portfolio changes by

$$d\Pi = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt - \Delta\, dS. \qquad (2)$$

The right-hand side of (2) contains two types of terms, the deterministic and the random.

- The deterministic terms are those with the dt.

- The random terms are those with the dS. These random terms are the risk in our portfolio.

Elimination of risk: Delta hedging

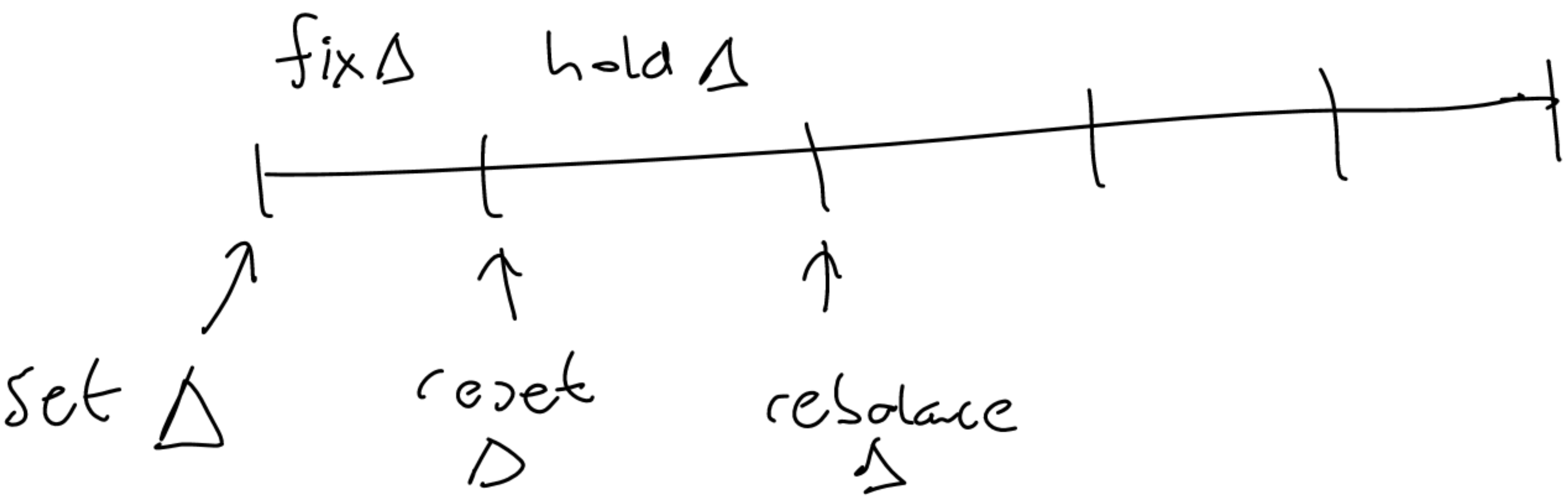
Is there any way to reduce or even eliminate this risk? This can be done in theory by carefully choosing Δ.

$$d(\Delta S) = \Delta\, dS$$

If we choose

$$\Delta = \frac{\partial V}{\partial S} \tag{3}$$

then the randomness is reduced to zero.



- Any reduction in randomness is generally termed **hedging**. The perfect elimination of risk, by exploiting correlation between two instruments (in this case an option and its underlying) is generally called **Delta hedging**.

Delta hedging is an example of a **dynamic hedging** strategy. From one time step to the next the quantity $\frac{\partial V}{\partial S}$ changes, since it is, like V a function of the ever-changing variables S and t.

This means that the perfect hedge must be continually rebalanced.

$$\Delta = \frac{\partial V}{\partial S} = \Delta(S, t)$$

No arbitrage

After choosing the quantity Δ as suggested above, we hold a portfolio whose value changes by the amount

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt. \tag{4}$$

$$\left(V - S \frac{\partial V}{\partial S} \right)$$

- This change is completely *riskless*.

$$d\Pi = r\,\Pi\, dt \quad (4a)$$

risk-free portfolio

$$\therefore \quad (4) \equiv (4a)$$

If we have a completely risk-free change $d\Pi$ in the value Π then it must be the same as the growth we would get if we put the equivalent amount of cash in a risk-free interest-bearing account:

$$d\Pi = r\Pi\, dt. \tag{5}$$

- This is an example of the **no arbitrage** principle.

The Black–Scholes equation

Substituting (1), (3) and (4) into (5) we find that

$$\underbrace{\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)}_{d\pi} dt = r \overbrace{\left(V - S\frac{\partial V}{\partial S}\right)}^{\Pi} dt.$$

On dividing by dt and rearranging we get

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \qquad (6)$$

-Merton

This is the **Black–Scholes equation**.

Observations:

- The Black–Scholes equation equation is a **linear parabolic partial differential equation**

- The Black–Scholes equation contains all the obvious variables and parameters such as the underlying, time, and volatility, but there is no mention of the drift rate μ.

- This means that if two people agree on the volatility of an asset they will agree on the value of its derivatives *even if they have differing estimates of the drift*.

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

Replication

$$V = \Pi + \Delta S$$

Another way of looking at the hedging argument is to ask what happens if we hold a portfolio consisting of just the stock, in a quantity Δ, and cash.

If Δ is the partial derivative of some option value then such a portfolio will yield an amount at expiry that is simply that option's payoff.

In other words, we can use the same Black–Scholes argument to **replicate** an option just by buying and selling the underlying asset.

- This leads to the idea of a **complete market**. In a complete market an option can be replicated with the underlying, thus making options redundant.

Final conditions

The Black–Scholes equation knows nothing about what kind of option we are valuing.

This is dealt with by the **final condition**. We must specify the option value V as a function of the underlying at the expiry date T. That is, we must prescribe $V(S,T)$, the payoff.

For example, if we have a call option then we know that

$$V(S,T) = \max(S - E, 0).$$

Options on dividend-paying equities

Assume that the asset receives a continuous and constant dividend yield, D.

- Thus in a time dt each asset receives an amount $DS\,dt$.

This must be built into the derivation of the Black–Scholes equation.

$$-\Delta D S\,dt$$

$\Pi = V - \Delta S$

Take up the Black–Scholes argument at the point where we are looking at the change in the value of the portfolio:

$$d\Pi = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt - \Delta\, dS - \underline{\underline{D\Delta S\, dt}}.$$

The last term on the right-hand side is the amount of the dividend per asset, $DS\,dt$, multiplied by the number of the asset held, $-\Delta$.

The Δ must still be the rate of change of the option value with respect to the underlying for the elimination of risk.

$dS = (\mu - D)S\,dt + \sigma S\,dX$

End result:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D)S\frac{\partial V}{\partial S} - rV = 0.$$

Solving the equation and the greeks

The Black–Scholes equation has simple solutions for calls, puts and some other contracts. Now we are going to go quickly through the derivation of these formulæ.

The 'delta,' the first derivative of the option value with respect to the underlying, occurs as an important quantity in the derivation of the Black–Scholes equation. In this lecture we see the importance of other derivatives of the option price, with respect to the variables and with respect to some of the parameters.

- These derivatives are important in the hedging of an option position, playing key roles in risk management.

Derivation of the formulæ for calls, puts and simple digitals

The Black–Scholes equation is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{7}$$

This equation must be solved with final condition depending on the payoff: each contract will have a different functional form prescribed at expiry $t = T$, depending on whether it is a call, a put or something else.

The first step in the manipulation is to change from present value to future value terms.

Recalling that the payoff is received at time T but that we are valuing the option at time t this suggests that we write

- $$V(S,t) = e^{-r(T-t)}U(S,t).$$

$$\frac{\partial V}{\partial t} = re^{-r(T-t)}U + e^{-r(T-t)}\frac{\partial U}{\partial t}.$$

$$\frac{\partial V}{\partial S} = e^{-r(T-t)}\frac{\partial U}{\partial S}$$

$$\frac{\partial^2 V}{\partial S^2} = e^{-r(T-t)}\frac{\partial^2 U}{\partial S^2}$$

This takes our differential equation to

B. K. E $$\frac{\partial U}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S} = 0.$$

(Remember this result for later, present valuing means that one of the terms disappears.)

The second step is really trivial. Because we are solving a backward equation we'll write

Time to expiry

- $\tau = T - t.$

$$\frac{\partial}{\partial t} = \frac{\partial \tau}{\partial t} \frac{\partial}{\partial \tau} = -\frac{\partial}{\partial \tau}$$

This now takes our equation to

F.K.E

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S}.$$

When we first started modeling equity prices we used intuition about the asset price *return* to build up the stochastic differential equation model. Let's go back to examine the return and write

- $$\xi = \log S.$$

With this as the new variable, we find that

$$\frac{\partial}{\partial S} = e^{-\xi}\frac{\partial}{\partial \xi} \quad \text{and} \quad \frac{\partial^2}{\partial S^2} = e^{-2\xi}\frac{\partial^2}{\partial \xi^2} - e^{-2\xi}\frac{\partial}{\partial \xi}.$$

Now the Black–Scholes equation becomes

$$\frac{\partial U}{\partial \tau} = \tfrac{1}{2}\sigma^2\frac{\partial^2 U}{\partial \xi^2} + \left(r - \tfrac{1}{2}\sigma^2\right)\frac{\partial U}{\partial \xi}.$$

The last step is simple, but the motivation is not so obvious. Write

- $$x = \xi + \left(r - \tfrac{1}{2}\sigma^2\right)\tau \quad \text{and} \quad U = W(x,\tau).$$

$\tau = \tau$

This is just a 'translation' of the co-ordinate system. It's a bit like using the forward price of the asset instead of the spot price as a variable.

After this change of variables the Black–Scholes becomes the simpler

$$\frac{\partial p}{\partial t'} = c^2 \frac{\partial^2 p}{\partial y'^2}$$

$$\boxed{\frac{\partial W}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 W}{\partial x^2}.} \tag{8}$$

And you've seen this equation before!

You've even solved it to find a special solution—out of the infinite number of possible solutions—and exactly the same solution will be needed here. (Lucky!)

Fundamental / Source sol^{n}

2 simple sol^{ns} of B.S.E are

(1) stock ie $V = S$

(2) Cash $V = S_0 e^{rt}$

To summarize,

$$V(S,t) = e^{-r(T-t)}U(S,t) = e^{-r\tau}U(S,T-\tau) = e^{-r\tau}U(e^{\xi},T-\tau)$$

$$= e^{-r\tau}U\left(e^{x-\left(r-\frac{1}{2}\sigma^2\right)\tau}, T-\tau\right) = e^{-r\tau}\underbrace{W(x,\tau)}.$$

where

W satisfies

(8) on

slide 26

We are going to derive an expression for the value of any option whose payoff is a known function of the asset price at expiry.

This includes calls, puts and digitals. This expression will be in the form of an integral.

For special cases, we'll see how to rewrite this integral in terms of the cumulative distribution function for the Normal distribution. This is particularly useful since the function can be found on spreadsheets, calculators and in the backs of books.

But there are two steps before we can write down this integral.

- The first step is to find a special solution of (8), called the fundamental solution. This solution has useful properties.

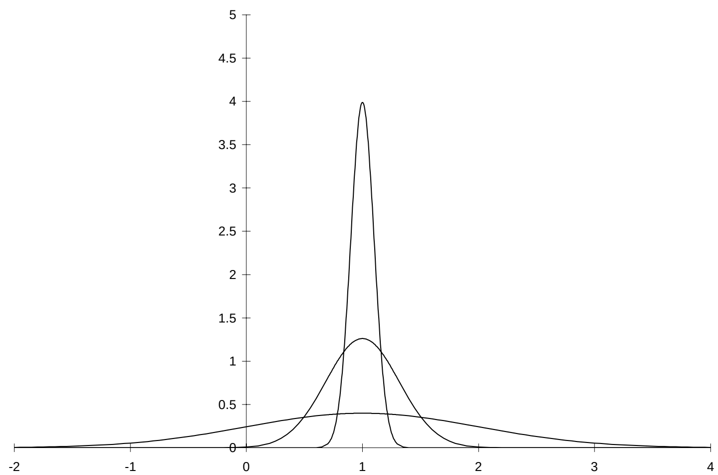
- The second step is to use the linearity of the equation and the useful properties of the special solution to find the *general solution* of the equation.

The first step is easy, just recall solving the equation from the earlier lecture. The solution we want is

- $$W_f(x,\tau;x') = \frac{1}{\sqrt{2\pi\tau}\,\sigma} e^{-\frac{(x-x')^2}{2\sigma^2\tau}}.$$

As you know, this is the probability density function for a Normal random variable x having mean of x' and standard deviation $\sigma\sqrt{\tau}$.

And this also strongly hints at a relationship between option values and probabilities. More anon!



Above is plotted W_f as a function of x' for several values of τ.

At $x' = x$ the function grows unboundedly, and away from this point the function decays to zero, as $\tau \to 0$.

Although the function is increasingly confined to a narrower and narrower region its area remains fixed at one.

- These properties of decay away from one point, unbounded growth at that point and constant area, result in a **Dirac delta function** $\delta(x' - x)$ as $\tau \to 0$.

The delta function has one important property, namely

$$\int \delta(x' - x) g(x') \, dx' = g(x)$$

where the integration is from any point below x to any point above x.

Thus the delta function 'picks out' the value of g at the point where the delta function is singular i.e. at $x' = x$.

In the limit as $\tau \to 0$ the function W becomes a delta function at $x = x'$. This means that

$$\lim_{\tau \to 0} \frac{1}{\sigma\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} e^{-\frac{(x'-x)^2}{2\sigma^2\tau}} g(x')\,dx' = g(x).$$

Whoa! This is tricky!

I am going to ‘cut to the chase’ and quote the solution:

$V(S,t) = e^{-r(T-t)} U(S,t)$

- $$V(S,t) = \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}}$$

$$\int_0^\infty e^{-\left(\log(S/S')+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2/2\sigma^2(T-t)} \text{Payoff}(S') \frac{dS'}{S'}. \qquad (9)$$

This ‘formula’ works for any European, non path-dependent, option on a single lognormal underlying asset, all you need to know is the payoff function.

probability of a stock S at time t

$\longrightarrow$ S' at expiry T

Observations

- This is a general formula (see above conditions on type of option)

- It is of the form of a) a discounting term multiplied by b) the integral of the payoff multiplied by c) another function

- This other ‘function’ is known as a Green’s function

- This function can be interpreted as a probability

- The whole expression can be interpreted as the present value of the expected payoff

Let's look at special cases.

<u>**Formula for a call**</u>

The call option has the payoff function

$$\text{Payoff}(S) = \max(S - E, 0).$$

Payoff(S')

Expression (9) can then be written as

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}}\int_E^\infty e^{-\left(\log(S/S')+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2/2\sigma^2(T-t)}(S'-E)\frac{dS'}{S'}.$$

$$\int_{-\infty}^{\cdots} e^{-\frac{1}{2}u^2}\,du$$

Return to the variable $x' = \log S'$, to write this as

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}}\int_{\log E}^{\infty} e^{-\left(-x'+\log S+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2/2\sigma^2(T-t)}\left(e^{x'}-E\right)dx'$$

$$= \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}}\int_{\log E}^{\infty} e^{-\left(-x'+\log S+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2/2\sigma^2(T-t)}e^{x'}\,dx'$$

$$-E\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}}\int_{\log E}^{\infty} e^{-\left(-x'+\log S+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2/2\sigma^2(T-t)}\,dx'.$$

Both integrals in this expression can be written in the form

$$\int_d^\infty e^{-\frac{1}{2}{x'}^2}\,dx'$$

for some d (the second is just about in this form already, and the first just needs a completion of the square).

Thus the option price can be written as two separate terms involving the cumulative distribution function for a Normal distribution:

$$\text{Call option value } = SN(d_1) - Ee^{-r(T-t)}N(d_2)$$

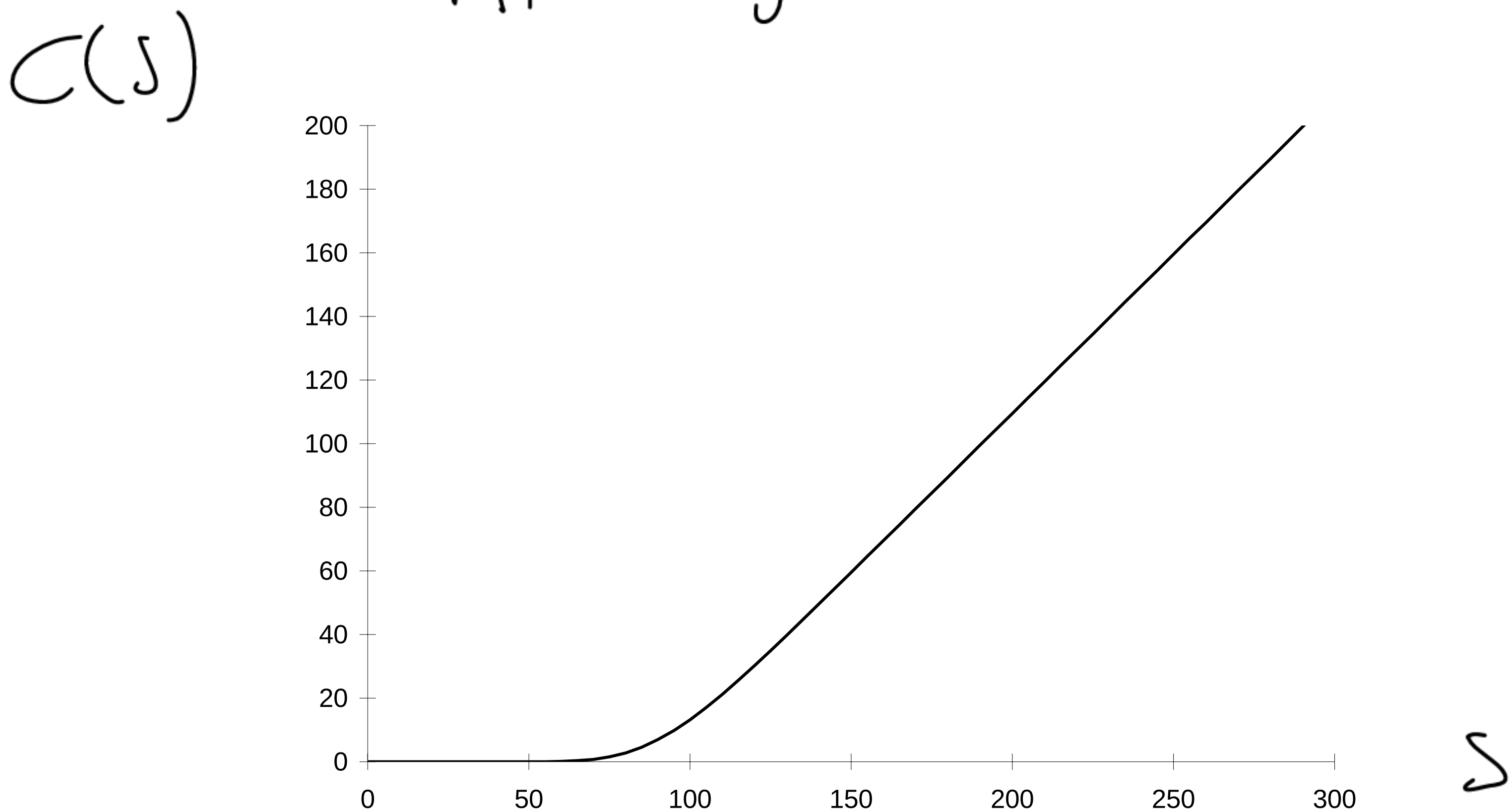
where

$$d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \quad \text{and}$$

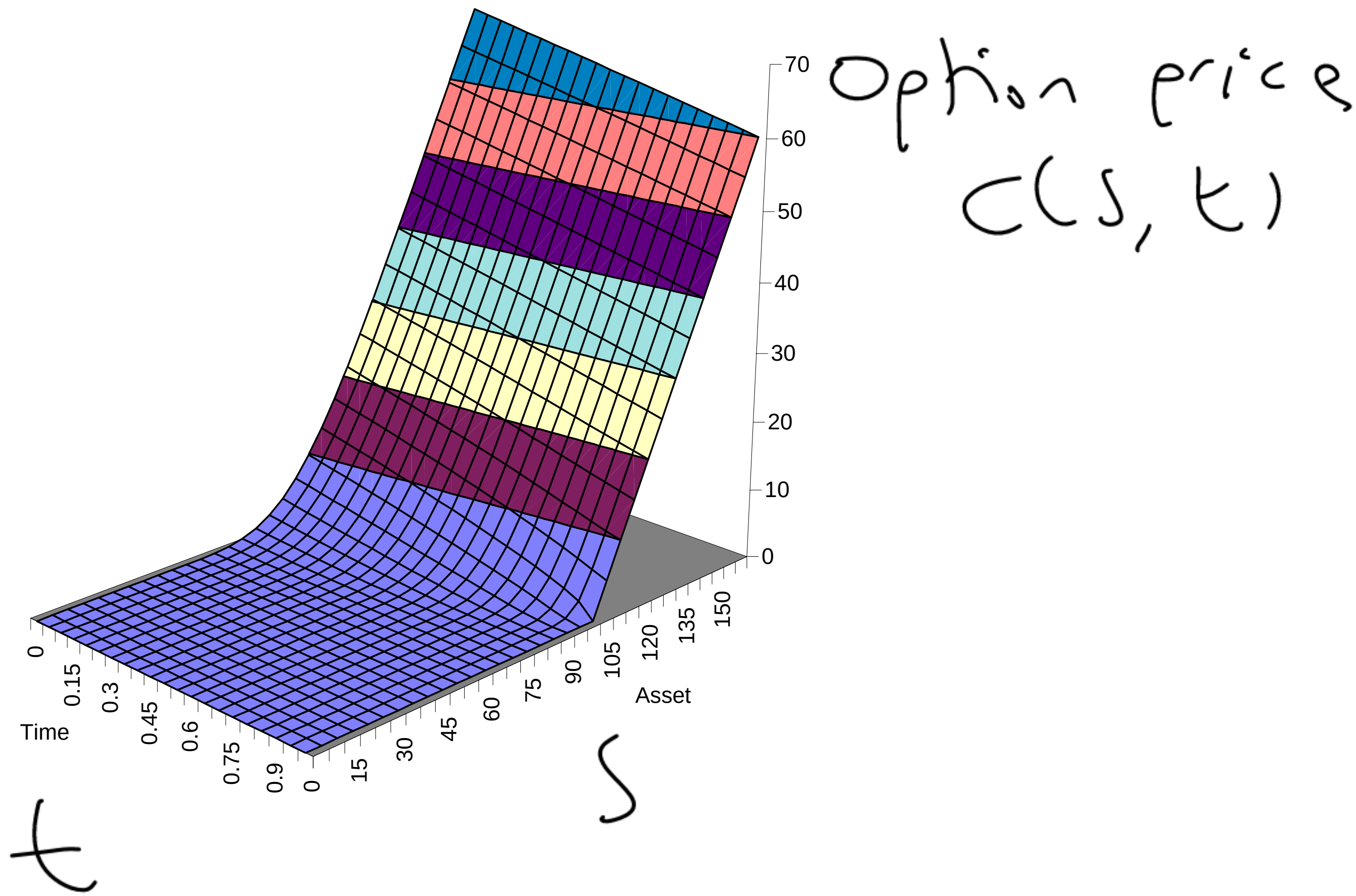
$$d_2 = \frac{\log(S/E) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\phi^2} d\phi.$$

At any time t



The value of a call option as a function of the underlying at a fixed time before expiry.



The value of a call option as a function of asset and time.

When there is continuous dividend yield on the underlying, or it is a currency, then

$f = S e^{(r-D)(T-t)}$

$\tau = T - t$

Call option value

$$Se^{-D(T-t)}N(d_1) - Ee^{-r(T-t)}N(d_2)$$

$$d_1 = \frac{\log(S/E)+(r-D+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\log(S/E)+(r-D-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t}$$

When the asset is 'at-the-money forward,' i.e. $S = Ee^{-(r-D)(T-t)}$, and the option is close to expiration then there is a simple approximation for the call value (Brenner & Subrahmanyam, 1994):

$$\text{Call} \approx 0.4\, Se^{-D(T-t)}\sigma\sqrt{T-t}.$$

$$\text{Ex: } C - P = S - Ee^{-r(T-t)}$$

Formula for a put

$$P = C - S + Ee^{-r(T-t)}$$

The put option has payoff

$$\text{Payoff}(S) = \max(E - S, 0).$$

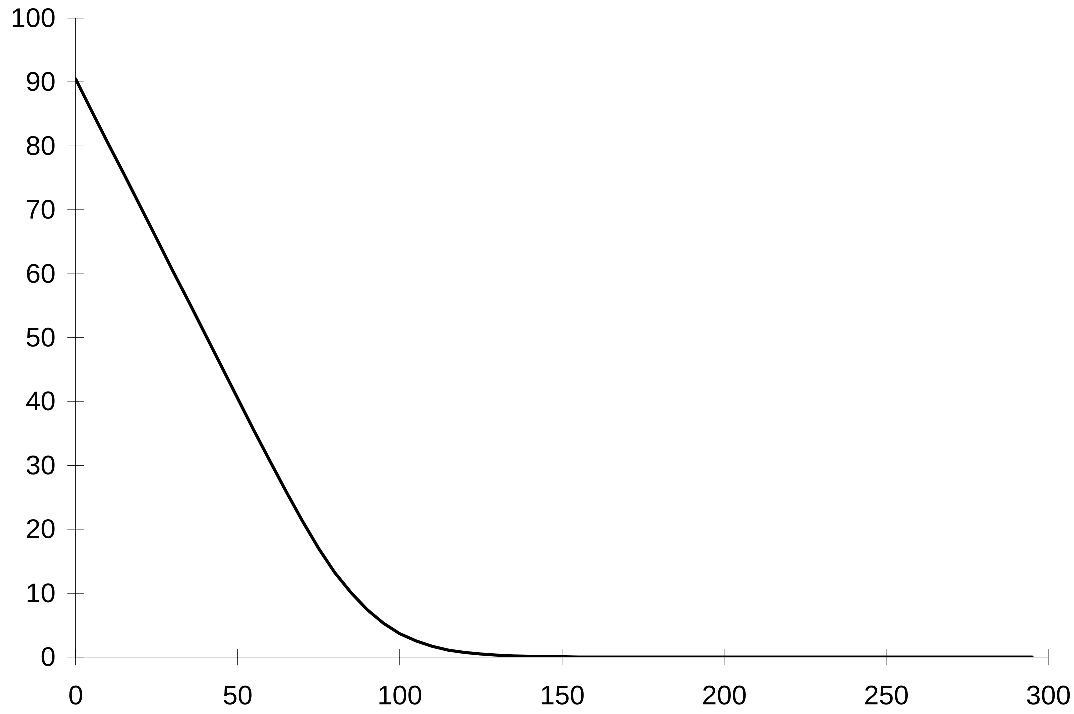
The value of a put option can be found in the same way as above, or using put-call parity

$$\text{Put option value } = -SN(-d_1) + Ee^{-r(T-t)}N(-d_2),$$

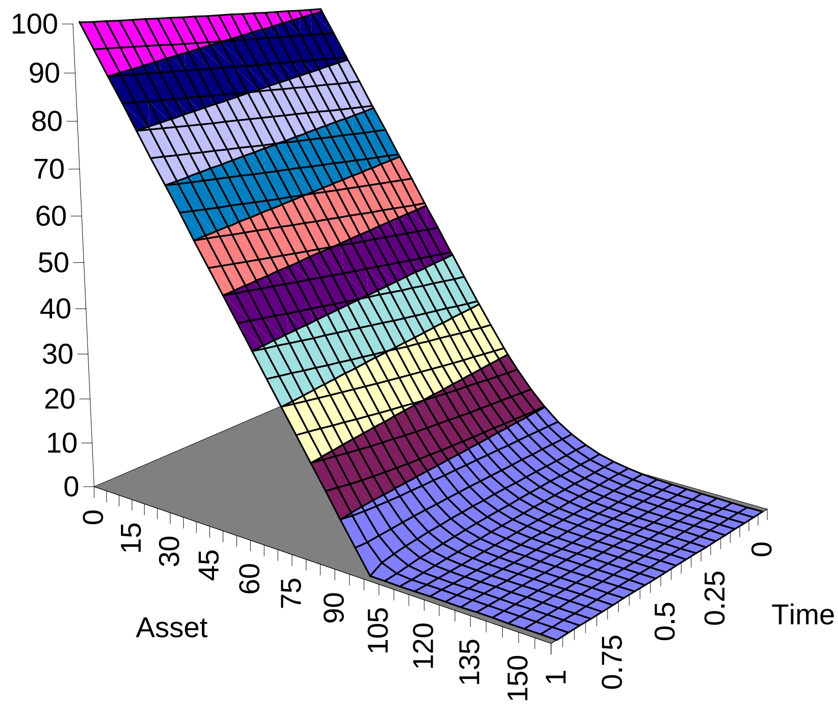
with the same d_1 and d_2.

$$N(x) + N(-x) = 1$$

$$\text{i.e. } N(-x) = 1 - N(x)$$



The value of a put option as a function of the underlying at a fixed time to expiry.



The value of a put option as a function of the underlying and time to expiry.

When there is continuous dividend yield on the underlying, or it is a currency, then

Put option value

$$-Se^{-D(T-t)}N(-d_1) + Ee^{-r(T-t)}N(-d_2)$$

When the asset is at-the-money forward and the option is close to expiration the simple approximation for the put value (Brenner & Subrahmanyam, 1994) is

$$\text{Put } \approx 0.4\, Se^{-D(T-t)}\sigma\sqrt{T-t}.$$

Formula for a binary call

The binary call has payoff

$$\text{Payoff}(S) = \mathcal{H}(S - E), = \begin{cases} 1 & S_T > E \\ 0 & \text{otherwise} \end{cases}$$

where $\mathcal{H}$ is the Heaviside function taking the value one when its argument is positive and zero otherwise.

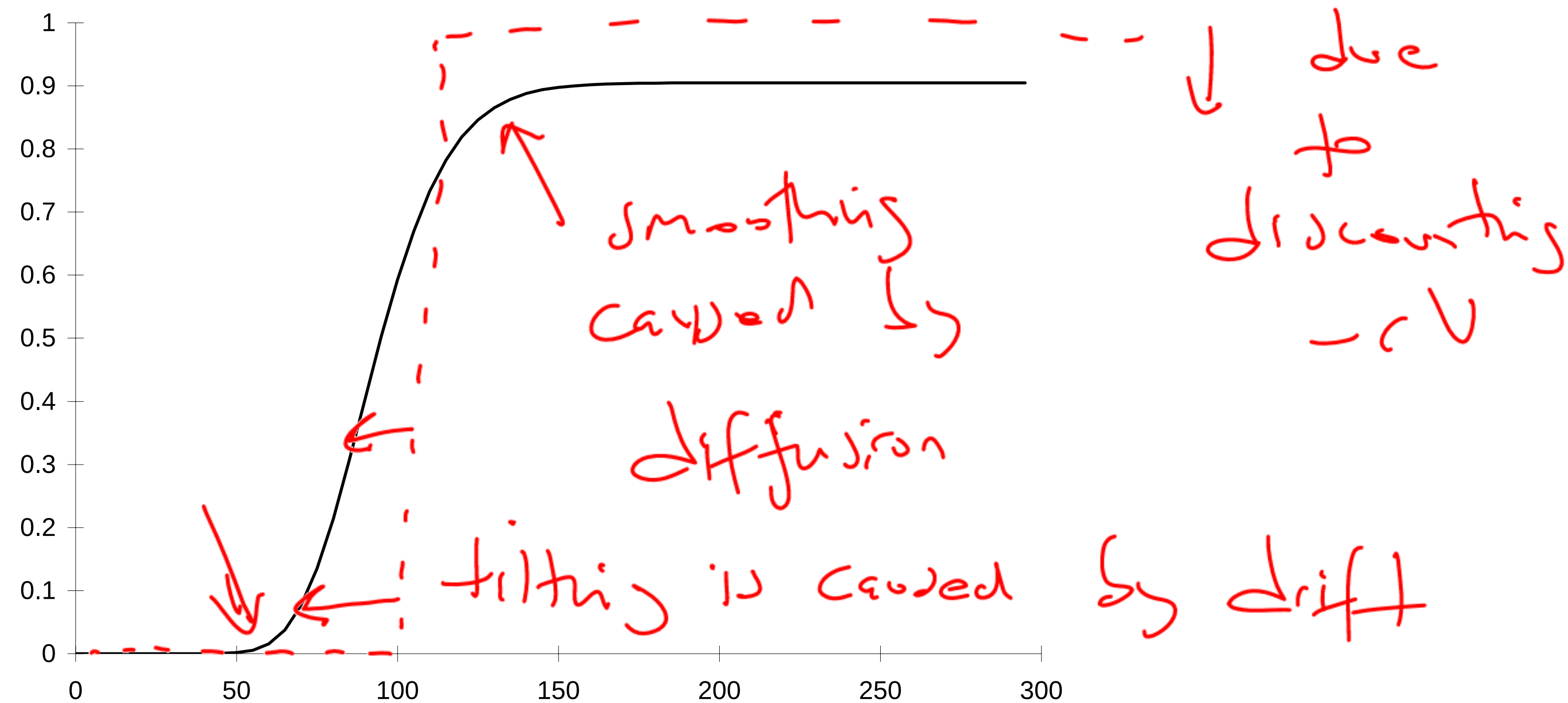
Incorporating a dividend yield, we can write the option value as

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_{\log E}^{\infty} e^{-\left(x' - \log S - \left(r - D - \frac{1}{2}\sigma^2\right)(T-t)\right)^2 / 2\sigma^2 (T-t)} \, dx'.$$

This term is just like the second term in the call option equation and so. . .

Binary call option value

$$e^{-r(T-t)}N(d_2)$$



The value of a binary call option.

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}+rS\frac{\partial V}{\partial S}-rV=0$$

$B_c + B_p =$

Formula for a binary put

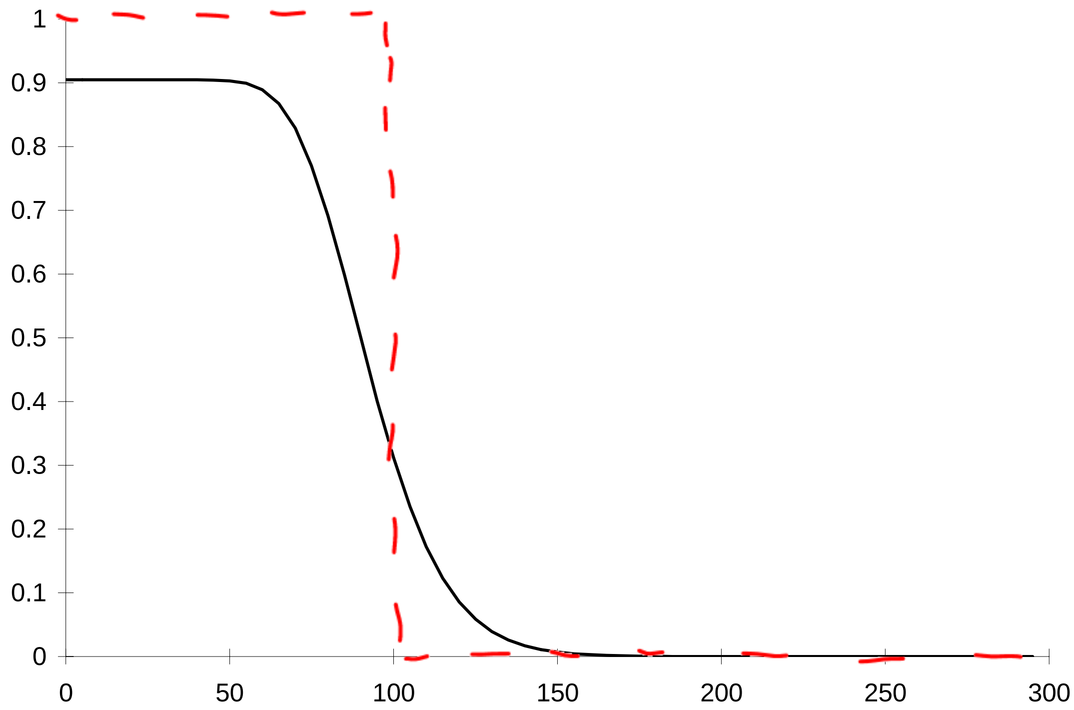
$e^{-r(T-t)} N(d_2) + e^{-r(T-t)} (1 - N(d_1))$

The binary put has a payoff of one if $S < E$ at expiry. It has a value of

Binary put option value $e^{-r(T-t)}(1 - N(d_2))$

A binary call and a binary put must add up to the present value of \$1 received at time T.

$$\mathcal{H}(E - S) = \begin{cases} 1 & \text{if } S_T < E \\ 0 & \text{otherwise} \end{cases}$$



The value of a binary put option.

Delta

The **delta** of an option or a portfolio of options is the sensitivity of the option or portfolio to the underlying. It is the rate of change of value with respect to the asset:

$$\boxed{\Delta = \frac{\partial V}{\partial S}}$$

Here V can be the value of a single contract or of a whole portfolio of contracts. The delta of a portfolio of options is just the sum of the deltas of all the individual positions.

- The theoretical device of delta hedging for eliminating risk is far more than that, it is a very important practical technique.

Delta hedging means holding one of the option and short a quantity Δ of the underlying.

Delta can be expressed as a function of S and t.

This function varies as S and t vary.

- This means that the number of assets held must be continuously changed to maintain a **delta neutral** position, this procedure is called **dynamic hedging**.

Changing the number of assets held requires the continual purchase and/or sale of the stock. This is called **rehedging** or **rebalancing** the portfolio.

Here are some formulæ for the deltas of common contracts (all formulæ assume that the underlying pays dividends or is a currency):

Deltas of common contracts

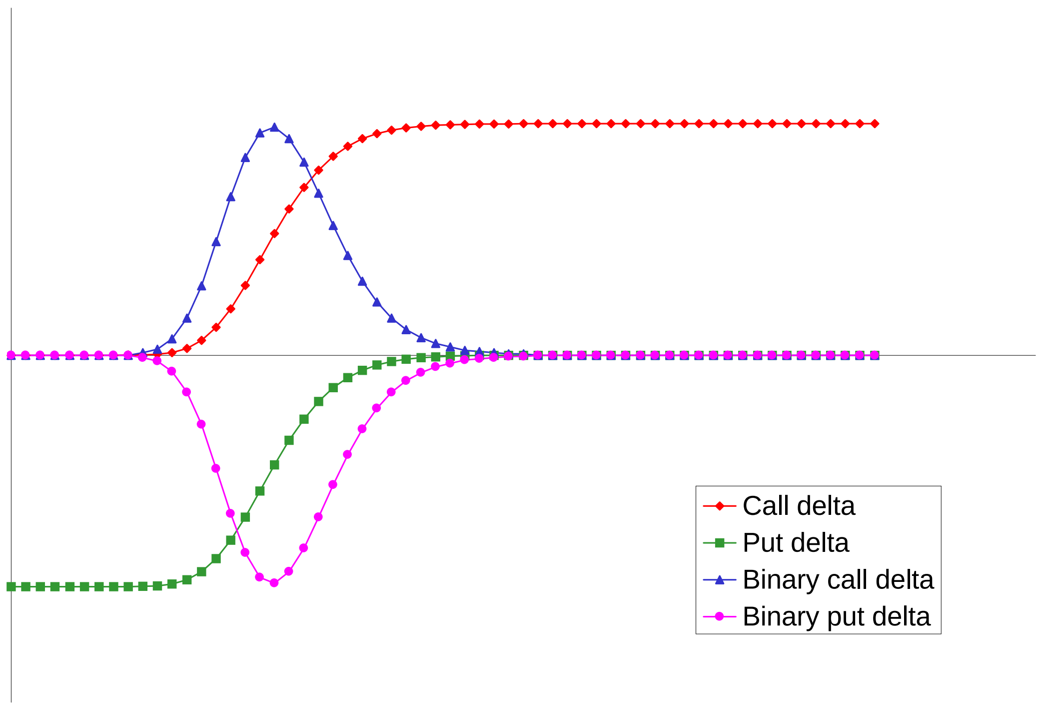
Call $e^{-D(T-t)}N(d_1)$

Put $e^{-D(T-t)}(N(d_1)-1)$

Binary call $\frac{e^{-r(T-t)}N'(d_2)}{\sigma S\sqrt{T-t}}$

Binary put $-\frac{e^{-r(T-t)}N'(d_2)}{\sigma S\sqrt{T-t}}$

$$N'(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$$



The deltas of a call, put and binary options. (The deltas of the binaries have been scaled.)

Gamma

The **gamma**, Γ, of an option or a portfolio of options is the second derivative of the position with respect to the underlying:

$$\boxed{\Gamma = \frac{\partial^2 V}{\partial S^2}}$$

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

This is, of course, just

$$\frac{\partial \left(\frac{\partial V}{\partial S} \right)}{\partial S}.$$

Since gamma is the sensitivity of the delta to the underlying it is a measure of by how much or how often a position must be rehedged in order to maintain a delta-neutral position.

Because costs can be large and because one wants to reduce exposure to model error it is natural to try to minimize the need to rebalance the portfolio too frequently.

Since gamma is a measure of sensitivity of the hedge ratio Δ to the movement in the underlying, the hedging requirement can be decreased by a gamma-neutral strategy.

- This means buying or selling more *options*, not just the underlying.

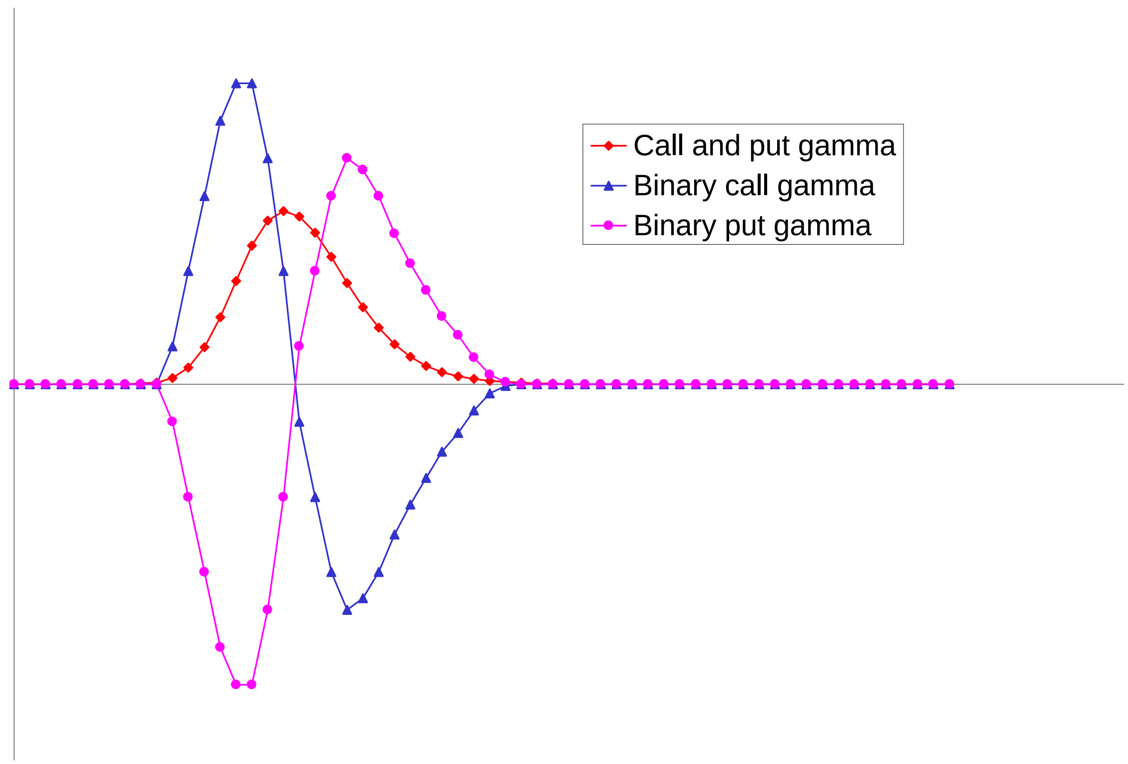
Because the gamma of the underlying (its second derivative) is zero, we cannot add gamma to our position just with the underlying.

- We can have as many options in our position as we want, we choose the quantities of each such that both delta and gamma are zero.

Here are some formulæ for the gammas of common contracts:

Gammas of common contracts

Call	$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$
Put	$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$
Binary call	$-\frac{e^{-r(T-t)}d_1 N'(d_2)}{\sigma^2 S^2(T-t)}$
Binary put	$\frac{e^{-r(T-t)}d_1 N'(d_2)}{\sigma^2 S^2(T-t)}$



The gammas of a call, put and binary options.

Theta

Theta, Θ, is the rate of change of the option price with time.

$$\boxed{\Theta = \frac{\partial V}{\partial t}}$$

$$\Theta = \frac{\partial V}{\partial \tau}$$

The theta is related to the option value, the delta and the gamma by the Black–Scholes equation. In a delta-hedged portfolio the theta contributes to ensuring that the portfolio earns the risk-free rate.

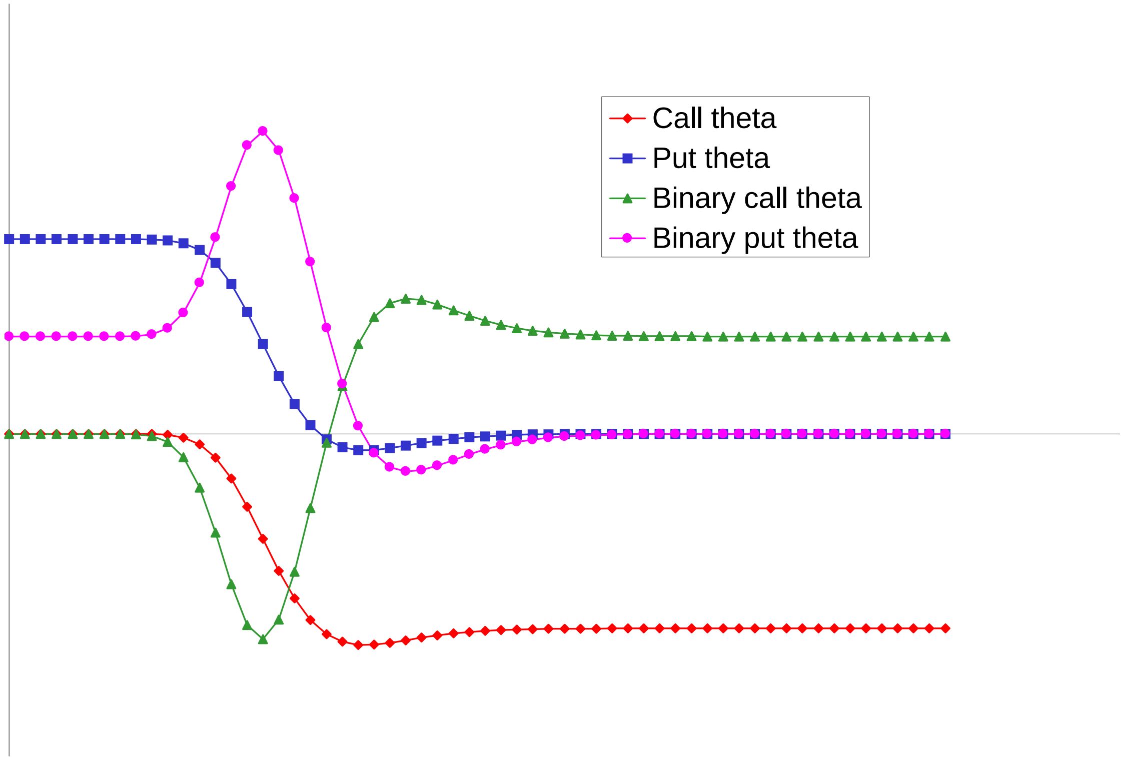
"Theta bleed"

$$\tau = T - t$$

Here are some formulæ for the thetas of common contracts:

Thetas of common contracts

Call	$-\frac{\sigma S e^{-D(T-t)} N'(d_1)}{2\sqrt{T-t}} + DSN(d_1)e^{-D(T-t)} - rEe^{-r(T-t)}N(d_2)$
Put	$-\frac{\sigma S e^{-D(T-t)} N'(-d_1)}{2\sqrt{T-t}} - DSN(-d_1)e^{-D(T-t)} + rEe^{-r(T-t)}N(-d_2)$
Binary call	$re^{-r(T-t)}N(d_2) + e^{-r(T-t)}N'(d_2)\left(\frac{d_1}{2(T-t)} - \frac{r-D}{\sigma\sqrt{T-t}}\right)$
Binary put	$re^{-r(T-t)}(1 - N(d_2)) - e^{-r(T-t)}N'(d_2)\left(\frac{d_1}{2(T-t)} - \frac{r-D}{\sigma\sqrt{T-t}}\right)$



The thetas of a call, put and binary options.

Summary

Please take away the following important ideas

- Using tools from stochastic calculus we can build up an option pricing model from our lognormal asset price random walk model

- There are some 'simple' formulæ for the prices of simple contracts

- The greeks are important measures of the sensitivities of the option value to variables and parameters

CQF Module 3: Martingales II

Black-Scholes All Over Again

CQF

In this lecture...

... we will apply probabilistic and martingale methods to the pricing of European stock and index options in complete markets. we will see:

- computing the price of a derivative as an expectation;
- Girsanov's theorem and change of measures;
- the fundamental asset pricing formula;
- the Black-Scholes Formula;
- the Feynman-Kăc formula;
- extensions to Black-Scholes: dividends and time-dependent parameters;
- Black's formula for options on futures;

Introduction

In Lecture 3.1. we developed the Black-Scholes derivative pricing approach from a PDE perspective. In this session, we will do the same but from the vantage point of probability theory.

While the concepts used to derive the Black-Scholes PDE are straightforward in their financial and economic interpretation, we will see that the probabilistic approach requires a greater level of abstraction.

However, where the probabilistic approach really shines is in the derivation of the Black-Scholes pricing formula where the fast functional manipulations of the PDE world are replaced by an elegant probabilistic interpretation.

1. The World of Black-Scholes

The setting of the probabilistic approach is not different from the setting seen in Lecture 3.1. It is simply expressed more formally.

We are in a financial market comprised of two traded assets, the risk-free asset B and the underlying security S. Our objective is to value a financial derivative (a.k.a. contingent claim) χ whose exercise value is a function $G(S(T))$ of the underlying security S at the expiry time T.

Since we are using a probabilistic framework, we will start by defining a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

1.1. Risk-Free Asset

The risk-free asset B evolves according to the dynamics

$$\boxed{dB_t = rB_t dt, \quad B(0) = B_0 = 1}$$

$$\int_0^T \frac{dB_t}{B_t} = r\int_0^T dt$$

written as an integral, the value of the risk-free asset at time t is given by:

$$\boxed{B(t) = B(0)e^{rt}}$$

$$\Leftrightarrow \ln B_T - \ln B(0) = rT$$

$$\Leftrightarrow B(T) = e^{rT}$$

The normal interpretation of the risk-free asset is as a short-term government bond or as a repo trade, r denoting the short-term (in fact instantaneous!) interest rate.

Note that the asset B is effectively risk-free since its dynamics only has a drift but no diffusion.

1.2. Underlying Asset

$(\Omega, \mathcal{F}, \mathbb{P})$

The dynamics of the underlying asset S is modelled through a geometric Brownian motion (GBM):

$$\boxed{dS_t = \mu S_t dt + \sigma S_t dX_t, \quad S(0) = S_0}$$

$Y(t) = \ln(S_t)$

Itô

$dY(t) = \dots$

where $X(s)$ is a Brownian motion over $[0, T]$. Written as an integral, the value of the underlying asset at time t is given by:

$$\boxed{S(t) = S_0 e^{\mu t - \frac{1}{2}\sigma^2 t + \sigma X_t}}$$

$Y(t) = \ln(S(t))$

Apply Itô to the function $f(x) = \ln(x)$ and to $Y(t)$

1.3. Derivative Security

To keep the discussion general but tractable, we will assume that the derivative security χ

1. matures at time T;
2. with payoff at maturity given by a function G of $S(T)$, i.e. $\chi(T, S(T)) = G(S(T))$;
3. is of European style;

1.4. Other Assumptions

We will also assume the usual at this stage:

1. short selling is allowed;
2. the market is frictionless, i.e. there is no transaction cost, taxes, lack of liquidity or constraints on shortselling;
3. as a consequence of 1. and 2., the assets can be bought or shorted in unlimited quantity;
4. fractional trading is authorized;
5. trading is conducted continuously;
6. there is no dividend and the coefficients r, μ and σ are constant.

2. The Fundamental Asset Pricing Formula

The most intricate part of this lecture is the derivation of the Fundamental Asset Pricing Formula.

The Fundamental Asset Pricing Formula basically states that

$$\text{Value of Asset} = \mathbb{E}^{\substack{\text{Some equivalent} \\ \text{probability measure}}}\left[\text{PV(Cash Flows)}\right]$$

This fundamental result not only provides a uniform view of asset valuation, but it is also the first step in establishing the Black-Scholes formula.

Admittedly, the Black-Scholes problem is special, because it can be solved analytically. In fact, many valuation problems do not have a closed-form solution. In this case, the Fundamental Asset Pricing Formula can be used as a base for the application of numerical methods, and in particular Monte Carlo simulations.

To derive the Fundamental Asset Pricing Formula, we will need to:

1. Define self-financing trading strategies and arbitrage strategies;
2. Discount the asset price;
3. Find an equivalent probability measure in which this asset price can be expressed as a simple (discounted) expectation;
4. Use the no-arbitrage condition to value the derivative;

$\Pi(t) = C(t) - \Delta S(t)$

$V(t) \Rightarrow$ Stocks & Bonds

2.1 Self-Financing Trading Strategies and Arbitrage Strategies

As in Lecture 3.1, our starting point will be the definition of a portfolio, or trading strategy which replicates the dynamics of the derivative.

# of shares that you own at time t

$\phi_t = (\phi_t^S, \phi_t^B)$ # of "units" held in the Bank account at time t

Definition (Trading Strategy)

A **trading strategy** $\phi_t = (\phi_t^S, \phi_t^B)$ is a pair of stochastic processes progressively measurable over the time interval $[0, T]$ where

- ϕ_t^S represents the number of units of the underlying asset S held at time $t \in [0, T]$;
- ϕ_t^B represents the number of units of the risk-free asset B held at time $t \in [0, T]$.

$$V(t) = \phi_t^S \cdot S(t) + \phi_t^B \cdot B(t)$$

This definition of trading strategies is quite wide as it only indicates the number of shares and the amount invested in/borrowed from the bank.

In particular, this definition does not say anything about flow to and from the portfolio such as:

- consumption - cash taken out of the portfolio to finance personal consumption (such as paying for grocery);
- contributions - cash added periodically to the portfolio (such as receiving a salary).

Because cash inflows and outflows are rather difficult to track accurately[1], we will concentrate on portfolios with **NO** cash inflows and **NO** cash outflows.

Such portfolios are called **self-financing**.

[1] Especially since one would need to actually track the *utility* generated by these cash flows rather than the flows themselves.

$dB(t) = r B(t) dt$

$dS(t) = \mu S(t) dt + \sigma S(t) dX_t$

drift

$\sigma S(t) \times \sqrt{dt} \times Z$

$Z \sim N(0,1)$

Formally,

Definition (Self-Financing Trading Strategy)

A trading strategy $\phi_t = (\phi_t^S, \phi_t^B)$ defined over the time interval $[0, T]$ is **self-financing** if its wealth process $V(\phi)$ given by

trading strategy →

$$\boxed{V_t(\phi) = \phi_t^S S_t + \phi_t^B B_t \quad \forall t \in [0, T]}$$

satisfies the condition:

$$\underbrace{V_t(\phi) = V_0(\phi)}_{V_t(\phi) - V_0(\phi)} + \int_0^t \phi_u^S dS_u + \int_0^t \phi_u^B dB_u \quad \forall t \in [0, T] \tag{1}$$

Self financ-ing condition

$\int_0^t dV(s) = V_t(\phi) - V_0(\phi)$

dynamics

$$\boxed{dV_t(\phi) = \phi_t^S dS_t + \phi_t^B dB_t}$$

In plain English, what condition (1) is saying is that no cash is ever added or taken out of the portfolio. Hence, any purchase (or sale) of the underlying asset will be financed by (or invested in) the risk-free asset. portfolio is therefore truly **self-financing**.

From a mathematical standpoint, note that since S is stochastic the first integral in condition (1) is understood in the Itô sense. The second integral is a "regular" pathwise Riemann integral.

As we have already seen in the binomial model and the PDE approach, the absence of arbitrage opportunities on our market is key to pricing assets.

But what is an arbitrage opportunity?

0 T

$V(0) = 0$

$\to V(T) > 0$

$\to V(T) = 0$

Do not lose money

In essence, an arbitrage opportunity exists if

- we can constitute a portfolio at time 0 at no cost;
- by time T, the portfolio cannot have lost money, but;
- the portfolio has a positive probability of having gained money by time T;

This definition is both wider and more realistic than the traditional academic definition of arbitrage which states that an arbitrage portfolio <u>must</u> deliver some gains by time T.

We now formalize our insight:

Definition (Arbitrage Opportunity)

An **arbitrage opportunity** is a self-financing portfolio ϕ such that

$$V_0(\phi) \quad = \quad 0$$

no money down

$$\underbrace{P\left[V_T(\phi) > 0\right]}_{\text{Proba of Profit}} > 0$$

and

$$\underbrace{P\left[V_T(\phi) < 0\right]}_{\text{Proba of a loss}} = 0$$

We say that a market is **arbitrage-free** if no arbitrage opportunity exists.

$$S^*(t) = \frac{S(t)}{B(t)} = \frac{S(t)}{e^{rt}} = e^{-rt} S(t)$$

2.2. Discount the Asset Process

Rather than using the nominal stock price $S(t)$ in our analysis, we will use the **discounted** stock price $S^*(t)$ defined as

$$S^*(t) = \frac{S(t)}{B(t)} = \frac{S(t)}{e^{rt}}$$

But why consider the discounted price rather than the current price?

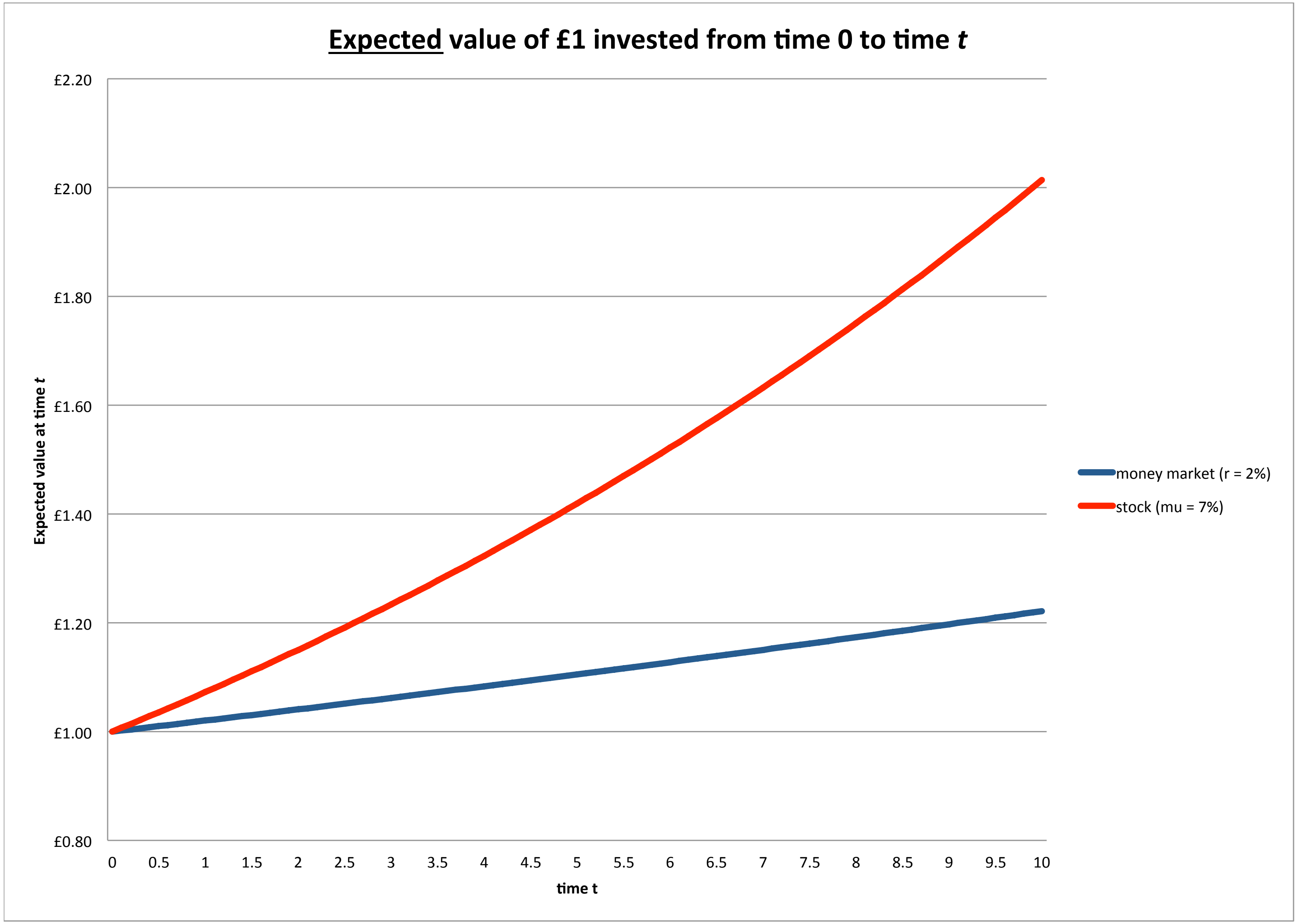
Because of the **time value of money**.

$$S(t) = S(0) \exp\left\{ (\mu - r)t - \frac{1}{2}\sigma^2 t + \sigma X(t) \right\}$$

Time value of money encodes a time dependence into the cash flow structure or the price dynamics of assets. It is helpful in determining time-dependent preferences between

1. receiving a cash flow C_0 at time 0; and,
2. receiving a cash flow C_T at time T.

In our model, time value of money is represented by the risk-free rate r. So an investor should be indifferent between the two cashflows if $C_0 = e^{-rt} C_T$.


Expected value of £1 invested from time 0 to time t
Expected value at time t
£2.20
£2.00
£1.80
£1.60
£1.40
£1.20
£1.00
£0.80
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5
8
8.5
9
9.5
10
time t
money market (r = 2%)
stock (mu = 7%)

Because of time value of money, the risk-free rate is already embedded inside the drift of all financial assets. So we want to remove it and consider the underlying asset's dynamics freed from the time-dependent relationship encoded in the time value of money.

For any time $t \in [0, T]$, the discounted asset process is given by

$Y(t) = \ln S^*(t)$

$$S^*(t) = S_0^* e^{\left(\mu - r - \frac{1}{2}\sigma^2\right)t + \sigma X(t)}$$

with an equivalent SDE given by

$$\boxed{dS^*(t) = (\mu - r)S^*(t)dt + \sigma S^*(t)dX(t), \quad S^*(0) = S_0^*}$$

Looking at the SDE, we see that the effect of discounting is actually to remove the risk-free rate from the drift, or said otherwise, to remove the portion of the stock returns "guaranteed" by the time value of money.

2.3. Change the Measure

In Lecture 2.5, we have seen that martingales are very nice processes to work with since they have many enviable properties:

- they are driftless, so all we have to think about is the randomness of the driving Brownian motion;
- their conditional expectation is easy to compute; $\mathbb{E}[\Pi_t | \mathcal{F}_s] = \Pi_s$
- Itô integrals are martingales;
- martingales can be represented as Itô integrals;

$$\left(Y(t) = \int_0^t g(.)\, dX(s) \right) \rightarrow \text{Mart}$$

$$dS^*(t) = \boxed{(\mu - r) S^*(t)\, dt} + \sigma S^*(t)\, dX(t)$$

For $S^*(t)$ to be a martingale, we need it to be driftless, that is, we need $\mu - r = 0 \iff \boxed{\mu = r}$

As a result, to ease our derivation we would like our main stochastic process, the discounted asset price $S^*(t)$, to be a martingale.

However, the discounted asset price $S^*(t)$ is not generally a martingale under the measure $\mathbb{P}$.

Martingales $\mathbb{E}[M(t) \mid \mathcal{F}_0] = M_0$

Sub-martingales $\mathbb{E}[M(t) \mid \mathcal{F}_0] > M_0$

Super-martingale $\mathbb{E}[M(t) \mid \mathcal{F}_0] < M_0$

Indeed, for the discounted asset price $S^*(t)$ to be a martingale under the measure $\mathbb{P}$, we would need $S^*(t)$ to be driftless, i.e.

$$(\mu - r)S^*(t)dt = 0 \Leftrightarrow \mu = r$$

By extension,

- if $\mu > r$ (as should be the case in financial markets) then $S^*(t)$ is a *submartingale*; and
- if $\mu < r$, $S^*(t)$ is a *supermartingale*.


Expected value of £1 invested from time 0 to time t and discounted back at the risk-free rate: the stock is not a martingale!
£1.70
£1.60
£1.50
£1.40
£1.30
£1.20
£1.10
£1.00
£0.90
£0.80
Expected value at time t
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 8.5 9 9.5 10
time t
discounted stock price (mu - r= 5%)
martingale: E[M_t |F_0] = M_0 = £1

Change of Measure → RN theorem

$\mathbb{P} \sim \mathbb{Q}$ $\qquad \mathbb{Q}(\underset{\text{event}}{A}) = \int_A \Lambda \, d\mathbb{P}$

If we want the process $S^*(t)$ to benefit from all the nice properties of martingales, we will need to move away from the measure $\mathbb{P}$ and into a new measure under which $S^*(t)$ is always a martingale.

Our broad objective is therefore to

- find a measure $\mathbb{Q}$ under which S^* is a martingale; and then,
- use the Radon Nikodým theorem to perform the change of measure.

$$\Lambda = \frac{d\mathbb{Q}}{d\mathbb{P}}$$

$$\mathbb{Q}(A) = \int_A \frac{d\mathbb{Q}}{\cancel{d\mathbb{P}}} \cancel{d\mathbb{P}} = \int_A d\mathbb{Q} = \mathbb{Q}(A)$$

The measure $\mathbb{Q}$ we are seeking is called a **martingale measure**.

Definition (Martingale Measure)

A probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ and equivalent to $\mathbb{P}$ is called a martingale measure for $S^*(t)$ if $S^*(t)$ is a martingale under $\mathbb{Q}$.

The trouble we encounter now is that while Radon Nicodým theorem we saw in the Martingale I lecture can be used to help us change measure once we know the measure we want to change into, it cannot help us to identify or define the martingale measure $\mathbb{Q}$ we are looking for.

To identify it, we need to use an additional result called **Girsanov's theorem**.

⇐ sanity checks

Key Fact (Novikov Condition)

A process θ satisfies the **Novikov condition** *if*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \theta_s^2 ds\right)\right] < \infty$$

Key Fact

If a process θ satisfies the Novikov condition, then the process M^{θ} defined as

$$M_t^{\theta} \quad = \quad \exp\left(-\int_0^t \theta_s dX_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right), \quad t \in [0, T]$$

is a martingale.

Key Fact (Girsanov's Theorem)

Given a process θ satisfying the Novikov condition, we can define the probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ equivalent to $\mathbb{P}$ through the Radon Nicodým derivative

Part I

Opportunistic

$$\Lambda = \frac{d\mathbb{Q}}{d\mathbb{P}} = \boxed{\exp\left(-\int_0^t \theta_s dX_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right)}, \quad t \in [0, T]$$

exponential martingale

In this case, the process $X^{\mathbb{Q}}$ defined as

Part II

Plan

$$X_t^{\mathbb{Q}} = X_t + \int_0^t \theta(s) ds, \quad t \in [0, T] \qquad (2)$$

is a standard Brownian Motion on $(\Omega, \mathcal{F}, \mathbb{Q})$.

So, What Does Girsanov Actually Do?

Girsanov effectively extends the Radon Nikodým result by

- giving an expression for the Radon Nikodým derivative;
- expliciting a correspondence between the $\mathbb{P}$ measure and the $\mathbb{Q}$ measure in terms of their respective Brownian motions.

The process θ is vital because it acts as the “key” enabling us to define the measure $\mathbb{Q}$ via the Radon Nikodým derivative and to travel in between the $\mathbb{P}$ measure and the $\mathbb{Q}$ measure via the Brownian motion correspondence.

The difficulty with Girsanov is that the theorem stops short of identifying the process θ. We therefore need to have an idea of what θ is like and make sure that it satisfies the Novikov condition.

So, how can we guess what θ is?

In our case, we want to find θ such that S^* is a $\mathbb{Q}$-martingale. Applying Girsanov for an arbitrary process θ, we see that under $\mathbb{Q}$, the dynamics of S^* is given by

$$\frac{dS_t^*}{S_t^*} = (\mu - r)dt + \sigma\left(-\theta dt + dX_t^{\mathbb{Q}}\right)$$

If we assume that θ is constant[2], then we have

$$\frac{dS_t^*}{S_t^*} = (\mu - r - \sigma\theta)dt + \sigma dX_t^{\mathbb{Q}}$$

under $\mathbb{Q}$.

For S_t^* to be a martingale under $\mathbb{Q}$, its dynamics must be driftless, which implies

$$\mu - r - \sigma\theta = 0$$

i.e.

$$\theta = \frac{\mu - r}{\sigma}$$

[2] After all why not? This is consistent with our assumptions made with respect to the other parameters of the problem.

We observe that θ satisfies the Novikov condition. Invoking **Girsanov's theorem**, we can finally define the equivalent martingale measure $\mathbb{Q}$ via the Radon Nikodým derivative:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\frac{\mu - r}{\sigma} X_t - \frac{1}{2}\frac{(\mu - r)^2}{\sigma^2} t\right), \quad t \in [0, T]$$

Moreover, the $\mathbb{Q}$-Brownian Motion, $X^{\mathbb{Q}}$, is defined as

$$X_t^{\mathbb{Q}} = X_t + \frac{\mu - r}{\sigma} t, \quad t \in [0, T]$$

and under $\mathbb{Q}$ the discounted asset process is indeed a martingale since

$$\frac{dS_t^*}{S_t^*} = \sigma dX_t^{\mathbb{Q}}$$

2.4. Derivative Valuation

$$V_t^*(\phi) = \frac{V_t(\phi)}{B(t)}$$

We will denote by $\chi(t, S_t)$ the time t arbitrage-free value of the derivative we are attempting to price.

By analogy with what we have done earlier with the share price, we will consider the discounted value of the replicating portfolio. We define the time t discounted portfolio value V_t^* by

$$V_t^*(\phi) = \frac{V_t(\phi)}{B_t}, \quad t \in [0, T]$$

To prevent arbitrage, the value of the replicating portfolio must be equal to the value of the derivative:

$$\boxed{\chi(t, S_t)} = V_t, \quad t \in [0, T]$$

Once discounted, this equation becomes

$$\frac{\chi(t, S_t)}{B_t} = V_t^*(\phi), \quad t \in [0, T] \tag{3}$$

In particular, for $t = T$, we have

$$\frac{\chi(T, S_T)}{B_T} = \frac{G(S_T)}{B_T} = V_T^*$$

Taking the conditional expectation under $\mathbb{Q}$, we get

$$\boxed{\mathbb{E}^{\mathbb{Q}}\left[V_T^*(\phi)|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[B_T^{-1}G(S_T)|\mathcal{F}_t\right],} \quad t \in [0, T] \qquad (4)$$

Now, how can we link expressions (3) and (4) in order to find the time t value of the derivative maturing at time T?

Answer: through our self-financing trading strategy!

$dV^{*}(t)$

First note that

$$dB_t^{-1} = -rB_t^{-1}dt$$

By the **Itô Product Rule**,

$$\begin{aligned} dV_t^* &= d(V_t B_t^{-1}) \\ &= V_t dB_t^{-1} + B_t^{-1} dV_t \\ &= \left(\phi_t^S S_t + \phi_t^B B_t\right) dB_t^{-1} + B_t^{-1} \left(\phi_t^S dS_t + \phi_t^B dB_t\right) \\ &= \phi_t^S \left(B_t^{-1} dS_t + S_t dB_t^{-1}\right) \\ &= \phi_t^S dS_t^* \end{aligned}$$

Integrating, we see that

$$\begin{aligned} V_t^* &= V_0^* + \int_0^t \phi_u^S dS_u^* \\ &= V_0^* + \int_0^t \phi_u^S \sigma S_u^* dX_u^{\mathbb{Q}} \end{aligned} \tag{5}$$

because under $\mathbb{Q}$,

$$dS_u^* = \sigma S_u^* dX_u^{\mathbb{Q}}$$

Now, note that

$$\int_0^t \phi_u^S \sigma S_u^* dX_u^{\mathbb{Q}}$$

is an Itô integral and therefore a martingale. Then, the discounted portfolio value V^* is a martingale.

Hence, under $\mathbb{Q}$, not only is S^* a martingale, but so is V^*!

Since V^* is a martingale, then by definition

$$V_t^* = \mathbb{E}^{\mathbb{Q}}\left[V_T^*(\phi)|\mathcal{F}_t\right], \quad t \in [0, T] \tag{6}$$

Considering in addition relationships (3) and (4), we obtain

No arbitrage

martingale under Q

$$\begin{aligned} \frac{\chi(t, S_t)}{B_t} &\overset{\text{No arbitrage}}{=} V_t^*(\phi) \\ &= \mathbb{E}^{\mathbb{Q}}\left[V_T^*(\phi)|\mathcal{F}_t\right] \\ &\underset{\text{No arb}}{=} \mathbb{E}^{\mathbb{Q}}\left[B_T^{-1} G(S_T)|\mathcal{F}_t\right], \quad t \in [0, T] \end{aligned} \tag{7}$$

No arb

$$\chi(t, S(t)) = B(t)\, \mathbb{E}^{\mathbb{Q}}\left[\frac{G(S(T))}{B(T)}\Big|\mathcal{F}_t\right]$$

Equation (7) is the cornerstone of asset valuation[3].

Key Fact

The value at time t of a derivative maturing at time T is the expected value under the $\mathbb{Q}$ measure of the discounted terminal cash flow of the contract.

$$\chi(t, S_t) = B_t \mathbb{E}^{\mathbb{Q}} \left[B_T^{-1} G(S_T) | \mathcal{F}_t \right], \quad t \in [0, T] \qquad (8)$$

$$\chi(t, S_t) = \mathbb{E}^{\mathbb{Q}} \left[\frac{B_t}{B_T} G(S_T) \middle| \mathcal{F}_t \right]$$

[3] and of the application of Monte-Carlo methods.

Aside: Is The Trading Strategy Considered Truly Self-Financing?

We have seen that under $\mathbb{Q}$, V^* is a martingale. Hence, by the **Martingale Representation Theorem** (see the Martingale I Lecture), there exists a process θ satisfying technical condition such that

$$
\begin{aligned}
V_t^* &= V_0^* + \int_0^t \theta_u dX_u^{\mathbb{Q}} \\
&= V_0^* + \int_0^t h_u dS_u^*, \quad t \in [0, T]
\end{aligned}
$$

where $h_t = \frac{\theta_t}{\sigma S_t^*} = \frac{\mu - r}{\sigma^2} \frac{1}{S_t^*}$.

Aside: $\frac{\mu - r}{\sigma^2}$ represents the optimal allocation to the log utility or Kelly criterion portfolio.

Consider a trading strategy ϕ defined as:

$$\begin{aligned}\phi_t^S &= h_t \\ \phi_t^B &= V_t^* - h_t S_t^* = B_t^{-1}(V_t - h_t S_t)\end{aligned}$$

We already know that $V_T(\phi) = G(S_T)$. We will now check that the strategy ϕ is self-financing. To do so, we need to go back to the current, or undiscounted, value of the replicating portfolio.

By the Itô Product Rule:

$$\begin{aligned} dV_t(\phi) &= d(B_t V_t^*) \\ &= B_t dV_t^* + V_t^* dB_t \\ &= B_t h_t dS_t^* + rV_t dt \\ &= B_t h_t \left(B_t^{-1} dS_t - rB_t^{-1} S_t dt\right) + rV_t dt \\ &= h_t dS_t + r(V_t - h_t S_t)dt \end{aligned}$$

which confirms the fact that the portfolio is indeed self financing.

Note that we do not know the specifics of the trading strategy, namely how much of the underlying asset to hold. All we know is that a strategy exists and that it is self-financing.

After seeing the importance of the Delta-hedging strategy in both the PDE approach and the Binomial model, it may seem quite strange to be dealing with an approach that do not require a specific knowledge of the strategy.

But this is precisely what the probabilistic approach does. It only checks that some technical conditions are fulfilled in order to guarantee that there exists an “appropriate” replicating strategy, without actually defining it.

End of aside.

The Black-Scholes Call Option Problem

We will now use the valuation formula (8) to solve the Black-Scholes European call option problem. A similar derivation could be done for European put options and for binary options.

Since r is constant, we can simplify (8) by writing

$$\chi(t, S_t) \;\; = \;\; e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[F(S_T)|\mathcal{F}_t\right], \quad t \in [0, T]$$

Going back to our *undiscounted* asset price process, S, we observe that under the measure $\mathbb{Q}$

$$\frac{dS_t}{S_t} = rdt + \sigma dX_t^{\mathbb{Q}}$$

Seems familiar? It is indeed what we saw during the previous lecture on Black-Scholes through PDE as the "Risk-Neutral" GBM.

In particular, at time T,

$$S_T = S_t \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\left(X_T^{\mathbb{Q}} - X_t^{\mathbb{Q}}\right)\right\}$$

3.1. Direct Derivation

Define a new random variable $Y_T := \ln \frac{S_T}{S_t}, \quad \forall t \in [0, T]$. Since S_t is Lognormally distributed, then the log return of the asset over the period $[0, T]$, Y_T, is normally distributed with mean

$$\left(r - \frac{1}{2}\sigma^2\right)(T - t)$$

and variance

$$\sigma^2(T - t)$$

i.e.

$$Y_T \sim \mathcal{N}\left(\left(r - \frac{1}{2}\sigma^2\right)(T - t), \sigma^2(T - t)\right)$$

Using the basic definition of expectations, the expectation in our pricing formula (8) can be rewritten in terms of the random variable Y_T as

$$\chi(t, S_t) \;=\; e^{-r(T-t)} \int_{-\infty}^{\infty} G(S_t e^{y}) p(y) dy$$

where p is the PDF of Y.

To simplify our notation, we define

$$\begin{aligned}\tilde{r} &= r - \frac{1}{2}\sigma^2 \\ \tau &= T - t\end{aligned}$$

With this new notation, we have

$$Y_T \sim \mathcal{N}\left(\tilde{r}\tau, \sigma^2\tau\right)$$

To further simplify our calculations, define the *standardized Normal random variable*[4] Z

$$Z := \frac{Y - \tilde{r}\tau}{\sigma\sqrt{\tau}} \Leftrightarrow Y = \tilde{r}\tau + Z\sigma\sqrt{\tau}$$

After a change of variable, the expectation (9) can be expressed in terms of Z as

$$\chi(t, S_t) \quad = \quad e^{-r\tau} \int_{-\infty}^{\infty} G\left(S_t e^{\tilde{r}\tau + \sigma\sqrt{\tau}z}\right) \varphi(z) dz$$

where φ is the standard normal PDF:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

[4] A standard or standardized Normal random variable is a Normally distributed random variable with mean 0 and variance 1

The payoff function for a call is given by

$$G(S_T) = \max\left[S_T - E, 0\right]$$

Substituting into the pricing equation (9), we get

$$\chi(t, S_t) \;\; = \;\; e^{-r\tau} \int_{-\infty}^{\infty} \max\left[S_t e^{\tilde{r}\tau + \sigma\sqrt{\tau} z} - E, 0\right] \varphi(z) dz$$

We can get rid of the max function by noticing that the integral vanishes when

$$S_t e^{\tilde{r}\tau + \sigma\sqrt{\tau}z} < E$$

i.e. when

$$z < z_0 := \frac{\ln\left(\frac{E}{S_t}\right) - \tilde{r}\tau}{\sigma\sqrt{\tau}}$$

The pricing formula now becomes

$$\begin{aligned} \chi(t, S_t) &= e^{-r\tau} \int_{z_0}^{\infty} \left(S_t e^{\tilde{r}\tau + \sigma\sqrt{\tau}z} - E \right) \varphi(z) dz \\ &= e^{-r\tau} \int_{z_0}^{\infty} S_t e^{\tilde{r}\tau + \sigma\sqrt{\tau}z} \varphi(z) dz - e^{-r\tau} \int_{z_0}^{\infty} E\varphi(z) dz \end{aligned} \quad (9)$$

The second term on the right-hand side of (9) seems easier to evaluate, so let's start with this one

$$\begin{aligned}-e^{-r\tau}\int_{z_0}^{\infty} E\varphi(z)dz &= -Ee^{-r\tau}\int_{z_0}^{\infty}\varphi(z)dz \\ &= -Ee^{-r\tau}P[Z \geq z_0]\end{aligned}$$

By symmetry of the normal distribution, this can also be written as

$$-Ee^{-r\tau}P[Z \leq -z_0] = -Ee^{-r\tau}N(-z_0)$$

where N is the standard normal CDF.

To evaluate the first term on the right side of (9), we need to complete the square in the exponent:

$$\begin{aligned}
e^{-r\tau}\int_{z_0}^{\infty} S_t e^{\tilde{r}\tau+\sigma\sqrt{\tau}z}\varphi(z)dz &= \frac{e^{(\tilde{r}-r)\tau}S_t}{\sqrt{2\pi}}\int_{z_0}^{\infty} e^{\sigma\sqrt{\tau}z-\frac{1}{2}z^2}dz \\
&= \frac{e^{(\tilde{r}-r)\tau}S_t}{\sqrt{2\pi}}\int_{z_0}^{\infty} e^{-\frac{1}{2}(z-\sigma\sqrt{\tau})^2+\frac{1}{2}\sigma^2\tau}dz \\
&= \frac{S_t}{\sqrt{2\pi}}\int_{z_0}^{\infty} e^{-\frac{1}{2}(z-\sigma\sqrt{\tau})^2}dz \\
&= S_t P[U \geq z_0-\sigma\sqrt{\tau}] \\
&= S_t P[U \leq -z_0+\sigma\sqrt{\tau}]
\end{aligned}$$

where $U = Z-\sigma\sqrt{\tau}$ so that $U \sim \mathcal{N}(-\sigma\sqrt{\tau},1)$. Standardizing, we see that the first term is actually equal to

$$S_t N(-z_0+\sigma\sqrt{\tau})$$

In order to write this in the more familiar form, all we need to do is to define d_1 and d_2 as

$$
\begin{aligned}
d_1 &= -z_0 + \sigma\sqrt{\tau} = \frac{\ln\left(\frac{S_t}{E}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \\
d_2 &= -z_0 = \frac{\ln\left(\frac{S_t}{E}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}
\end{aligned}
$$

and to substitute the first and the second term into the pricing equation:

$$\chi(t, S_t) = S_t N(d_1) - E e^{-r(T-t)} N(d_2)$$

3.2. Alternative Derivation Through Change of Measure

In the case of a call, the valuation equation (8) can be expressed as

$S_T = S_t \exp\{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma(X_T^{\mathbb{Q}} - X_t)\}$

$$
\begin{aligned}
X(t, S(t)) = \quad & B_t \mathbb{E}^{\mathbb{Q}} \left[B_T^{-1} [S_T - K]^+ \,|\mathcal{F}_t \right] \\
= \quad & B_t \mathbb{E}^{\mathbb{Q}} \left[B_T^{-1} \left[S_t e^{\sigma(X_T^{\mathbb{Q}} - X_t) + (r - \frac{1}{2}\sigma^2)(T-t)} - K \right]^+ |\mathcal{F}_t \right]
\end{aligned}
$$

where the strike price is now denoted by K to avoid any confusion with $\mathbb{E}$, the expectation.

Focusing on the case $t = 0$, we can drop the conditional expectation and deal with an unconditional expectation:

$$
\begin{aligned}
& \mathbb{E}^{\mathbb{Q}} \left[B_T^{-1} [S_T - K]^+ \right] \\
= \quad & \mathbb{E}^{\mathbb{Q}} \left[B_T^{-1} S_T \mathbf{1}_{\{S_T > K\}} \right] - \mathbb{E}^{\mathbb{Q}} \left[B_T^{-1} K \mathbf{1}_{\{S_T > K\}} \right]
\end{aligned}
$$

Tackling the second expectation on the RHS,

$$\begin{aligned}
& \mathbb{E}^{\mathbb{Q}}\left[B_T^{-1} K \mathbf{1}_{\{S_T > K\}}\right] \\
= \;& e^{-rT} K P^{\mathbb{Q}}\left[S_T > K\right] \\
= \;& e^{-rT} K P^{\mathbb{Q}}\left[S_0 e^{\sigma X_T^{\mathbb{Q}} + (r - \frac{1}{2}\sigma^2)T} > K\right] \\
= \;& e^{-rT} K P^{\mathbb{Q}}\left[\ln\left(\frac{S_0}{K}\right) + (r - \frac{1}{2}\sigma^2)T > -\sigma X_T^{\mathbb{Q}}\right] \\
= \;& e^{-rT} K P^{\mathbb{Q}}\left[\frac{\ln\left(\frac{S_0}{K}\right) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} > \xi\right] \\
= \;& e^{-rT} K N(d_2)
\end{aligned}$$

where we have emphasized the fact that the probability P is taken with respect to the measure $\mathbb{Q}$ and have defined $\xi = -X_T^{\mathbb{Q}}/\sqrt{T}$, which is standard Normal random variable: $\xi \sim \mathcal{N}(0, 1)$.

As for the first expectation on the RHS,

$$
\begin{aligned}
& \mathbb{E}^{\mathbb{Q}}\left[B_T^{-1} S_T \mathbf{1}_{\{S_T>K\}}\right] \\
= \; & \mathbb{E}^{\mathbb{Q}}\left[S_T^{*} \mathbf{1}_{\{S_T>K\}}\right] \\
= \; & \mathbb{E}^{\mathbb{Q}}\left[S_0 \exp\left\{\sigma X_T^{\mathbb{Q}} - \frac{1}{2}\sigma^2 T\right\} \mathbf{1}_{\{S_T>K\}}\right]
\end{aligned}
$$

Unless we can find a trick, it does not seem that this expectation can be computed analytically...

But, as happens frequently in mathematics, we can find a simple trick!

Aside: The Stochastic Exponential

Sometimes, in the literature, you will encounter the notation

$$\mathcal{E}\left(\int_0^t \theta_s dX_s\right)$$

This "curly E" denotes the **stochastic** (or **Doléans**) **exponential**, which is defined as

$$\mathcal{E}\left(\int_0^t \theta_s dX_s\right) = \exp\left(\int_0^t \theta_s dX_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right)$$

The Girsanov Theorem can also be defined in terms of the stochastic exponential.

Key Fact (Girsanov's Theorem in terms of stochastic exponential)

Given a process θ satisfying the condition

$$\mathbb{E}\left[\mathcal{E}\left(\int_0^T \theta_s dX_s\right)\right] = 1$$

we can define the probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ equivalent to $\mathbb{P}$ through the Radon Nicodým derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}\left(\int_0^t \theta_s dX_s\right), \quad t \in [0, T]$$

$$= \exp\left\{+\int_0^t \theta_s dX(s) - \frac{1}{2}\int_0^t \theta^2(s) ds\right\}$$

In this case, the process $X^{\mathbb{Q}}$ defined as

$$X_t^{\mathbb{Q}} = X_t - \int_0^t \theta(s) ds, \quad t \in [0, T] \qquad (10)$$

is a standard Brownian Motion on $(\Omega, \mathcal{F}, \mathbb{Q})$.

Observe that condition (10) requires that the stochastic exponential is a martingale. It is therefore equivalent to the Novikov condition in our standard presentation of Girsanov.

Also, note the sign change between (2) and (10), which is due to the difference in sign of the stochastic integral in between the Novikov condition and the stochastic exponential.

End of aside and back to our problem...

Notice that the term

$$\exp\left\{\sigma X_T^{\mathbb{Q}} - \frac{1}{2}\sigma^2 T\right\} \tag{11}$$

inside the expectation looks reminiscent of the Doléans exponential we introduced in our second formulation of Girsanov's theorem:

$$\mathcal{E}\left(\int_0^t \theta_s dX_s\right) = \exp\left(\int_0^t \theta_s dX_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right)$$

In fact, we can reformulate (11) as

$$\exp\left\{\int_0^T \sigma dX_t^{\mathbb{Q}} - \frac{1}{2}\int_0^T \sigma^2 dt\right\}$$

and check that this is indeed the the Doléans exponential with $\theta = \sigma$!

This is an important observation, because it means that we can get rid of this bothersome term in our expectation by defining a new measure via Girsanov's theorem.

All we do do is check that

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left\{\sigma X_T^{\mathbb{Q}} - \frac{1}{2}\sigma^2 T\right\}\right] = 1$$

i.e. that (11) is an exponential martingale.

(check left to the reader since we have done something very similar at the end of Lecture 2.5)

We now define a new probability measure $\bar{\mathbb{Q}}$ via the Radon Nikodym derivative

$$\bar{\Lambda} = \frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} = \exp\left\{\sigma X_T^{\mathbb{Q}} - \frac{1}{2}\sigma^2 T\right\}$$

Note that under the $\bar{\mathbb{Q}}$ measure,

$$X_t^{\bar{\mathbb{Q}}} = X_t^{\mathbb{Q}} - \sigma t, \qquad t \in [0, T]$$

is a Brownian motion and

$$S_T^* = S_0 \exp\left\{\sigma X_T^{\bar{\mathbb{Q}}} + \frac{1}{2}\sigma^2 T\right\} \tag{12}$$

Therefore,

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[S_T^* \mathbf{1}_{\{S_T>K\}}\right] &= \mathbb{E}^{\mathbb{Q}}\left[S_0 \exp\left\{\sigma X_T^{\mathbb{Q}} - \frac{1}{2}\sigma^2 T\right\} \mathbf{1}_{\{S_T>K\}}\right] \\
&= S_0 \int \frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} \mathbf{1}_{\{S_T>K\}} d\mathbb{Q} \\
&= S_0 P^{\bar{\mathbb{Q}}}\left[S_T > K\right] \\
&= S_0 P^{\bar{\mathbb{Q}}}\left[S_T^* > K B_T^{-1}\right] \\
&= S_0 P^{\bar{\mathbb{Q}}}\left[S_0 \exp\left\{\sigma X_T^{\bar{\mathbb{Q}}} + \frac{1}{2}\sigma^2 T\right\} > K e^{-rT}\right] \\
&= S_0 P^{\bar{\mathbb{Q}}}\left[\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right) T > -\sigma X_T^{\bar{\mathbb{Q}}}\right] \\
&= S_0 P^{\bar{\mathbb{Q}}}\left[\frac{\ln\left(\frac{S_0}{K}\right) + (r + \frac{1}{2}\sigma^2) T}{\sigma\sqrt{T}} > \xi\right] \\
&= S_0 N(d_1)
\end{aligned}
$$

where we have emphasized the fact that the probability P is taken with respect to the measure $\bar{\mathbb{Q}}$ and have defined $\xi = -X_T^{\mathbb{Q}}/\sqrt{T}$.

3.3. A few concluding notes on this section...

This derivation provides a wealth of probability related information on the derivative and on the relationship between the underlying price and the exercise price:

$S(t) \sim Ln$

- $N(d_1)$ and $N(d_2)$ are probabilities based on a normal distribution;
- the normal distribution itself is a consequence of our assumption that the price of the underlying asset is lognormally distributed and hence that its log returns are normally distributed;
- the **probability under $\mathbb{Q}$ of exercising the option at maturity** is explicitly given: it is N(d2)!
- the **probability under $\bar{\mathbb{Q}}$ of exercising the option at maturity** is explicitly given: it is N(d1)!

While this derivation is quite logical overall, the most counter-intuitive aspect is certainly the fact that you do not need to know the specifics of the replicating strategy. One of the direct implications of this fact is that, since delta-hedging has not been considered and the Greeks have not been defined, the probabilistic approach presented here cannot help you with **local** risk management.

4. The Numéraire Pair

The **Fundamental Asset Pricing Formula** is much more general than what we have seen so far and there are some excellent reasons to call it "fundamental"...

We can extend the Fundamental Asset Pricing Formula very naturally to take advantage of the idea of a **numéraire pair** $(N_t, \mathbb{Q}^N)$.

$(B(t), \mathbb{Q})$

Key Fact

The numéraire pair $(N_t, \mathbb{Q}^N)$ is comprised of:

Price of a security

- *a numéraire process N_t: any stochastic process $N_t > 0$ that can be viewed as a price can be used as a numéraire;*
- *an equivalent martingale measure $\mathbb{Q}^N$ under which any asset price discounted using the numéraire is a martingale.*

... and here is where the numéraire pair comes in handy:

Key Fact (Fundamental Asset Pricing Formula (Revisited))

The value at time t of a derivative maturing at time T is the expected value under the $\mathbb{Q}^N$ measure of the terminal cash flow of the contract, discounted using the numéraire asset.

$$\chi(t, S_t) = N_t \mathbb{E}^{\mathbb{Q}^N}\left[N_T^{-1} G(S_T) | \mathcal{F}_t\right], \quad t \in [0, T]$$

In fact we have already seen two numéraire pairs:

- $(B_t, \mathbb{Q})$: the risk-free asset and the risk-neutral measure; $\to N(d_2)$
- $(S_t, \bar{\mathbb{Q}})$: the stock and the auxiliary measure $\bar{\mathbb{Q}}$ we used to compute $N(d_1)$. $\to N(d_1)$

We can also price a derivative under the **real-world $\mathbb{P}$ measure**. In this case, the numéraire asset is the **log-optimal (or Kelly) portfolio**:

- Portfolio maximizing the log return on (or log utility of) wealth;
- In the Black-Scholes universe, the proportion of the log-optimal portfolio invested in the stock is $\frac{\mu - r}{\sigma^2}$;
- More on this portfolio in ~~Module 6~~!
  Extra lectures

5. The Feynman-Kač Formula

During our investigation of the martingale methods, we derived the fundamental asset pricing equation (8).

Since in the Black-Scholes problem we have assumed that the interest rate is constant, we can rewrite this formula as:

$$\chi(t, S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[G(S_T)|\mathcal{F}_t\right], \quad t \in [0, T] \tag{13}$$

It turns out that this type of expectation has a PDE representation, thanks to the **Feynman-Kač formula**.

Key Fact (The Feynman-Kač Formula)

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

$$\begin{aligned} \frac{\partial V}{\partial t}(t,s) + \mu(t,s)\frac{\partial V}{\partial s}(t,s) + \frac{1}{2}\sigma^2(t,s)\frac{\partial^2 V}{\partial s^2}(t,s) - rV(t,s) &= 0 \\ V(T,s) = G(s) \qquad (14) \end{aligned}$$

and that the process $S(t)$ 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

$$V(t,S_t) = e^{-r(T-t)}\mathbb{E}\left[G(S_T)|\mathcal{F}_t\right] \qquad (15)$$

Application

In the Black-Scholes model, the option value under the risk-neutral measure can be expressed as the expectation:

$$V(t, S_t) \quad = \quad e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[G(S_T)|\mathcal{F}_t\right], \quad t \in [0, T]$$

where S_t follows the dynamics:

$$dS_t = rS_t dt + \sigma S_t dX^{\mathbb{Q}}(t) \tag{16}$$

in which $X^{\mathbb{Q}}(t)$ is a Brownian motion under $\mathbb{Q}$.

By the **Feynman-Kač formula**, the value $V(t, S_t)$ of the option solves the boundary value problem

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV \quad = \quad 0$$
$$V(T, s) = G(s)$$

which is the **Black-Scholes PDE**.

A Few Remarks...

1. **the Feynman-Kač formula works both ways:** we can represent a PDE of the form (14) as an expectation, and we can represent an expectation of the form (15). However, since historically Feynman-Kač was established to represent PDEs as expectation, we generally quote the formula as the representation of a PDE;
2. **the Feynman-Kač formula is independent from the measure:** Feynman-Kač works for any measure and does not imply any change of measure. Indeed, in the previous slide, we have used Feynman-Kač in the measure of the expectation, i.e the risk-neutral measure. **Tip:** make sure that you are using the "correct" dynamics for the process $S(t)$ (i.e. the dynamics under the same measure as the expectation), otherwise the $\frac{\partial V}{\partial s}$ coefficient in the PDE will be wrong!

A Few Remarks...

3. since we have not gone through the Δ-hedging argument of the PDE method, we do not know that $\frac{\partial V}{\partial s}$ represents the number of stocks to be held to hedge/replicate an option.

6. Extensions of the Basic Framework

In this section we consider three extensions to the basic Black-Scholes framework

1. Dividend paying stocks;
2. Stock price process with time-dependent parameters;
3. Valuations of options on futures (Black's formula);

5.1. Constant Dividends

If the underlying asset pays a constant dividend yield D, then its evolution becomes

$$dS_t = (\mu - D)S_t dt + \sigma S_t dX_t, \quad S_0 > 0$$

In this case, self-financing condition needs to be adjusted from

$$V_t(\phi) = V_0(\phi) + \int_0^t \phi_u^S dS_u + \int_0^t \phi_u^B dB_u, \quad \forall t \in [0, T]$$

to

$$\begin{aligned} V_t(\phi) &= V_0(\phi) + \int_0^t \phi_u^S dS_u + \int_0^t D\phi_u^S S_u d_u + \int_0^t \phi_u^B dB_u, \\ & \forall t \in [0, T] \end{aligned}$$

in order to account for the dividend.

In this case, it is generally best to adjust the asset for the dividend paid. Define $\tilde{S}$ as dividend-adjusted asset dynamics, given by

$$\tilde{S}_t = S_t e^{Dt}$$

$\tilde{S}$ is the value of an investment in the stock if all dividends paid are automatically reinvested.

We can also define $\mu_D = \mu + D$ in order to write the evolution of $\tilde{S}$ as

$$d\tilde{S}_t = \mu_D \tilde{S}_t dt + \sigma \tilde{S}_t dX_t$$

The derivation we did for S now fully applies to $\tilde{S}$ and we quickly conclude that, since $\tilde{S}_t = e^{Dt} S_t$, the value of a European call on a dividend-paying asset is given by:

$$\begin{aligned}\chi(t, S_t) &= \tilde{S}_t N(d_1) - E e^{-r(T-t)} N(d_2) \\ &= S_t e^{-D(T-t)} N(d_1) - E e^{-r(T-t)} N(d_2)\end{aligned}$$

with d_1 and d_2 given by

$$\begin{aligned}d_1 &= \frac{\ln\left(\frac{S_t}{E}\right) + \left(r - D + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= \frac{\ln\left(\frac{S_t}{E}\right) + \left(r - D - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\end{aligned}$$

5.2. Time-Dependent Parameters

The Black-Scholes formula is not materially affected if instead of evolving according to a geometric Brownian motion, the underlying asset evolves according to a more general, time-dependent dynamics

$$dS_t = \mu(t, S_t)dt + \sigma(t)S_t dX_t, \quad S_0 > 0$$

and the risk-free rate is itself time-dependent, so that

$$dB_t = r(t)B(t)dt, \quad B_0 = 1$$

The price of the risk-free asset is now given by

$$B_t = e^{\int_0^t r(u)du}$$

However, the martingale measure $\mathbb{Q}$ is still unique, with the process θ used in the change of measure defined as

$$\theta = \frac{\frac{\mu(t,S_t)}{S_t} - r(t)}{\sigma(t)}$$

Note that, as expected, under the martingale measure $\mathbb{Q}$, the dynamics of S_t is given by

$$dS_t = r(t)S_t dt + \sigma(t)S_t dX_t^{\mathbb{Q}}, \quad S_0 > 0$$

In addition, the no-arbitrage pricing equation (7) is still valid, and as a consequence, the value of a derivative is given by:

$$\chi_t = e^{-\int_t^T r(u)du}\mathbb{E}\left[G(S_T)|\mathcal{F}_t\right]$$

Hence, the Black-Scholes formula that we obtained in the constant parameter case is also valid provided that we make two substitutions:

- replace $r(T-t)$ with $\int_t^T r(u)du$;
- replace σ^2 with $v^2 = \int_t^T \sigma^2(u)du$

Note that since the drift of the underlying asset $\mu(t, S_t)$ does not have any impact on the solution of the problem, we could actually choose any functional form we could think of to represent it.

5.3. Black's Model for Options on Futures

In 1976, Black published an option model in which the underlying asset is not traded spot (i.e. bought and sold for value today) but is a futures contract (i.e. a derivative bought and sold today for value at a later date).

What is particularly interesting for us is that by comparing the valuation of options on spot instruments and on futures or forwards, we can see the important role played by the Time Value of Money in holding together instruments and market across time.

This in turn, is an added motivation for looking at better ways to model interest rates instead of keeping them constant or simply time-dependent.

5.3.1. Futures Forwards and Forward Price

Forward contracts are OTC derivatives securities in which the long party has the obligation to buy an agreed upon quantity of an underlying asset (securities, commodities or others) at an agreed upon time and at an agreed upon price called the forward price. Forward contracts are symmetrical contracts. Therefore, the obligations of the short party mirror those of the long party. The contract is settled at maturity and typically no cash flow is exchanged in the meantime. As they are OTC derivatives, forward contracts are subject to counterparty risk.

Futures contracts are an exchange-traded type of forwards. Since they are traded on an exchange, futures are heavily standardized. Counterparty risk is mitigated by the exchange's clearinghouse. In particular the clearinghouse requires that exchange participants posts margin and to settle the their contracts daily. The combination of daily settlement and margins generates a stream of daily cash-flow through the life of the contract.

Forward contracts can be priced easily through the no-arbitrage condition. The time t forward price for a contract maturing at time T on an underlying S is given by:

$$F(t; T) = S_t e^{r(T-t)}$$

where r is the (constant) risk-free rate.

As a result of the cash flows generated by margin and settlement, futures and forward prices only coincides when interest rates are modelled as constant or time dependent, but not when interest rates are stochastic. However, although the forward and futures price do not often coincide, the forward price still plays a role, at least as an approximation, since it enables us to bridge the time gap between time t and maturity T.

For more information on pricing forwards and futures, please refer to Paul Wilmott's book.

5.3.2. Pricing Options on Futures

We will keep the same setting and assumptions as for the Black-Scholes model, including the hypothesis that the asset's dynamics is given by a geometric Brownian motion.

The dynamics of the time t futures price, f_t, evolves according to a Geometric Brownian Motion:

$$df_t = \mu_f f_t dt + \sigma_f f_t dX, \quad f(0) = f_0$$

where μ_f and σ_f represent respectively the drift and diffusion of the futures.

Since interest rates are deterministic, the futures price is equal to the forward price and we can write:

$$f_t = F(t; T) = S_t e^{r(T-t)} \tag{17}$$

and in particular

$$f_0 = S_0 e^{rT}$$

Recall that

$$dS_t = \mu S_t dt + \sigma S_t dX, \quad S(0) = S_0$$

Applying Itô to the relationship (17), we can now express the dynamics of f_t as

$$df_t = (\mu - r) f_t dt + \sigma f_t dX, \quad f(0) = S_0 e^{rT}$$

and thus we can see now that

$$\begin{aligned} \mu_f &= \mu - r \\ \sigma_f &= \sigma \end{aligned}$$

namely, the volatility of the futures is equal to the volatility of the spot and the drift of the futures is the discounted drift of the spot.

As a result we can also see clearly that the dynamics (SDE) for f_t is of the same form as the dynamics (SDE) for $\frac{S_t}{B_t} = S_t^*$. In fact, we even have

$$f(t) = \frac{S(t)}{B(t)} e^{rT} = S^*(t) e^{rT}$$

where T is fixed by the contract. This relationship reveals that the futures price is already a (naturally) discounted process. The immediate conclusion from this observation is that we will not need to consider the discounted futures price process. We can just go ahead with the futures price process as it stands.

We will now proceed as in the Black-Scholes model, first defining a self-financing futures strategy with equation

$$V_t(\phi) = V_0(\phi) + \int_0^t \phi_u^f df_u + \int_0^t \phi_u^B dB_u, \quad \forall t \in [0, T]$$

Please note that now, the trading strategy involves the futures and the risk-free asset. It does not involve the underlying asset S.

The martingale measure $\mathbb{Q}$ is still unique, with the process θ used in the change of measure defined as

$$\theta = \frac{\mu_f}{\sigma}$$

As expected, under the martingale measure $\mathbb{Q}$, the dynamics of f_t is given by

$$df_t \quad = \quad \sigma f_t dX_t^{\mathbb{Q}}, \quad S_0 > 0$$

In addition, the no-arbitrage pricing equation (7) is still valid, and as a consequence, the value of a derivative is given by:

$$\chi_t = e^{-r(T-t)}\mathbb{E}\left[G(S_T)|\mathcal{F}_t\right]$$

Solving this equation leads us to Black's formula for a European call on a futures:

$$\chi(t, f_t) = e^{-r(T-t)}\left[f_t N(d_1) - EN(d_2)\right]$$

with d_1 and d_2 given by

$$d_1 = \frac{\ln\left(\frac{f_t}{E}\right) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\ln\left(\frac{f_t}{E}\right) - \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}$$

In this lecture, we have seen...

- the probabilistic approach to solving the Black-Scholes problem;
- Girsanov's theorem and how to use it to change measure;
- what an equivalent martingale measure is;
- the derivation of the fundamental asset pricing formula;
- how to get the Black-Scholes fromula from the fundamental asset pricing formula;
- how to use the Feynman-Kǎc formula to go from the fundamental asset pricing formula to a the Black-Scholes PDE;
- extensions to the original Black-Scholes pricing problem: dividends; time-dependent coefficient and options of futures.

CQF Module 3: Option Valuation Models

Connecting the Dots

CQF

In this lecture...

We conclude on the first half of the course by bringing together the three valuation models we have seen so far:

- the binomial model;
- the PDE approach;
- the martingale approach.

1. A Review of the Binomial Model

The binomial model is

- neither a great model of asset behaviour,
- ... nor a great numerical method,
- ... but is one of the best ways of understanding how the various asset pricing techniques we have seen so far interact;

1.1. Recall the general 1-period setup:

- stock priced at S_0;
- time period δt;
- stock can only go up to $S_u = uS_0$ with probability p or down to $S_d = dS_0$ with probability $1 - p$. *Formally, we denote by S_T the random variable taking value S_u in an up-move and value S_d in a down-move*;
- condition that $d < 1 < u$;
- discount rate r. Hence discount factor $D = (1 + r\delta t)^{-1}$ and bank account factor $B = \frac{1}{D} = (1 + r\delta t)$;
- Our objective is to find the current value of an option, denoted by V_0.

There are two ways of finding V_0:

- Delta hedging;
- Risk-neutral valuation.

Obviously, both views must yield the same answer, so it is often useful to use one to check the other (especially during exams).

1.2. Delta Hedging

Delta hedging means finding Δ such that the value of a portfolio long a call option and short Δ stocks is independent on the stock value or risk. We can think of delta has the sensitivity of the option to the change in the stock (i.e. same Delta as in Black Scholes).

Since the portfolio is **delta hedged**, the value of the portfolio must be the same in the upper and lower state. Hence,

$$V_u - \Delta S_u = V_d - \Delta S_d \iff \Delta = \frac{V_u - V_d}{(u-d)S_0}$$

With this choice of Δ, the portfolio is deterministic, i.e. **risk-free**. We can use **no-arbitrage valuation** and state that to prevent arbitrage, the return of the portfolio over the time period δt must be the risk-free rate:

$$V_u - \Delta S_u = V_d - \Delta S_d = \frac{1}{D}(V_0 - \Delta S_0)$$

Therefore,

$$
\begin{aligned}
V_0 &= \frac{V_u - V_d}{u-d} + D\frac{uV_d - dV_u}{u-d} \\
&= \frac{(1-Dd)V_u - (Du-1)V_d}{u-d}
\end{aligned} \tag{1}
$$

1.3. Risk-Neutral Valuation

The **risk-neutral world** has the following characteristics:

- We don't care about risk and don't expect any extra return for taking unnecessary risk;
- We don't ever need statistics to estimate probabilities of events happening;
- We believe that everything is priced using simple expectations.

The **risk-neutral probabilities** are computed as follows. Since in the risk-neutral world investors are not compensated for the risk they are taking, then

$$S_0 = D\left[p^* S_u + (1-p^*)S_d\right] \iff p^* = \frac{S_0/D - S_d}{S_u - S_d} = \frac{1/D - d}{u - d}$$

Hence,

$$V_0 = D\left[p^* V_u + (1 - p^*) V_d\right]$$

1.4. Tying in Delta Hedging and Risk-Neutral Valuation

Equation (1) can be written more elegantly as

$$V_0 = D\left[p^* V_u + (1 - p^*) V_d\right]$$

where

$$p^* = \frac{\frac{1}{D} - d}{u - d}$$

The parameter p^* can be interpreted as a probability. In fact, the Cox Ross Rubinstein model is built so that the p^* is the probability in use.

2. Martingale Properties of the Binomial Model

In the binomial model, we defined a new parameter, p^* and identified it as the "risk-neutral probability," while in the martingale method, we performed a change of measure to formally define an equivalent martingale measure $\mathbb{Q}$.

What is the link here?

2.1. Probability Measures in the Binomial Model

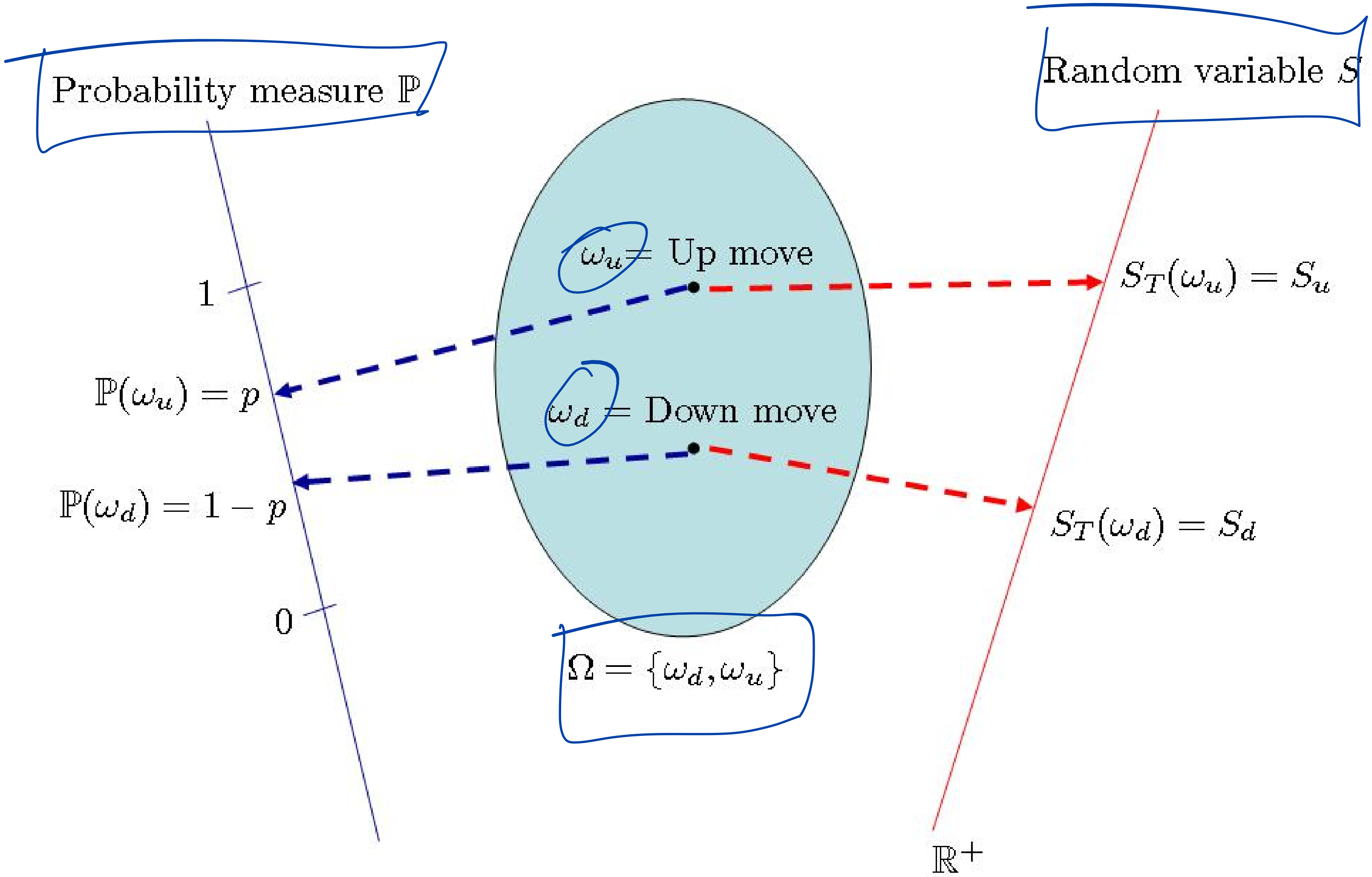
As usual with the binomial model, concepts and their applications are very simple.

Since we just have two possible events[1], namely an up move and a down move, the parameter $p \in (0, 1)$ is enough to uniquely define a probability measure $\mathbb{P}$ as

$$\begin{cases} \mathbb{P}[\text{up move}] & = & p \\ \mathbb{P}[\text{down move}] & = & 1 - p \end{cases}$$

[1]Mathematically, $\boxed{\Omega} = \{\text{up move}, \text{down move}\}$

Figure : Probabilistic setting of the binomial model



We have the physical measure $\mathbb{P}$. What about the risk-neutral measure $\mathbb{P}^*$?

Before answering this question, let's try the same approach as in Lecture 3.3, i.e. determine the equivalent martingale measure $\mathbb{Q}$ and price the option in this measure.

2.2. The Equivalent Martingale Measure in the Binomial Model

Recall from Lecture 3.3. that the martingale measure $\mathbb{Q}$ is defined as the equivalent probability measure such that the stock price, when discounted at the risk-free rate, follows a martingale.

Since we are in such a simple setting, we will not need to use any "big" result.

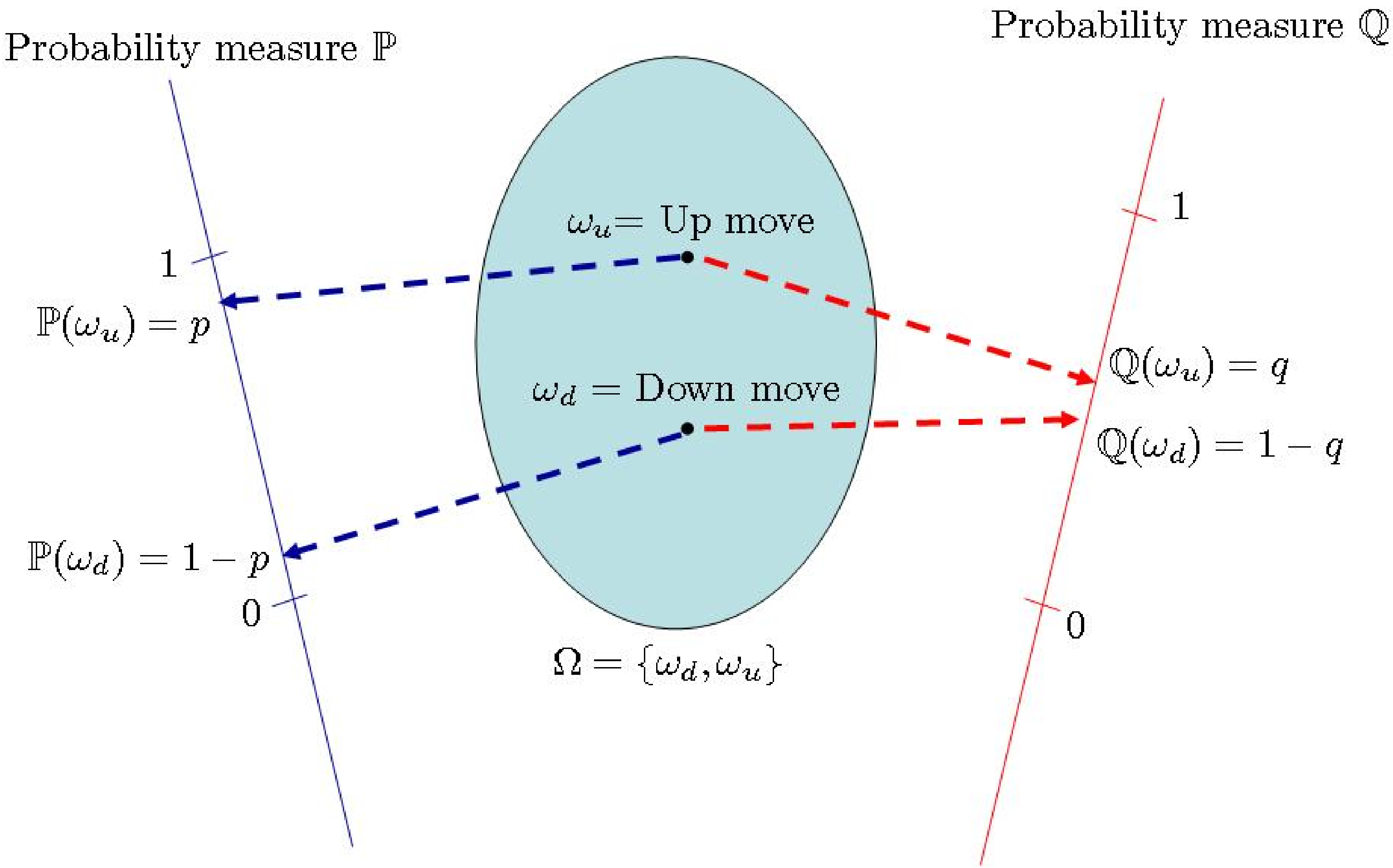
As in the case of the physical measure $\mathbb{P}$, all it takes to define an equivalent measure $\mathbb{Q}$ is a parameter $q \in (0,1)$ such that

$$\begin{cases} \mathbb{Q}\,[\text{up move}] & = \quad q \\ \mathbb{Q}\,[\text{down move}] & = \quad 1-q \end{cases}$$

We can see that since $p \in (0,1)$ and $q \in (0,1)$ the $\mathbb{P}$-measure and $\mathbb{Q}$-measure are indeed equivalent.

$\to$ Find ~~Q~~ q such DS_T is a martingale

Figure : $\mathbb{P}$ measures and $\mathbb{Q}$ measure in the binomial model



The next step is to find the parameter q such that $\mathbb{Q}$ is the equivalent martingale measure.

Under $\mathbb{Q}$, the discounted stock price is a martingale and hence

$$\begin{aligned} S_0 &= \mathbf{E}^{\mathbb{Q}}\left[DS_T\right] \\ &= D\left(qS_u + (1-q)S_d\right) \end{aligned} \tag{2}$$

FAPF for the stock (under the Bernoulli distribution)

Solving this equation for q, we find that

$$q = \frac{1/D - d}{u - d} = p^* \tag{3}$$

This derivation leads us to three observations:

1. The equivalent martingale probability q and the risk neutral probability p^* are equal.
2. q is *unique*: there is only one equivalent martingale measure.
3. the risk-neutral valuation formula (2) is exactly the "fundamental" asset pricing formula we derived in Lecture 3.3

$$
\begin{aligned}
V_0 &= D\left[p^* V_u + (1 - p^*) V_d\right] \\
&= \mathbf{E}^{\mathbb{Q}}\left[D \max(S_T - E, 0)\right]
\end{aligned}
$$

FAPF

Key points of this section...

- There is *one and only one* equivalent martingale measure $\mathbb{Q}$ and it coincides with the "risk-neutral" measure $\mathbb{P}^*$ induced by the parameter p^*.
- The binomial model is equivalent to the martingale approach we developed in ~~Lecture 3.3~~ last night. The option price can be obtained by evaluating the risk-neutral/equivalent martingale expectation:

$$V_0 = \mathbf{E}^{\mathbb{Q}}\left[D \max\left[S_T - E\right]\right] \tag{4}$$

FAPF

3. The PDE approach and the Binomial Model

In this section, we will explore the connection between the binomial model and the PDE approach by deducing the Black-Scholes PDE from the Cox-Ross-Rubinstein implementation of the binomial model.

3.1. The Cox-Ross-Rubinstein Implementation

There are different implementations of the binomial model, due to Cox, Ross and Rubinstien (79), Jarrow-Rudd (83) and Leisen-Reiner (96). Each of these implementation distinguishes itself with a specific choice of parameters for u, d and sometimes p.

The implementation we will use in this lecture is the Cox, Ross and Rubinstien (CRR). In the CRR model,

- $u := \boxed{e^{\sigma\sqrt{\delta t}}};$
- $d := e^{-\sigma\sqrt{\delta t}} = \frac{1}{u};$
- $D := e^{-r\delta t}.$

as $\delta t \to 0$ → This implementation should converge to the RN SDE for S:

where r is the instantaneous risk-free rate and σ is the instantaneous volatility of the stock (log) return.

$$\boxed{dS_t = r S_t\,dt + \sigma S_t\,dX^Q(t)}$$

What is so special about the CRR implementation?

The parameters u, d and D have been chosen so that, in the limit as $\delta t \to 0$, the stock price S_t follows a geometric Brownian motion.

With this choice, our binomial model is consistent with the assumptions made by Black and Scholes.

3.2. From CRR to Black-Scholes PDE

Our starting point is the delta hedging valuation formula (1):

$$V = \frac{V_u - V_d}{u - d} + D\frac{uV_d - dV_u}{u - d}$$

which can be written more conveniently as

$$(1/D)(u - d)V = (1/D)(V_u - V_d) + (uV_d - dV_u) \qquad (5)$$

$$u = e^{\sigma\sqrt{\delta t}} \approx 1 + \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2 \delta t$$

$$d = e^{-\sigma\sqrt{\delta t}} \approx 1 - \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2 \delta t$$

Next, we perform a Taylor expansion of u, d and $1/D$ up to order δt (anything beyond that is too small) to obtain

Ignore (δt^{α}) $\alpha > 1$

$$
\begin{aligned}
u &= e^{\sigma\sqrt{\delta t}} \approx 1 + \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t \\
d &= e^{-\sigma\sqrt{\delta t}} \approx 1 - \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t \\
1/D &= e^{r\delta t} \approx 1 + r\delta t
\end{aligned}
$$

$$D = e^{-r\delta t}$$

$$\rightarrow \frac{1}{D} = e^{r\delta t} \approx 1 + r\delta t$$

$$u - d \approx \underbrace{\left(\cancel{1} + \sigma\sqrt{\delta t} + \cancel{\tfrac{1}{2}\sigma^2\delta t}\right) - \left(\cancel{1} - \sigma\sqrt{\delta t} + \cancel{\tfrac{1}{2}\sigma^2\delta t}\right)}_{2\sigma\sqrt{\delta t}}$$

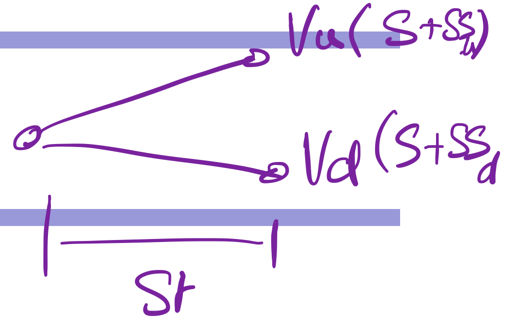
$$\frac{1}{D} \approx \cancel{1} + r\,\delta t$$

$$-d \approx -\cancel{1} + \sigma\sqrt{\delta t} - \frac{1}{2}\sigma^2\delta t$$

Referring to equation (5) we therefore have

$$
\begin{aligned}
u - d &\approx 2\sigma\sqrt{\delta t} \\
1/D - d &\approx r\delta t + \sigma\sqrt{\delta t} - \frac{1}{2}\sigma^2\delta t \\
u - 1/D &\approx \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t - r\delta t
\end{aligned}
$$

$V_u = V(t+\delta t, S+\delta S_u)$

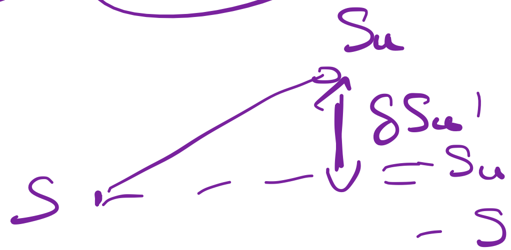


We now perform a Taylor expansion of V_u:

Taylor

$$V_u \approx V + \frac{\partial V}{\partial t}\delta t + \frac{\partial V}{\partial S}\delta S_u + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(\delta S_u)^2 + \cdots$$

where



$$\begin{aligned}\delta S_u &= S_u - S\\ &= S\left(\underbrace{e^{\sigma\sqrt{\delta t}}}_{u} - 1\right)\\ &\approx S\left(\sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t\right)\end{aligned}$$

$S_u = uS = S e^{\sigma\sqrt{\delta t}}$

and hence

$\approx S^2\left(\sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t\right)^2$

$$(\delta S_u)^2 \approx S^2\sigma^2\delta t$$

$V_d = V(t+\delta t, S+\delta S_d)$

Similarly, for V_d:

$$V_d \approx V + \frac{\partial V}{\partial t}\delta t + \frac{\partial V}{\partial S}\delta S_d + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(\delta S_d)^2$$

where

$$\begin{aligned} \delta S_d &= S_d - S \\ &= S\left(e^{-\sigma\sqrt{\delta t}} - 1\right) \\ &\approx S\left(-\sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t\right) \end{aligned}$$

$S \xrightarrow{\delta t}$ $\delta S_d = S_d - S$

S_d $= S \times d - S$

$= S(e^{-\sigma\sqrt{\delta t}} - 1)$

and hence

$$(\delta S_d)^2 \approx S^2\sigma^2\delta t$$

Substituting all these terms into equation (5), we get

$$
\begin{aligned}
&(2\sigma\sqrt{\delta t})(1 + r\delta t)V \\
\approx \quad & \left.\left(V + \frac{\partial V}{\partial t}\delta t + \frac{\partial V}{\partial S}S\left(\sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t\right) + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}S^2\sigma^2\delta t\right)\right] V_u \\
&\times(r\delta t + \sigma\sqrt{\delta t} - \frac{1}{2}\sigma^2\delta t) \\
&\left.+\left(V + \frac{\partial V}{\partial t}\delta t + \frac{\partial V}{\partial S}S\left(-\sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t\right) + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}S^2\sigma^2\delta t\right)\right] V_d \\
&\times(\sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t - r\delta t)
\end{aligned}
$$

$$\frac{\delta t^{\alpha}}{\delta t} = \delta t^{\alpha - 1}$$

$\beta = \alpha - 1$

You are Here

Developing (make sure to keep all the terms), rearranging and dividing by $2\sigma\sqrt{\delta t}$, we get

$$+\frac{\partial V}{\partial t}\delta t + rS\frac{\partial V}{\partial S}\delta t + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\delta t - rV\delta t \approx 0$$

divide by δt

$O(\delta t^{\alpha})$ $\alpha > 1$

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} - rV \approx 0$$

$\approx O(\delta t^{\beta})$

limit $\delta t \to 0$

$\beta = \alpha - 1$, $\beta > 0$

Dividing by δt and taking the limit as $\delta t \to 0$, we finally obtain the **Black-Scholes PDE**:

equality = victory

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0$$

in the limit as $\delta t \to 0$

Key points of this section...

We have just seen that the Black-Scholes PDE is "contained" in the binomial model's delta-hedging valuation equation (1).

In fact, we could venture further and assert that **absence of arbitrage**, the necessary condition for delta hedging to be possible, implies the existence of a "Black-Scholes"-type PDE.

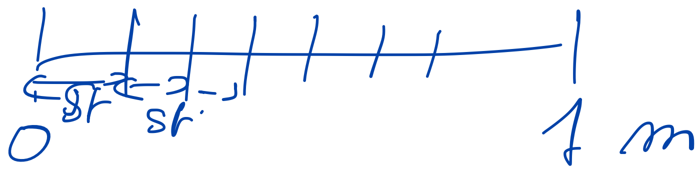
In other terms, in a valuation problem, IF you have absence of arbitrage, you can delta-hedge and thus derive a PDE for the dynamics of your derivative.

On a technical note, and as could have been expected, deriving the Black-Scholes PDE from the binomial model made only use of local arguments (Taylor).

4. The binomial model and the Black-Scholes formula

In this section, we will explore the connection between Cox-Ross-Rubinstein implementation of the binomial model and the Black-Scholes formula.

4.1. The *N*-period binomial model

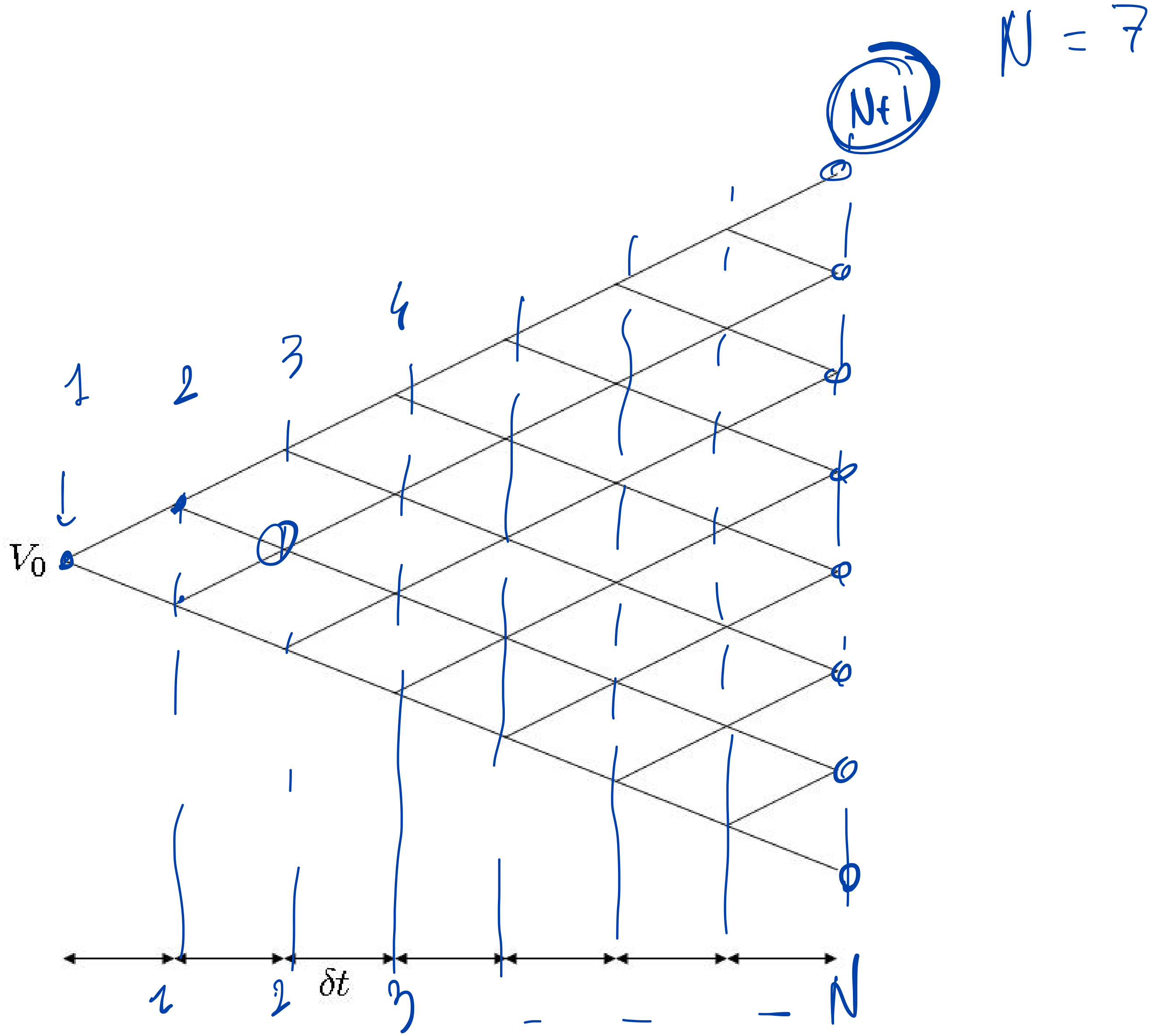


Let's say that we want to price a European call on a stock using the binomial model. The call has exercise price E and matures at time T. The stock has instantaneous volatility σ and the instantaneous risk-free rate is r.

For pricing purpose, a one-period binomial model is clearly not good enough. We will therefore consider a N-period binomial model by splitting the lifetime T of the option into N steps of "length" δt with

$$\delta t = \frac{T}{N}$$

Figure : N-step binomial tree



At each time step, the stock can go up with probability p and down with probability $1 - p$.

Our model of choice will once again be the CRR, so that the up-move, down-move and discount factor for each time step δt is given by:

- $u := e^{\sigma\sqrt{\delta t}}$;
- $d := e^{-\sigma\sqrt{\delta t}} = \frac{1}{u}$;
- $D := e^{-r\delta t}$.

4.2. Recursive computation of the option price

Since we know the one-step binomial valuation model very well, the natural way of disentangling the N-step pricing problem is to proceed recursively, one step at a time, from expiry backwards in time to time 0.

For each time step, we need to solve the option valuation problem a given number of times. For example, at time step $N - k$, we have $N - k + 1$ option values to compute.

Let's introduce some notation so that we can find our way more easily through the binomial tree.

Figure : N-step binomial tree

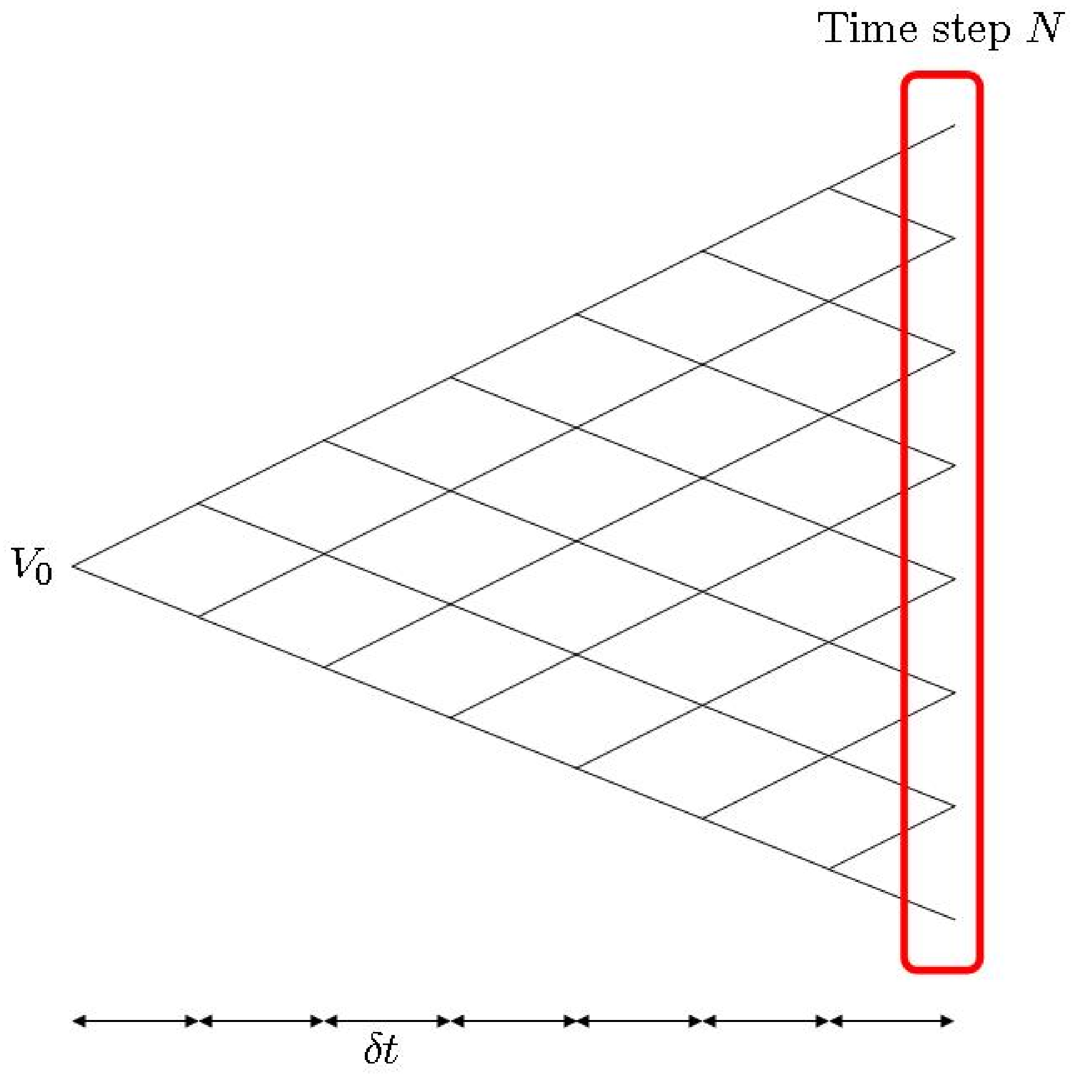
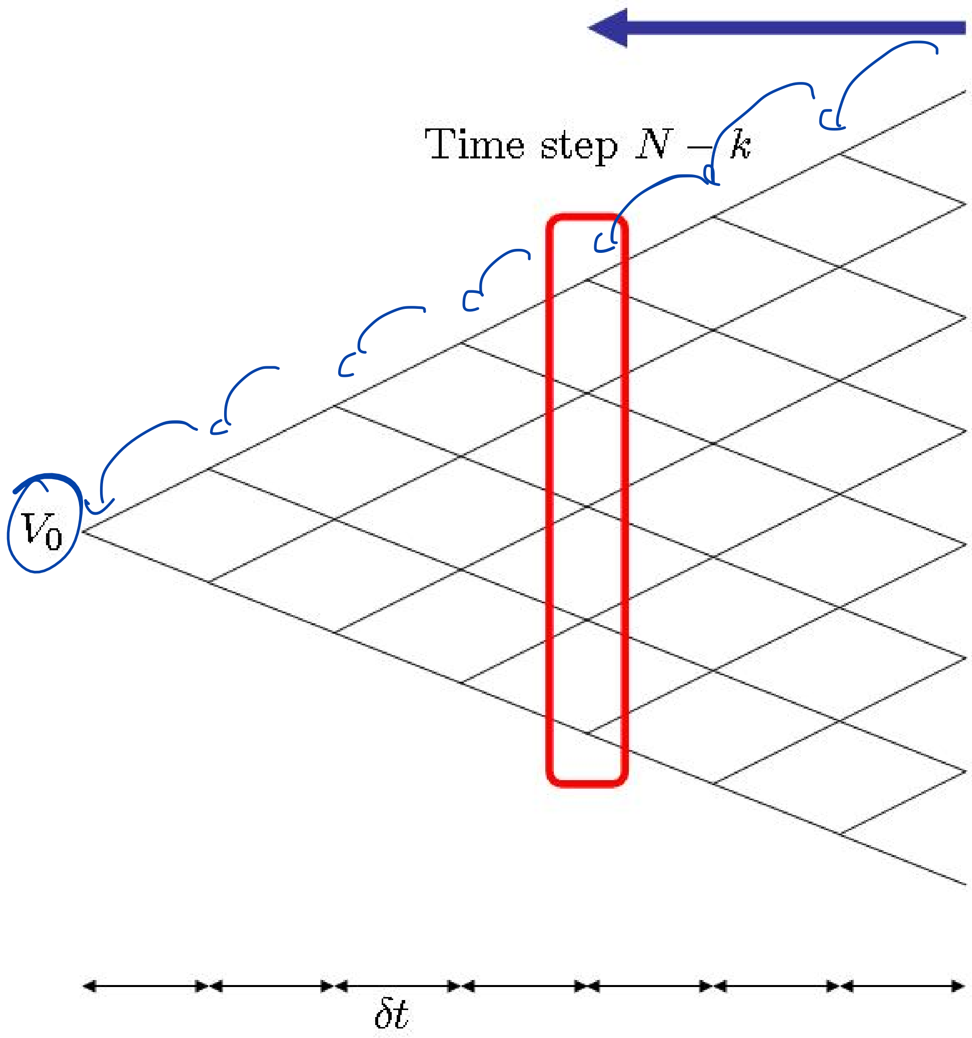


Figure : *N*-step binomial tree



$k = 3$

Figure : *N*-step binomial tree

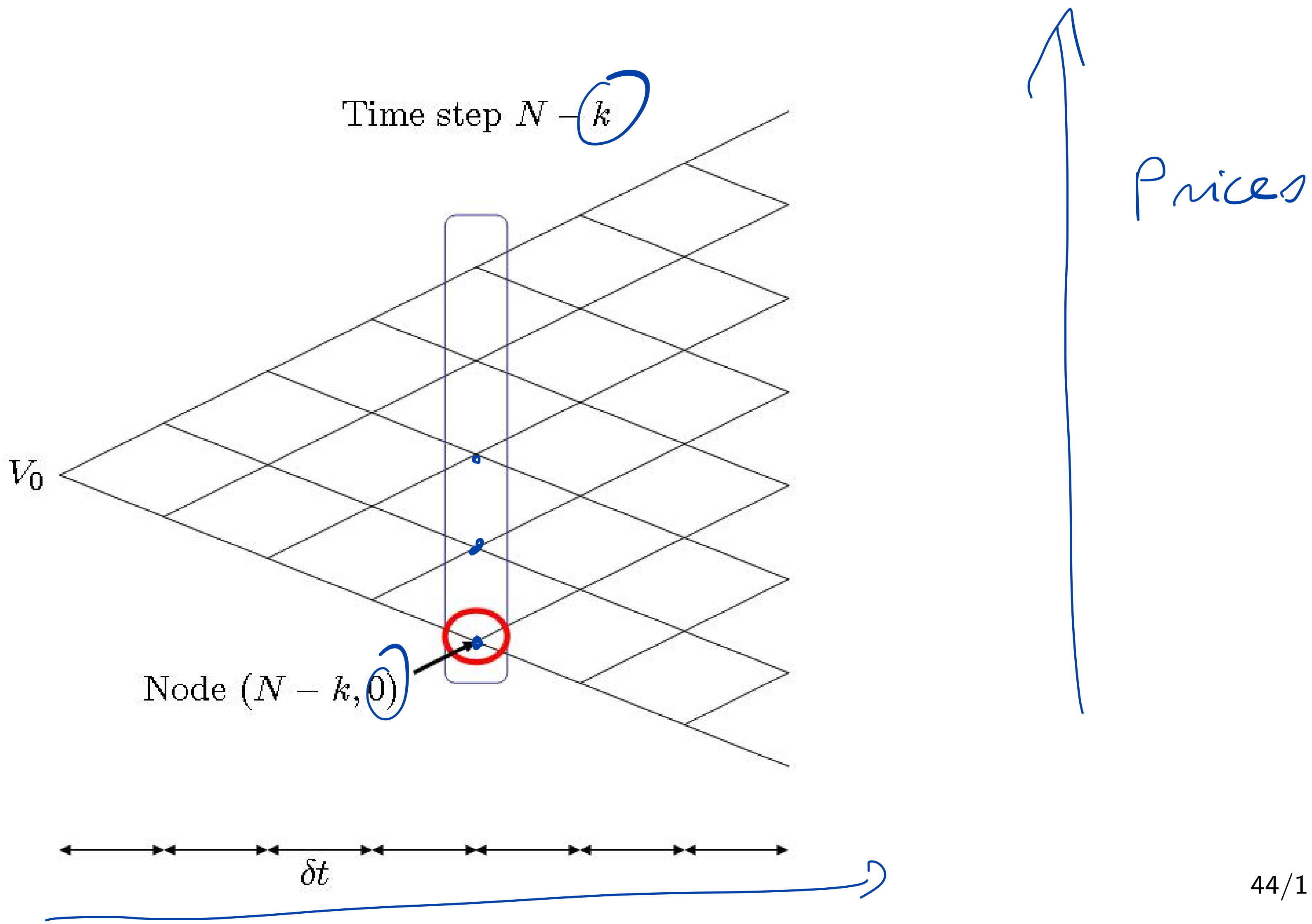


Figure : *N*-step binomial tree

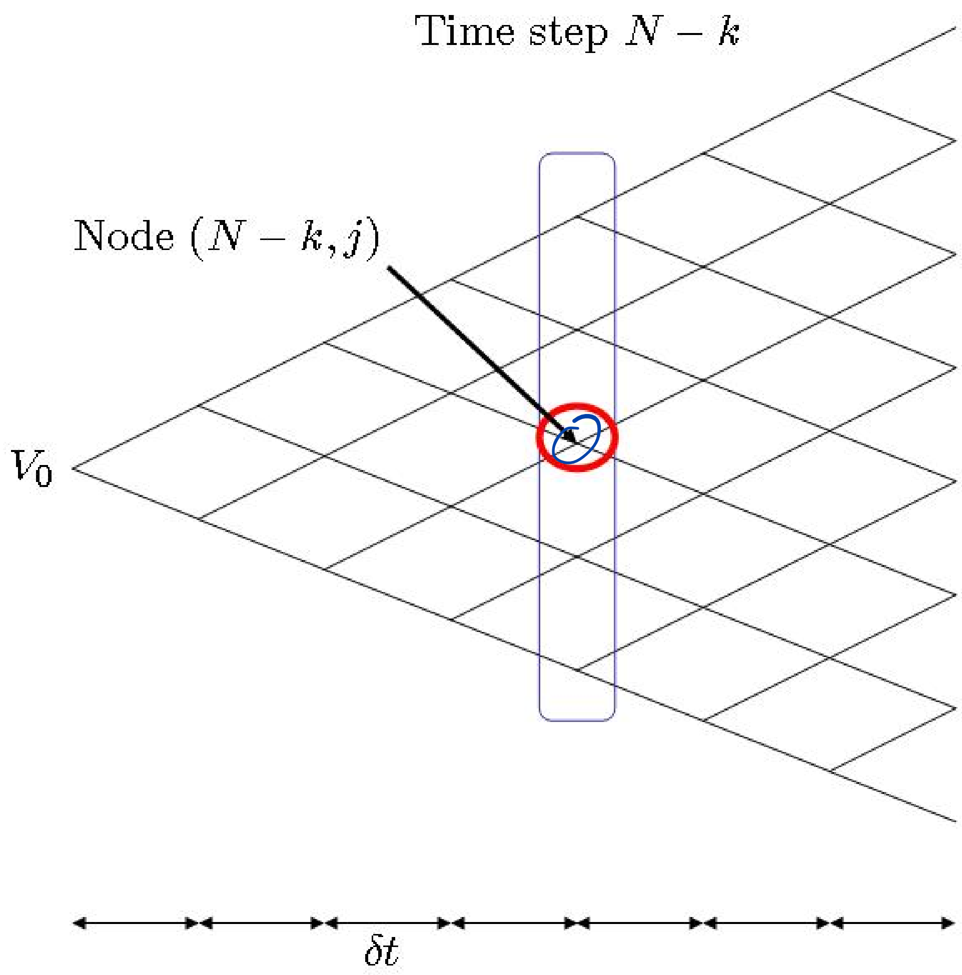
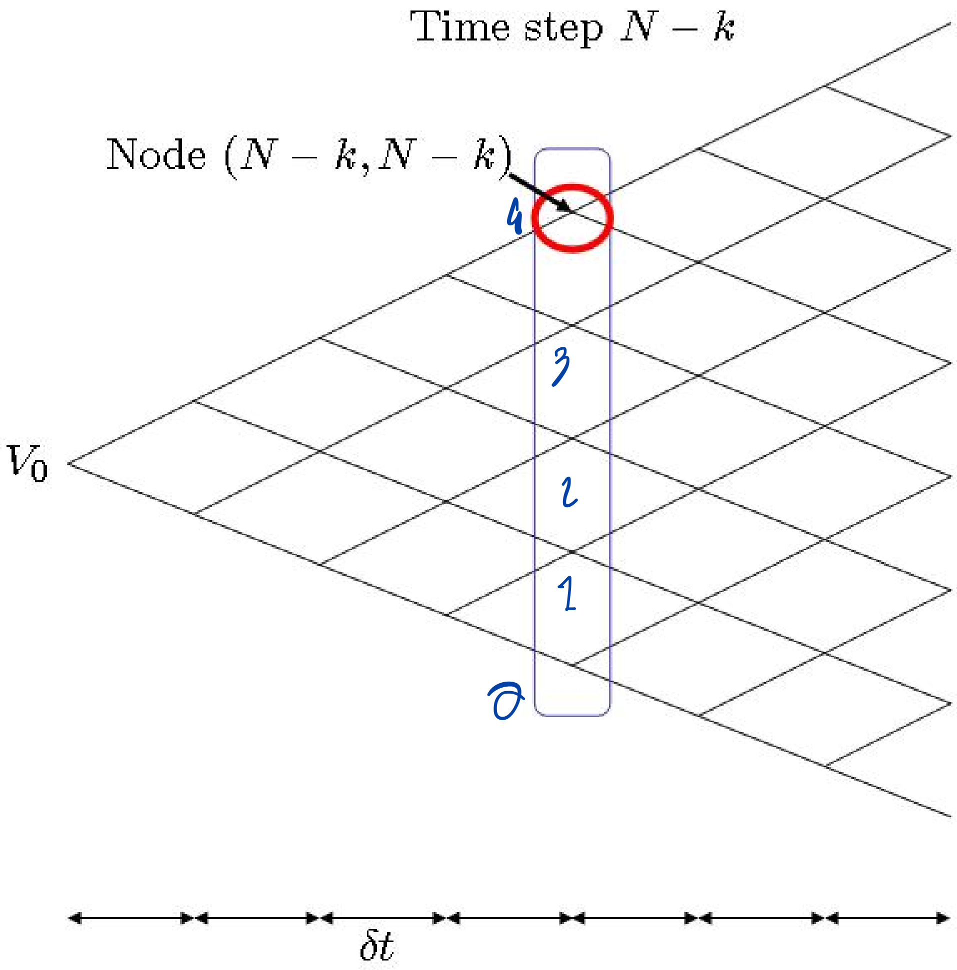


Figure : *N*-step binomial tree

$N = 7$ $k = 3$

$N - k = 4$



We will use the number of time steps (counted backwards) and the stock price movement as a system of coordinates of the form

$$\text{(Time Steps, Stock Movement)}$$

to uniquely identify each node in the tree.

Hence, $(N-k, j)$ refers to the node

- k steps back in time from expiry;
- j asset steps up from the lowest stock price for this time step;

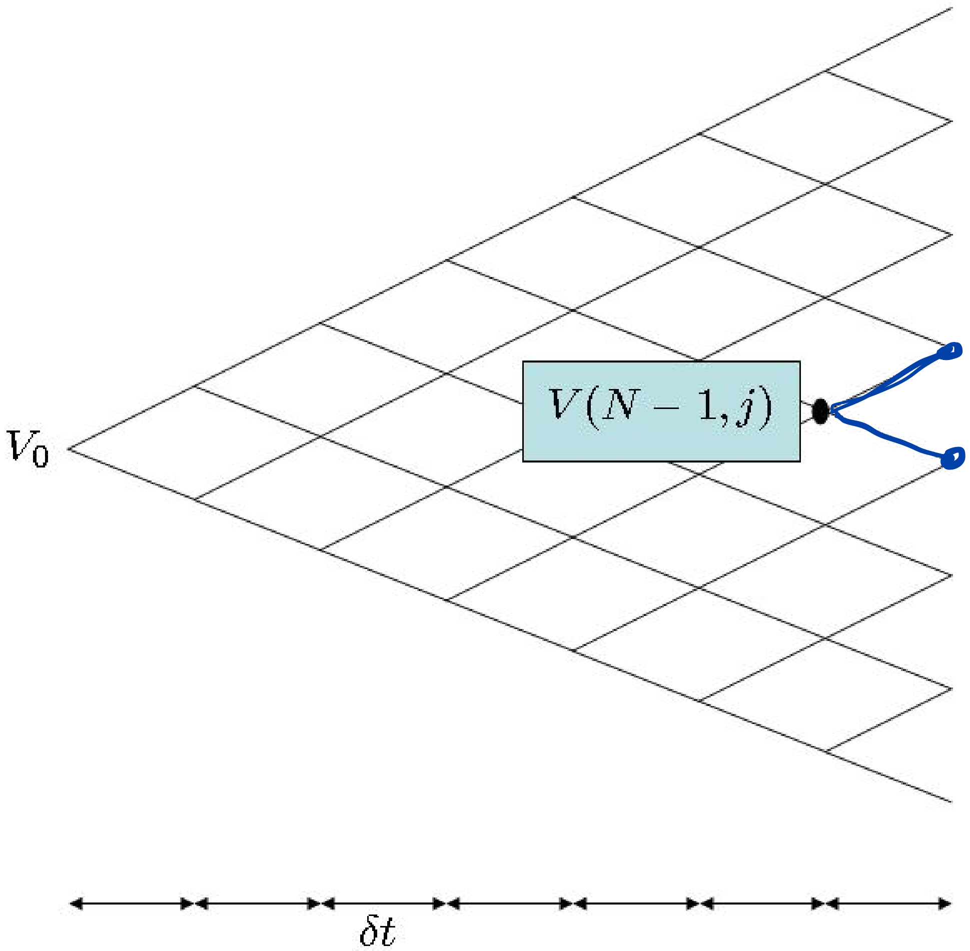
4.2.1 Pricing the Option at Step $N-1$

Let's pick a decision node $(N-1, j)$ with $0 \leq j \leq N-1$ to illustrate.

At time step $N-1$, we just have one hedging decision to make before expiry (corresponding to time N).

Figure : Pricing the Option at Step $N-1$

$V(N-1, j)$



$V(N, j+1) = \text{Payoff}$

$V(N, j) = \text{Payoff}$

Considering our binomial option pricing formula (2), and adapting it to our N-period setting we get the value of the option 1 step from expiry.

Payoff in up state

Payoff in downstate

elementary, Bernoulli!, FAPF

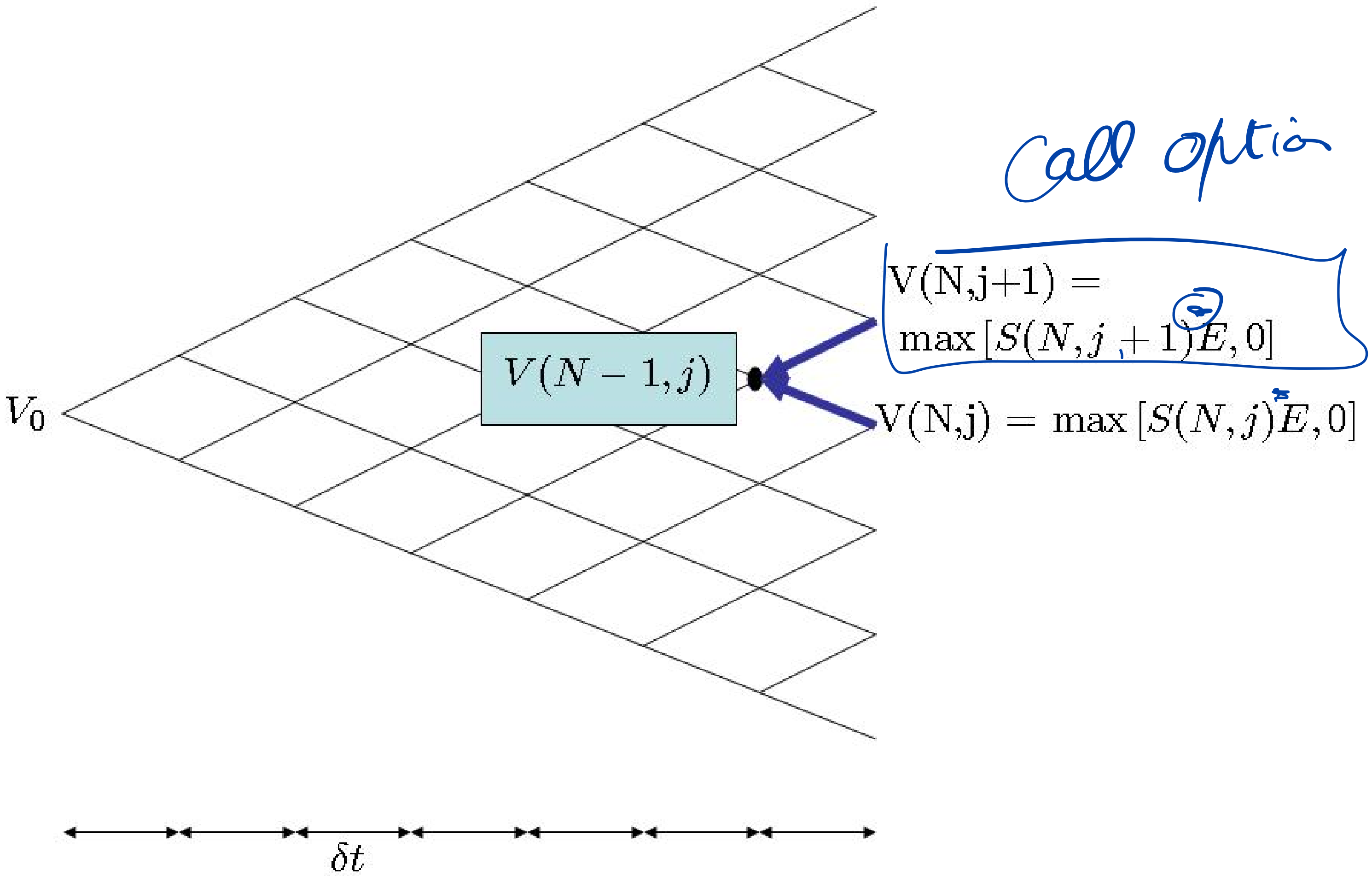
$$V(N-1,j) = D\left[p^* \boxed{V(N,j+1)} + (1-p^*)V(N,j)\right]$$

where

$$p^* = \frac{e^{r\delta t} - d}{u - d}$$

Hence, to compute $V(N-1,j)$, we need to know the value of $V(N,j+1)$ and $V(N,j)$

Figure : Pricing the Option at Step $N-1$

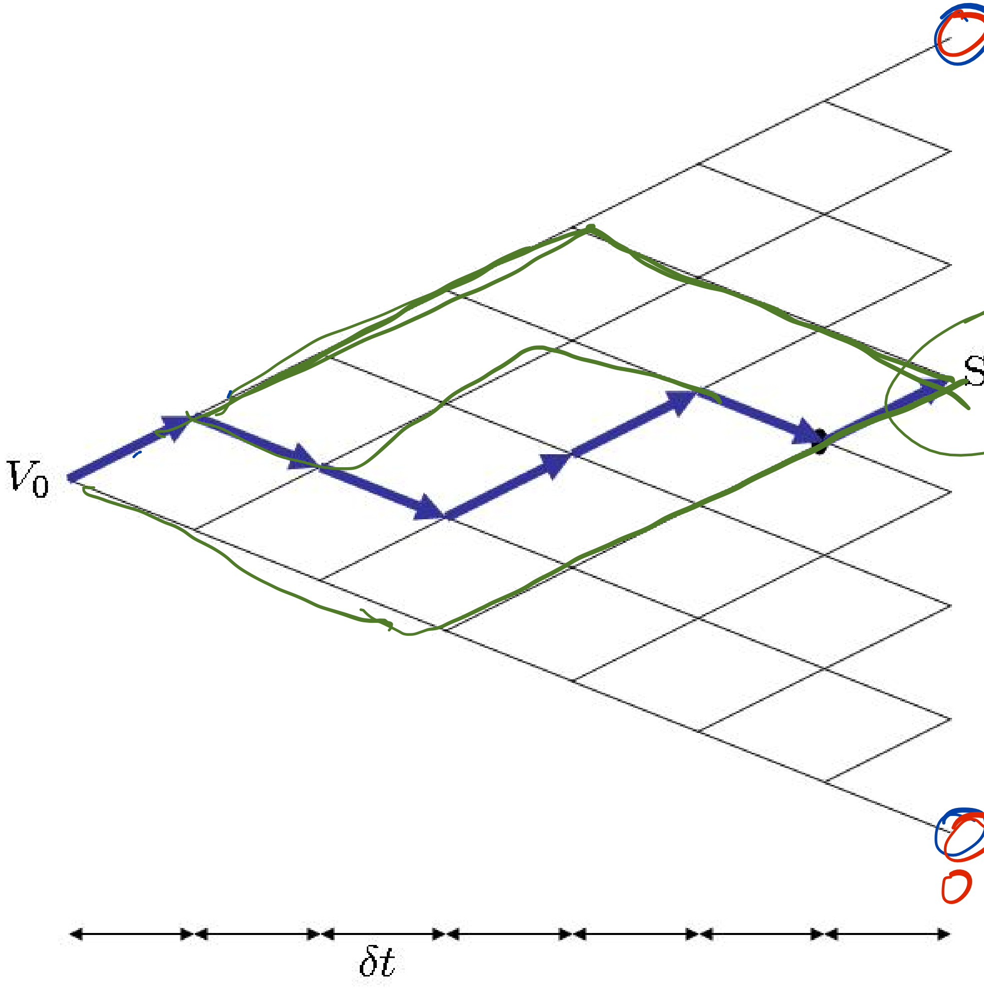


This is rather simple since time step N represents the terminal time T, and hence $V(N, j+1)$ and $V(N, j)$ are respectively the option payoffs:

$$\begin{aligned} V(N, j+1) &= \max\left[S(N, j+1) - E, 0\right] \\ V(N, j) &= \max\left[S(N, j) - E, 0\right] \end{aligned}$$

The next question is: how do we compute $S(N, j)$?

Figure : A possible path leading to $S(N,j)$



N

$S(N,N) \to$ N up
$\to$ 0 down

$S(N,N) = S_0 u^N d^0$

$S(N,j) = \boxed{S_0 u^j d^{N-j}}$

$S(N,0) = S_0 u^0 d^N$

0

$S(N,0) \Rightarrow$ N down
$\Rightarrow$ 0 up

Note that $S(N, N)$ corresponds to N up-moves and zero down-moves.

Conversely, $S(N, 0)$ corresponds to zero up-moves and N down-moves.

Thus, $S(N, j)$ corresponds

- j up-moves;
- $N - j$ down-moves.

which implies that

$$S(N, j) = S_0 u^j d^{N-j}$$

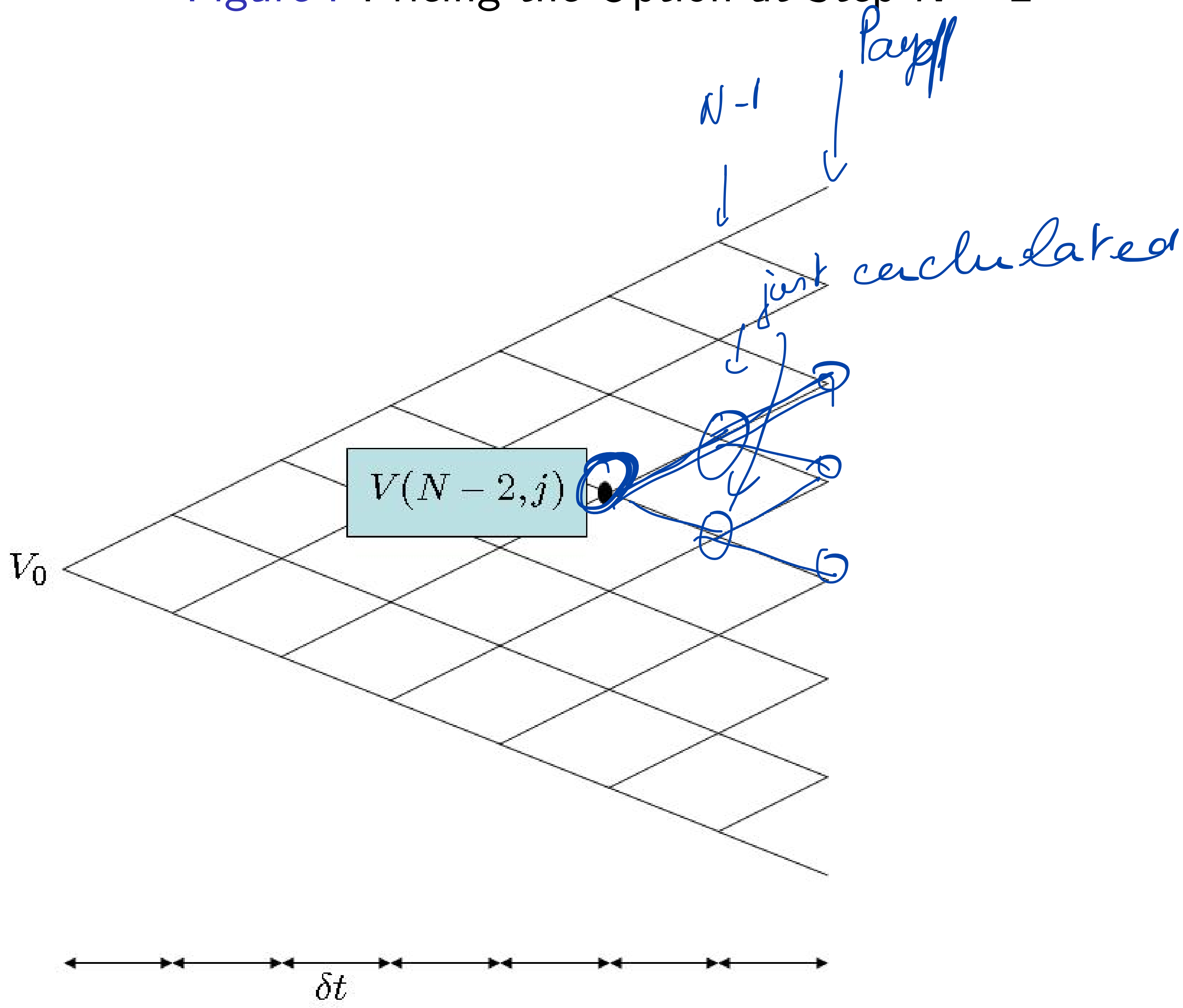
To conclude for this step, the option value $V(N-1,j)$ for $0 \leq j \leq N$ can be computed as:

$$\begin{aligned} V(N-1,j) &= D\left[p^* \max\left[S(N,j+1)-E,0\right]\right. \\ &\left. +(1-p^*)\max\left[S(N,j)-E,0\right]\right] \end{aligned}$$

4.2.2 Pricing the Option at Step $N-2$

Let's pick a decision node $(N-2, j)$ with $0 \leq j \leq N$ to illustrate.

Figure : Pricing the Option at Step $N-2$

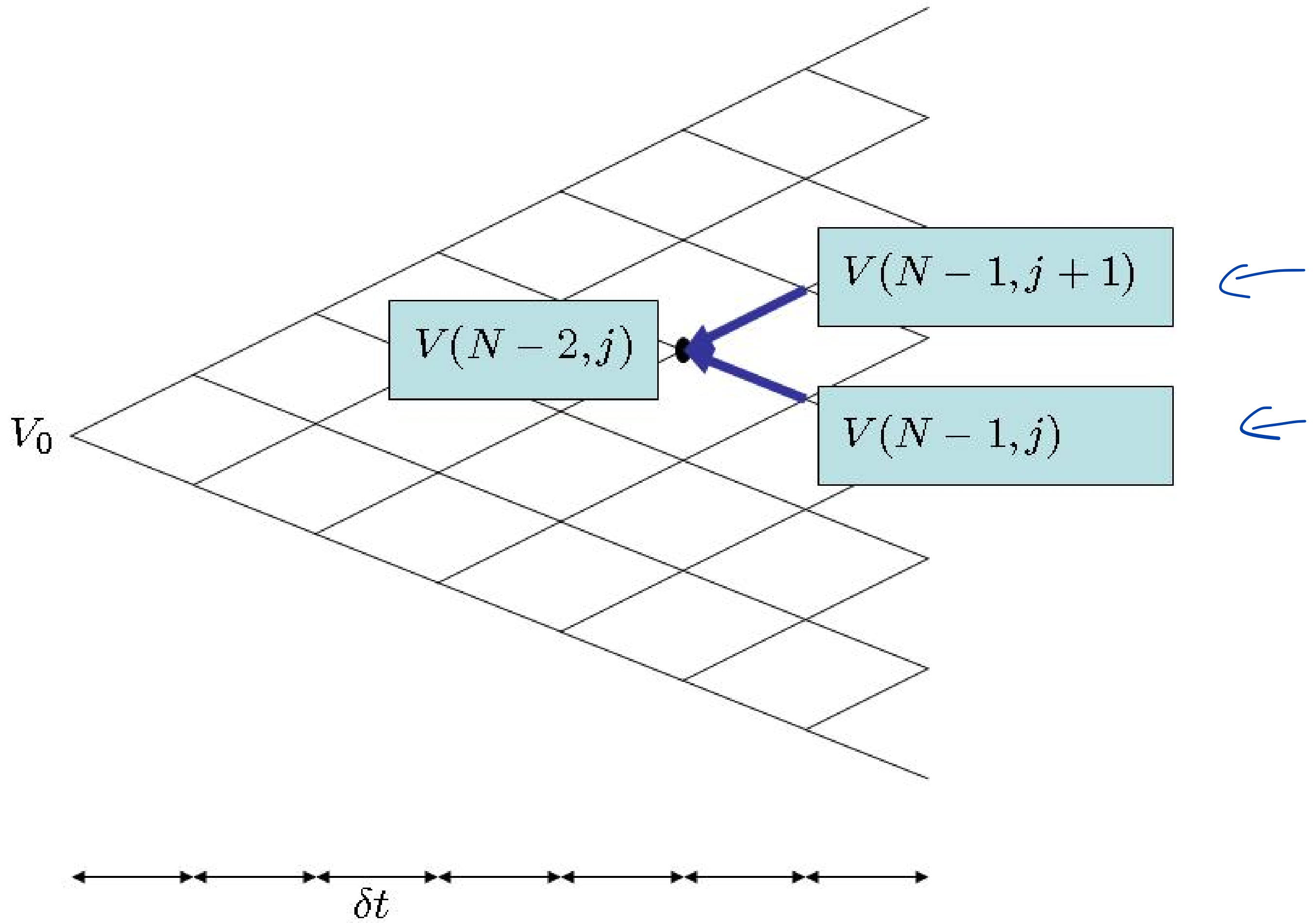


Again, we use our binomial option pricing formula (2). Adapting it to our N-period setting, we get

$$V(N-2,j) = D\left[p^* V(N-1,j+1) + (1-p^*)V(N-1,j)\right]$$

Hence, to compute $V(N-2,j)$, we need to know the value of $V(N-1,j+1)$ and $V(N-1,j)$...

Figure : Pricing the Option at Step $N-2$



... and we have just computed these two values in the previous step. Developing our expression for the option value:

$$
\begin{aligned}
V(N-2,j) &= D\left[p^* V(N-1,j+1) + (1-p^*)V(N-1,j)\right] \\
&= D\left[p^*\left(D\left[p^* \max\left[S(N,j+2)-E,0\right]\right.\right.\right. \\
&\quad + \left.\left.(1-p^*)\max\left[S(N,j+1)-E,0\right]\right]\right) \\
&\quad +(1-p^*)\left(D\left[p^* \max\left[S(N,j+1)-E,0\right]\right.\right. \\
&\quad \left.\left.\left. +(1-p^*)\max\left[S(N,j)-E,0\right]\right]\right)\right] \\
&= D^2\left((p^*)^2 \max\left[S(N,j+2)-E,0\right]\right. \\
&\quad +2p^*(1-p^*)\max\left[S(N,j+1)-E,0\right] \\
&\quad \left.+(1-p^*)^2\max\left[S(N,j)-E,0\right]\right) \qquad (6)
\end{aligned}
$$

As far as intuition goes, this development hints at a binomial expansion of the form:

$$(x+y)^n = x^n + nx^n y + \ldots = \sum_{i=0}^{n} C_n^i x^i y^{(n-i)}$$

where N is an integer and

$$C_n^i := \binom{n}{i} := \frac{n!}{i!(n-i)!}$$

$S(N, j)$

collection of N moves

pick j up moves

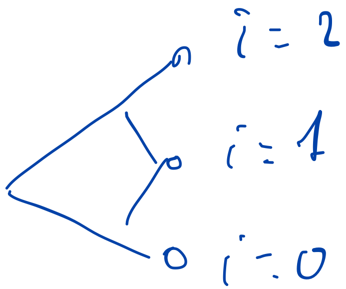
is the number of combinations (i.e. the number of ways one can pick i objects from a bag of n objects when the order in which the objects are picked does not matter).

$C_2^0 = 1$ $\quad$ $C_2^2 = 1$

With this in mind, and noting that $C_m^0 = C_m^m = 1$, we could express formula (6) as

Value of the Option

$$V(N-2,j) = D^2\left(\sum_{i=0}^{2} C_2^i (p^*)^i (1-p^*)^{2-i} \max\left[S(N,j+i)-E,0\right]\right) \quad (7)$$



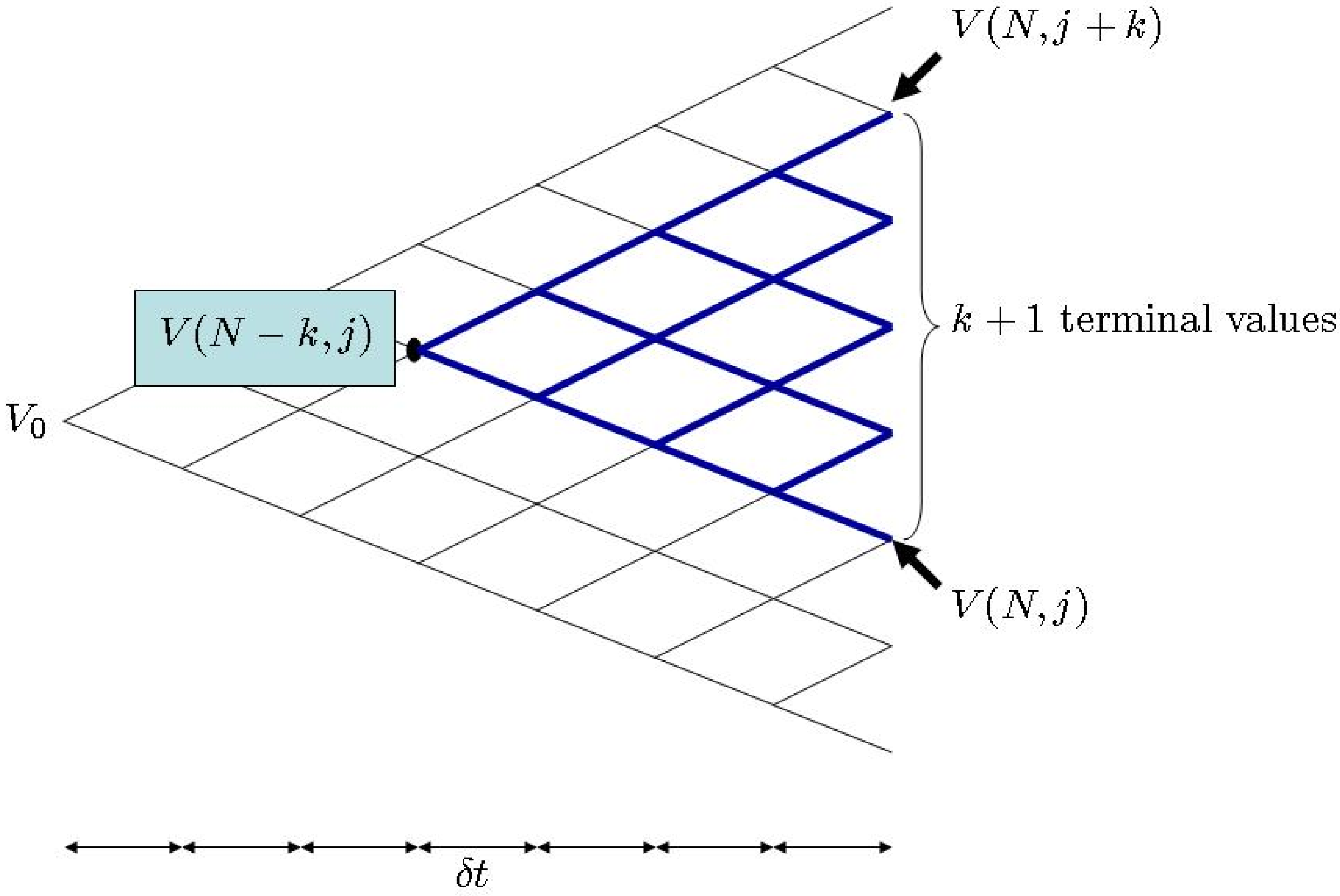
4.2.4 Step $N - k$

If our intuition is correct, then the option value $V(N-k,j)$ at the time step $N-k$ and asset step j with $0 \le j \le N$ can be expressed as:

$$
\begin{aligned}
& V(N-k,j) \\
= \quad & D^k \left(\sum_{i=0}^{k} C_k^i (p^*)^i (1-p^*)^{k-i} \max\left[S(N,j+i)-E,0\right] \right) \quad (8)
\end{aligned}
$$

Let's verify that this statement is true.

Figure : Intuition for the option value at time step $N - k$



Using the one-step formula for the binomial model (2), we know that the option value $V(N-k,j)$ at the time step $N-k$ and asset step j with $0 \leq j \leq N$ is equal to:

$$\begin{aligned} V(N-k,j) &= D\,(p^* V(N-(k-1),j+1) \\ &\quad +(1-p^*)V(N-(k-1),j)) \end{aligned} \tag{9}$$

If our intuition is correct, then the same type of formula as (8) applies to the option values at nodes $(N-(k-1), j+1)$ and $(N-(k-1), j)$. Hence, we conjecture that

$$\begin{aligned} & V(N-(k-1), j+1) \\ = \quad & D^{k-1}\left(\sum_{i=0}^{k-1} C_{k-1}^{i}(p^*)^i(1-p^*)^{k-1-i}\max\left[S(N, j+1+i)-E, 0\right]\right) \end{aligned} \tag{10}$$

and

$$\begin{aligned} & V(N-(k-1), j) \\ = \quad & D^{k-1}\left(\sum_{i=0}^{k-1} C_{k-1}^{i}(p^*)^i(1-p^*)^{k-1-i}\max\left[S(N, j+i)-E, 0\right]\right) \end{aligned} \tag{11}$$

Substituting (10) and (11) into (9), we get

$$
\begin{aligned}
& V(N-k,j) \\
= \quad & D\left[p^* D^{k-1}\left(\sum_{i=0}^{k-1} C_{k-1}^{i}(p^*)^{i}(1-p^*)^{k-1-i} \max\right.\right. \\
& \times \left[S(N,j+1+i)-E,0\right]\Big) \\
& +(1-p^*)D^{k-1} \\
& \left.\left.\times\left(\sum_{i=0}^{k-1} C_{k-1}^{i}(p^*)^{i}(1-p^*)^{k-1-i} \max\left[S(N,j+i)-E,0\right]\right)\right]
\end{aligned}
$$

This can be rewritten as

$$
\begin{aligned}
& V(N-k,j) \\
= \;& D^{k}\left(\sum_{i=0}^{k-1} C_{k-1}^{i}(p^*)^{i+1}(1-p^*)^{k-1-i}\max\left[S(N,j+1+i)-E,0\right]\right. \\
& \left.+\sum_{i=0}^{k-1} C_{k-1}^{i}(p^*)^{i}(1-p^*)^{k-i}\max\left[S(N,j+i)-E,0\right]\right) \qquad (12)
\end{aligned}
$$

The first term on the right-hand side is a term in

$$\max\left[S(N, j+1+i) - E, 0\right]$$

while the second term on the right-hand side is a term in

$$\max\left[S(N, j+i) - E, 0\right]$$

To be able to add these two terms together, we will need to change the indexing of the first term from

$$\sum_{i=0}^{k-1} C_{k-1}^{i}(p^*)^{i+1}(1-p^*)^{k-1-i}\max\left[S(N,j+1+i)-E,0\right]$$

to

$$\sum_{i=1}^{k} C_{k-1}^{i-1}(p^*)^{i}(1-p^*)^{k-i}\max\left[S(N,j+i)-E,0\right]$$

by setting "$i+1 \to i$".

With the new indexing, equation (12) can be written as

$$
\begin{aligned}
& V(N-k, j) \\
= \; & D^k \left(\sum_{i=1}^{k} C_{k-1}^{i-1} (p^*)^i (1-p^*)^{k-i} \max\left[S(N, j+i) - E, 0\right] \right. \\
& \left. + \sum_{i=0}^{k-1} C_{k-1}^{i} (p^*)^i (1-p^*)^{k-i} \max\left[S(N, j+i) - E, 0\right] \right)
\end{aligned}
$$

Developing, we get

$$
\begin{aligned}
& V(N-k,j) \\
= \quad & D^k \Big((p^*)^k \max\left[S(N,j+k)-E,0\right] \\
& + \sum_{i=1}^{k-1} C_{k-1}^{i-1} (p^*)^i (1-p^*)^{k-i} \max\left[S(N,j+i)-E,0\right] \\
& + \sum_{i=1}^{k-1} C_{k-1}^{i} (p^*)^i (1-p^*)^{k-i} \max\left[S(N,j+i)-E,0\right] \\
& + (1-p^*)^k \max\left[S(N,j)-E,0\right] \Big)
\end{aligned}
$$

Recalling that

- $C_m^{j-1} + C_m^j = C_{m+1}^j$;
- $C_m^0 = C_m^m = 1$

we can now add term by term to obtain:

$$\begin{aligned} & V(N-k, j) \\ = \; & D^k \left(\sum_{i=0}^{k} C_k^i (p^*)^i (1-p^*)^{k-i} \max\left[S(N, j+i) - E, 0 \right] \right) \end{aligned}$$

which confirms our intuition.

4.2.5 Option Price at time 0

We can get the pricing formula for time 0 by setting $k = N$ and $j = 0$ into (13):

$$V(0,0) = \underbrace{D^N}_{\text{Discounting}} \left(\underbrace{\sum_{i=0}^{N} \overbrace{C_N^i (p^*)^i (1-p^*)^{N-i}}^{\text{Probability}}}_{\text{Expected Payoff}} \underbrace{\overbrace{\max\left[S(N,i) - E, 0\right]}^{\text{function = Payoff}}}_{\text{Payoff.}} \right)$$

4.3. From binomial model to Black-Scholes formula

Looking at formula (13) closely, we notice that the terms

$$\boxed{C_N^i (p^*)^i (1-p^*)^{N-i}}$$

are probabilities under the Binomial distribution with N trials and probability of success p^*, $\mathscr{B}(N, p^*)$. Specifically, if $X \sim \mathscr{B}(N, p^*)$, then

$$P[X = i] = C_N^i (p^*)^i (1-p^*)^{N-i}$$

X = count of up moves in a Binomial model
that has N periods

Since the expectation of a function $f(x)$ with respect to a *discrete* random variable X is defined as

$$\mathbf{E}[f(X)] = \sum_{i=1}^{N} p_i \boxed{f(X_i)}$$

then formula (13) can be viewed as the following expectation under a Binomial distribution:

$$V(0,0) \quad = \quad D^N \boxed{\mathbf{E}^{\mathscr{B}(N,p^*)}} \underbrace{[\max[S_T - E, 0]]}_{\text{Payoff}} \qquad (13)$$

Equation (13) is simply an application of the “fundamental” asset pricing formula we derived in Lecture 3.3 and which state that

$$\text{Value of a Derivative} \quad = \quad \mathbf{E}^{\substack{\text{equivalent}\\ \text{martingale}\\ \text{measure}}} \text{[PV of Future Cash Flows]}$$

Since $S(N,i) = S_0 u^i d^{N-i}$, we can go even further by writing formula (13) as:

$$
\begin{aligned}
& V(0,0) \\
= \quad & S_0 D^N \sum_{i=0}^{N} C_N^i (p^*)^i (1-p^*)^{N-i} u^i d^{N-i} \mathbf{1}_{S(N,i)>E} \\
& - E D^N \sum_{i=0}^{N} C_N^i (p^*)^i (1-p^*)^{N-i} \mathbf{1}_{S(N,i)>E}
\end{aligned}
$$

where $\mathbf{1}_{y>a}$ is the indicator function returning 1 if $y > a$ and 0 otherwise.

Regrouping,

$$
\begin{aligned}
& V(0,0) \\
= \quad & S_0 \sum_{i=0}^{N} C_N^i (Dp^* u)^i (D\left[1-p^*\right] d)^{N-i} \mathbf{1}_{S(N,i)>E} \\
& -ED^N \sum_{i=0}^{N} C_N^i (p^*)^i (1-p^*)^{N-i} \mathbf{1}_{S(N,i)>E}
\end{aligned}
$$

Define

$$\bar{p} := Dup^* = Du\frac{\frac{1}{D} - d}{u - d}$$

and note that

$$\begin{aligned} Dd(1 - p^*) &= Dd\frac{u - \frac{1}{D}}{u - d} \\ &= 1 - \bar{p} \end{aligned}$$

then,

$$
\begin{aligned}
& V(0,0) \\
= \quad & S_0 \sum_{i=0}^{N} C_N^i (\bar{p})^i (1-\bar{p})^{N-i} \mathbf{1}_{S(N,i)>E} \\
& -ED^N \sum_{i=0}^{N} C_N^i (p^*)^i (1-p^*)^{N-i} \mathbf{1}_{S(N,i)>E}
\end{aligned}
$$

and the parameter $\bar{p}$ uniquely defines a new measure $\bar{\mathbb{P}}$ (and a new Binomial distribution $\mathscr{B}(N, \bar{p})$).

The previous equation can then be written as

$$\begin{aligned} V(0,0) &= S_0 \mathbf{E}^{\mathscr{B}(N,\bar{p})}\left[\mathbf{1}_{S(N,i)>E}\right] - ED^N \mathbf{E}^{\mathscr{B}(N,p^*)}\left[\mathbf{1}_{S(N,i)>E}\right] \\ &= S_0 \bar{\mathbb{P}}\left[S_T > E\right] - ED^N \mathbb{P}^*\left[S_T > E\right] \qquad (14) \end{aligned}$$

since $\mathbf{E}^{\mathbb{P}}\left[\mathbf{1}_{X\in A}\right] = \mathbb{P}\left[X \in A\right]$.

$B(N, \bar{p})$ $\quad$ $B(N, p^*)$

$N \to \infty$ $\quad$ $N(.)$ $\quad$ $N(.)$

We are now very close to the Black-Scholes formula...

It turns out that it can be shown that as $N \to \infty$ (i.e. $\delta t \to 0$), the formula (14) converges to the Black-Scholes formula:

$$V(0,0) = S_0 N(d_1) - Ee^{-rT} N(d_2)$$

The intuition for this convergence is that by the Central Limit Theorem, the Binomial distribution converges to the Normal distribution as $N \to \infty$. Most of the remaining work consists in verifying that we effectively get d_1 and d_2 out of $\bar{p}$ and p^*.

Key points of this section...

We have just seen that the binomial model's risk-neutral valuation equation (2) tends to the Black-Scholes formula in the limit as the number of time steps tends to infinity.

While deriving and interpreting the risk-neutral valuation equation proved relatively easy, obtaining the Black-Scholes formula as a result proved more difficult.

The reason is that the derivation of the risk-neutral valuation equation is relatively model-independent (subject to probabilistic assumptions), while the Black-Scholes formula is very much model-dependent. Adapting a model-independent result to a specific model often require a significant amount of work because you need to impose all of the model assumptions and structure onto the result.

5. The **BIG** idea...

So far, we have seen that the binomial model connects two seemingly different aspects of derivatives valuation:

- the binomial model and the martingale approach coincide regardless of the specific implementation we choose for the binomial model.
- when we adopt the CRR implementation, the binomial model's *no-arbitrage pricing formula* (1) converges to the Black-Scholes *PDE*;
- when we adopt the CRR implementation, the binomial model's *risk-neutral pricing formula* (2) converges to the Black-Scholes *formula*;

So, what is the **BIG** idea here?

The BIG Idea...

... is that in the Black-Scholes setting, the no-arbitrage approach and the martingale measure approach are strictly equivalent!

In other words, if one of the two methods works on a specific valuation problem then the other method works as well. It is then up-to-you to choose the appropriate method and/or to combine both methods in order to get the most out of the valuation problem.

5.1. No Arbitrage implies that we can find a Martingale Measure

① No Arb → EMM
② EMM → No Arb.

At the begining of the lecture (slide 14), we noted that the no-arbitrage valuation equation (1) could be rewritten as the risk-neutral valuation equation (2) by defining a probability p^*

This probability p^* uniquely defines the risk-neutral measure $\mathbb{P}^*$ which coincide with the martingale measure $\mathbb{Q}$.

Hence, **no arbitrage leads to the existence of a unique equivalent martingale measure**.

5.2 Finding a Martingale Measure implies that there is no arbitrage

For an equivalent martingale measure to exist, we need $0 < p^* < 1$. Indeed, if $p^* = 0$, then $\mathbb{P}^*[S_T = uS] = 0$, but $\mathbb{P}[S_T = uS] > 0$. Hence $\mathbb{P}$ and $\mathbb{P}^*$ are not equivalent. A similar conclusion arises if $p^* = 1$.

$$p^* = \frac{\frac{1}{D} - d}{u - d} > 1 \iff \frac{\frac{1}{D} - d}{u - d} > 1 \iff \frac{1}{D} - \cancel{d} > u - \cancel{d} \iff \boxed{\frac{1}{D} > u}$$

To generalize a little, let's just imagine that no equivalent martingale measure exists because $p^* \geq 1$. This implies that

$$p^* = \frac{\frac{1}{D} - d}{u - d} \geq 1 \Leftrightarrow u \leq \frac{1}{D}$$

Hence $d < u \leq \frac{1}{D}$.

No EMM because $p^* \geq 1$
$\Rightarrow$ Arbitrage exists.

Since D is risk-free, an arbitrage opportunity may exist: shortselling the stock and investing in the bank account guarantees a minimum return equal to $\frac{1}{D} - u \geq 0$. And even if academic arbitrage does not exist, then the "probabilistic" arbitrage does exists: why invest in a risky asset that can at best give you a risk-free return when you could get the risk-free return for sure[2]!

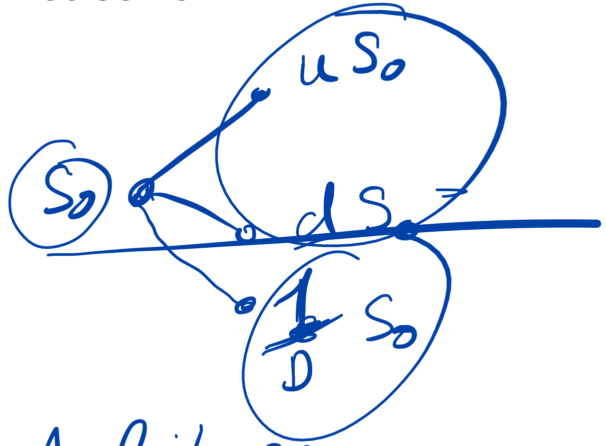
[2] remember that p^* and p are not the same thing

$p^* \le 0$

$p^* \le 0 \iff \dfrac{\frac{1}{D} - d}{u - d} < 0 \iff \dfrac{1}{D} - d < 0 \iff \boxed{\dfrac{1}{D} < d}$

Similarly, let's imagine that no equivalent martingale measure exists because $p^* \le 0$. This implies that

$$p^* = \frac{\frac{1}{D} - d}{u - d} \le 0 \Leftrightarrow \frac{1}{D} \le d$$



Hence $\frac{1}{D} \le d < u$.

When no EMM exists because $p^* \le 0$ $\rightarrow$ Arbitrage

An arbitrage opportunity may also exist: buying the stock and financing the purchase by borrowing from the bank account guarantees a minimum return equal to $d - \frac{1}{D} \ge 0$ with no risk of loss.

$$0 < \frac{\frac{1}{D} - d}{u - d} < 1$$

$$0 < \frac{1}{D} - d < u - d$$

$$\boxed{d < \frac{1}{D} < u}$$

Now, what happens if $\boxed{0 < p^* < 1}$?

This implies that

$$0 < \frac{\frac{1}{D} - d}{u - d} < 1 \Leftrightarrow d < \frac{1}{D} < u$$

$r < 0$

and no arbitrage opportunity exists.

$$\frac{1}{D} = e^{r\delta t}$$

Hence, **the existence of a unique equivalent martingale measure leads to the absence of arbitrage opportunities**.

6. Complete and Incomplete Markets

The BIG idea that the No Arbitrage and Equivalent Martingale Measure approaches are equivalent works well enough in the (continuous) Black-Scholes setting and the binomial model because they are examples of **complete markets**.

6.1. Complete Markets

Broadly speaking, a complete market is a market in which we have enough tradeable instruments to perfectly hedge all of the risk(s) of the derivative we are intending to price.

For example, in the Black-Scholes setting, we are trying to replicate a European option using the underlying stock and the risk-free rate.

The option is exposed to *one* source of risk, the risk of the underlying share modelled as the Brownian motion $X(t)$. Luckily, we can trade and hedge this risk by trading the stock.

Complete markets are so special because in a complete market the derivative is a *redundant security* in the sense that

1. it does not add any new source of risk to the market;
2. it could be fully and efficiently replicated by trading the other securities available on this market.

This is exactly the case in the Black-Scholes world: assuming no trading costs or other market friction, the European option is redundant...

... and this is precisely why No-Arbitrage pricing works so flawlessly.

6.2. Incomplete Markets

Incomplete markets are markets in which you cannot find enough tradeable securities to hedge the risk(s) of the derivative you are trying to price.

The simplest example of an incomplete market is the bond market. One can view bond as a derivative on some (stochastic) interest rates. Since interest rates are not directly tradeable, the only way we have to hedge/replicate a bond B_1 is by trading a related bond B_2 exposed to the same interest rate risk.

The problem here is since we do not know how to price a bond to start with, using bond B_2 as a hedging instrument for bond B_1 is far from solving all our problems, and we will have to make some assumptions if we want to solve the pricing problem.

More on this in **Module 4**!

In this lecture, we have seen...

- the binomial model and the martingale approach coincide regardless of the specific implementation we choose for the binomial model.
- when we adopt the CRR implementation, the binomial model's *risk-neutral pricing formula* (2) converges to the Black-Scholes *formula*;
- when we adopt the CRR implementation, the binomial model's *no-arbitrage pricing formula* (1) converges to the Black-Scholes *PDE*;
- In complete markets, the no-arbitrage approach and the martingale measure approach are strictly equivalent.

Introduction to
Numerical Methods

In this lecture. . . 60%

- The justification for pricing by Monte Carlo simulation M.C
- Grids and discretization of derivatives
- The explicit finite-difference method

F.D.M 30%

By the end of this lecture you will be able to

- implement the Monte Carlo method for simulating asset paths and pricing options

- implement the explicit finite-difference method for pricing options

Introduction

More often than not we must solve option-pricing problems by numerical means.

It is rare to be able to find closed-form solutions for prices unless both the contract and the model are very simple.

The most useful numerical techniques are Monte Carlo simulations and finite-difference methods.

Monte Carlo Simulations

Relationship between derivative values and simulations

Theory says:

- The fair value of an option is the present value of the expected payoff at expiry under a *risk-neutral* random walk for the underlying.

The risk-neutral random walk for S is

- $$dS = rS\,dt + \sigma S\,dX.^{\mathbb{Q}}$$

This is simply our usual lognormal random walk but with the risk-free rate instead of the real growth rate.

Justification:

- Binomial method

- Black–Scholes Equation similar to backward Kolmogorov equation

- Martingale theory

We can therefore write

- $$\text{option value } = e^{-r(T-t)} E\,[\text{payoff}(S)]$$

provided that the expectation is with respect to the risk-neutral random walk, not the *real* one.

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s\, ds}\, \text{Payoff}(S) \right]$$

The algorithm:

1. Simulate the risk-neutral random walk starting at today's value of the asset S_0 over the required time horizon. This gives one realization of the underlying price path.

2. For this realization calculate the option payoff.

3. Perform many more such realizations over the time horizon.

4. Calculate the average payoff over all realizations.

5. Take the present value of this average, this is the option value.

How do we simulate the asset?

Two ways:

1. **If** the s.d.e. for the asset path is integrable **and** the contract is not path dependent (or American) **then** simulate in 'one giant leap'

2. **Otherwise** you will have to simulate time step by time step, the entire path

One giant leap: A method that works in special cases

For the lognormal random walk we are lucky that we can find a simple, and exact, time stepping algorithm.

We can write the risk-neutral stochastic differential equation for S in the form

$$d(\log S) = \left(r - \tfrac{1}{2}\sigma^2\right) dt + \sigma\, dX.$$

This can be integrated exactly to give

$$S(t) = S(0) \exp\left(\left(r - \tfrac{1}{2}\sigma^2\right) t + \sigma \int_0^t dX\right).$$

i.e.

- $$S(T) = S(0) \exp\left(\left(r - \tfrac{1}{2}\sigma^2\right) T + \sigma \sqrt{T}\, \phi\right).$$

Because this expression is exact and simple it is the best time stepping algorithm to use... but only if we have a payoff that only depends on the final asset value, i.e. is European and path independent.

We can then simulate the final asset price in one giant leap, using a time step of T if both of these are true

- the s.d.e. is integrable and
- the contract is European and not path dependent

Simulating the entire path: A method that always works

Price paths are simulated using a discrete version of the stochastic differential equation for S.

Euler - Maruyama

An obvious choice is to use

$$\delta S = rS\,\delta t + \sigma S\,\sqrt{\delta t}\,\phi,$$

$$S_{i+1} = S_i\left(1 + r\,\delta t + \sigma\,\phi\sqrt{\delta t}\right)$$

where ϕ is from a standardized Normal distribution.

- This way of simulating the time series is called the **Euler method**. This method has an error of $O(\delta t)$.

Errors

There are two (at least) sources of error in the Monte Carlo method:

- If the size of the time step is δt then we may introduce errors of $O(\delta t)$ by virtue of the discrete approximation to continuous events

- Because we are only simulating a finite number of an infinite number of possible paths, the error due to using N realizations of the asset price paths is $O(N^{-1/2})$.

Generating Normal variables

ϕ ?

- **Quick 'n' dirty**: A useful distribution that is easy to implement on a spreadsheet, and is fast, is the following *approximation* to the Normal distribution:

$$\sqrt{\frac{12}{n}}\left[\sum_{i=1}^{N} \mathrm{RAND}() - \frac{N}{2}\right]$$

$$\left(\sum_{i=1}^{12} \psi_i\right) - 6,$$

where the ψ_i are independent random variables, drawn from a uniform distribution over zero to one.

There are other methods such as **Box–Muller**, more later.

Polar Marsaglia

Accuracy and computational time

Let's use ϵ to represent the desired accuracy in a MC calculation.

We know that errors are $O(\delta t)$ and $O(1/\sqrt{N})$. It makes sense to have errors due to the time step and to the finite number of simulations to be of the same order (no point in having one link in a chain stronger than another!). So we would choose:

$$\delta t = O(\epsilon) \quad \text{and} \quad N = O(\epsilon^{-2}).$$

The time taken is then proportional to number of calculations, therefore

$$\text{Time taken} = O(\epsilon^{-3}).$$

If you want to halve the error it will take eight times as long.

In higher dimensions...

Suppose you have a basket option with D underlyings. The time taken now becomes

$$\text{Time taken} = O(D\epsilon^{-3}).$$

(Think of having one Excel spreadsheet per asset.)

This is surprisingly insensitive to dimension!

Other issues

- Greeks

- Early exercise (and other decisions)

Advantages of Monte Carlo simulations

- The mathematics that you need to perform a Monte Carlo simulation can be very basic

- Correlations can be easily modeled, and it is easy to price options on many assets (high-dimensional contracts)

- It is computationally quite efficient in high dimensions

- There is plenty of software available, at the very least there are spreadsheet functions that will suffice for most of the time

- To get a better accuracy, just run more simulations

- The effort in getting *some* answer is very low

- The models can often be changed without much work

- Complex path dependency can often be easily incorporated

- Many contracts can be priced at the same time

- People accept the technique, and will believe your answers

Disadvantages of Monte Carlo simulations

- The method is very slow, you need a lot of simulations to get an accurate answer

- Finding the greeks can be hard

- The method does not cope well with early exercise

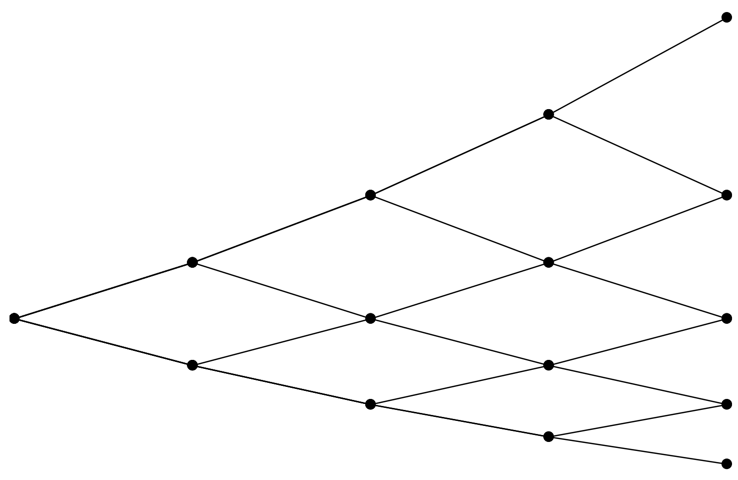
Finite Difference Methods

Monte Carlo simulations can be very slow to converge to the answer, and they do not give us the greeks without further effort.

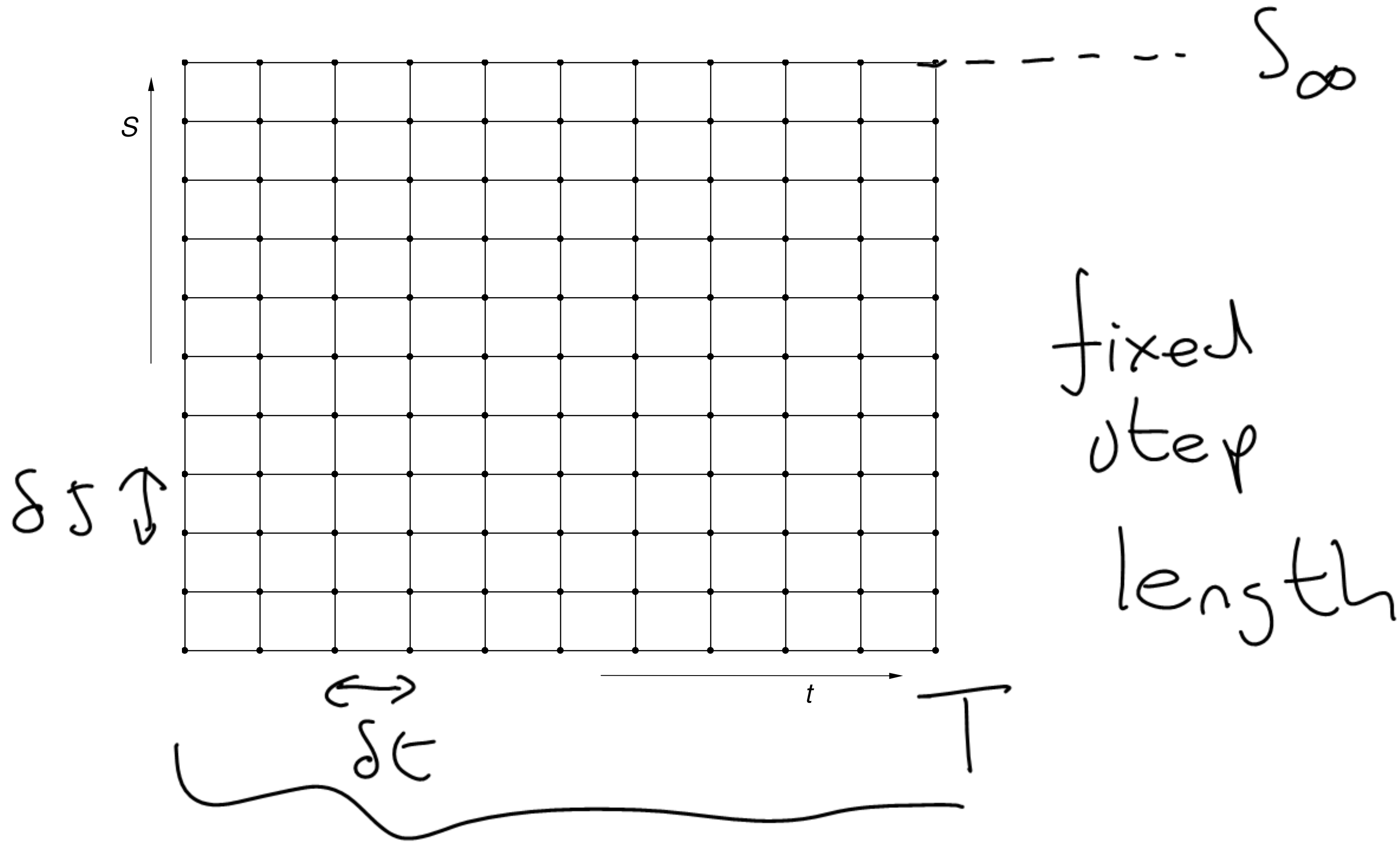
There is a method that is very similar to the binomial tree method which is the method of choice for certain types of problem.

Grids

Recall the shape of the binomial tree...



The shape of the tree is determined by the asset volatility.



The finite-difference grid.

The finite-difference grid usually has equal time steps and equal S steps.

Differentiation using the grid

Notation: time step δt and asset step δS. The grid is made up of the points at asset values

$$\delta t = \frac{T}{K}$$

$$S = i\,\delta S$$

and times

$$t = T - k\,\delta t$$

$$\delta S = \frac{S_\infty}{I}$$

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

We will be solving for the asset value going from zero up to the asset value $I\,\delta S$.

The Black–Scholes equation is to be solved for $0 \leq S < \infty$ so that $I\,\delta S$ is our approximation to infinity.

Write the option value at each of these grid points as

$V(S,t) = V(i\,\delta S, T - k\,\delta t)$

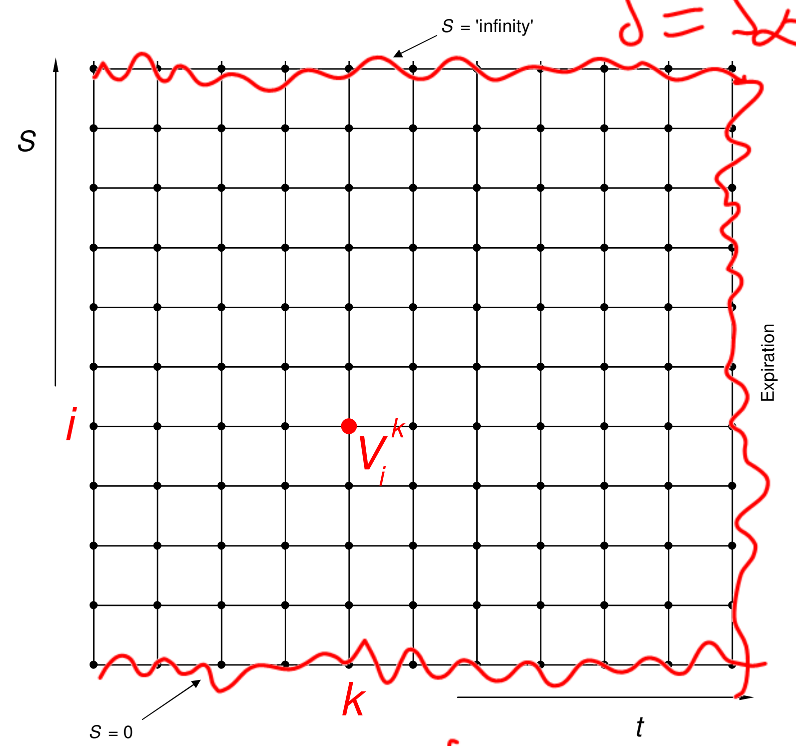
$$V_i^k = V(i\,\delta S, T - k\,\delta t).$$

$V(S,t)$

- The superscript is the time variable and the subscript the asset variable.

$S = \infty$

Unknown



Payoff

$S = 0$

Approximating θ

$\frac{\partial V}{\partial t}$

$\tau = T - t$

$\frac{\partial}{\partial t} = -\frac{\partial}{\partial \tau}$

The definition of the first time derivative of V is simply

$$\frac{\partial V}{\partial t} = \lim_{h \to 0} \frac{V(S,t) - V(S,t-h)}{h}.$$

It follows naturally that we can approximate the time derivative from our grid of values using

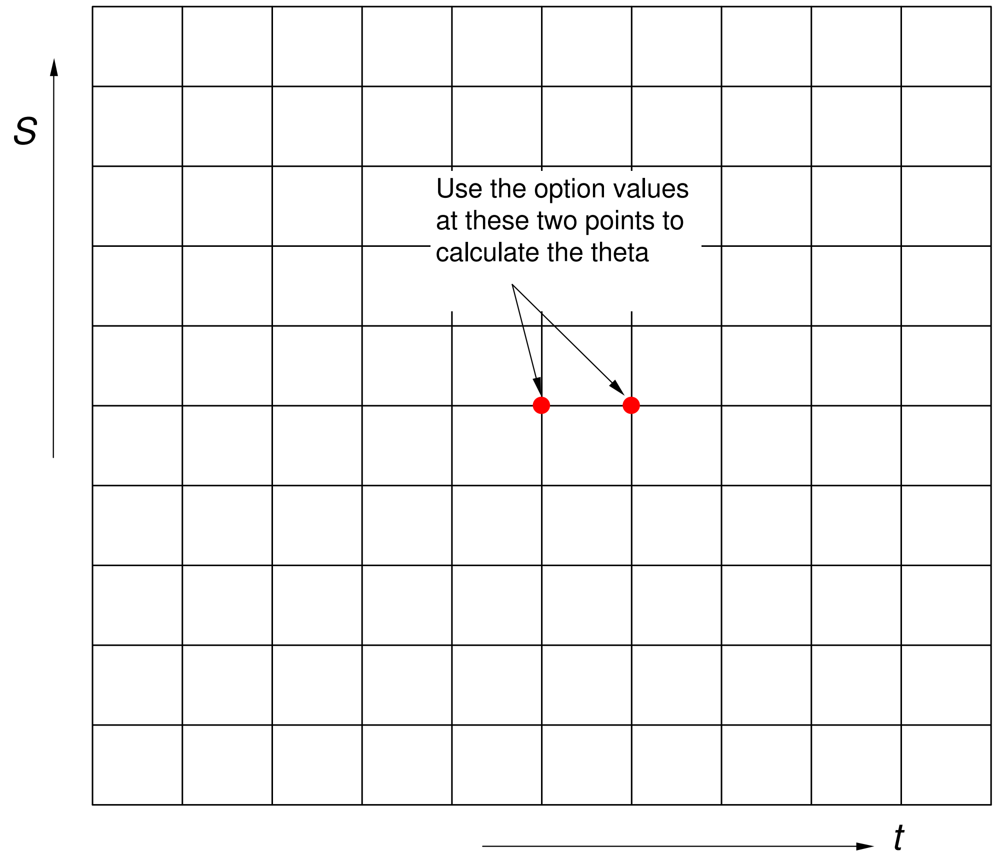
$t = T - k\delta t$

$$\frac{\partial V}{\partial t}(S,t) \approx \frac{V_i^k - V_i^{k+1}}{\delta t}.$$

k

$k+1$

This is our approximation to the option's theta.



Approximating the theta.

How accurate is this approximation?

We can expand the option value at asset value S and time $t - \delta t$ in a Taylor series about the point S, t as follows.

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

In terms of values at grid points this is just

$$V_i^k = V_i^{k+1} + \delta t \frac{\partial V}{\partial t}(S, t) + O(\delta t^2).$$

Which, upon rearranging, is

$t - \delta t \qquad t \qquad t + \delta t$

$$\frac{\partial V}{\partial t}(S, t) = \frac{V_i^k - V_i^{k+1}}{\delta t} + O(\delta t).$$

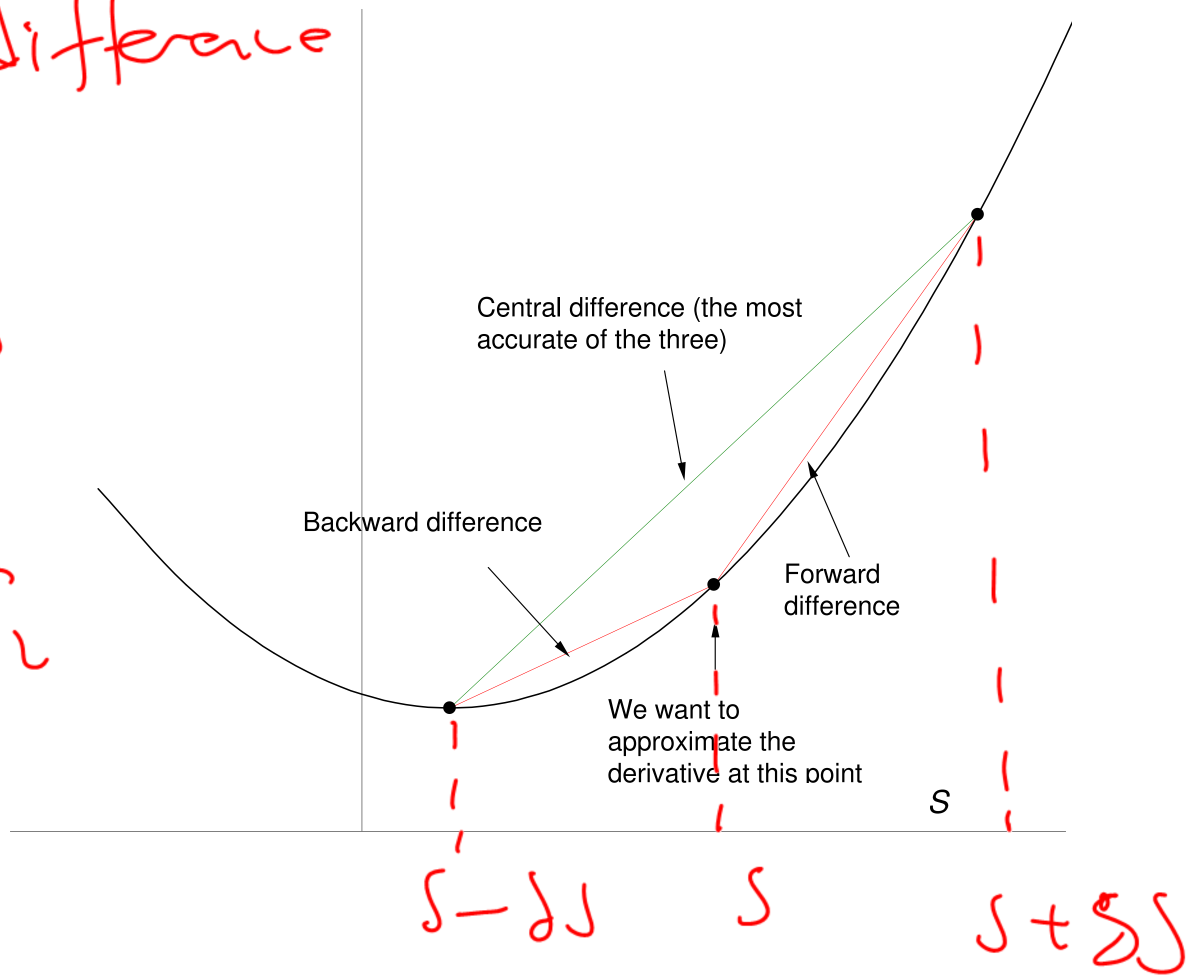
- The error is $O(\delta t)$.

$$\frac{\partial V}{\partial t} \sim \frac{V_i^k - V_i^{k+1}}{\delta t}$$

Approximating Δ

Examine a cross section of the grid at one of the time steps.

Central difference for $\dfrac{\partial V}{\partial S}$ and $\dfrac{\partial^2 V}{\partial S^2}$



These three approximations are

$$\frac{V_{i+1}^k - V_i^k}{\delta S},$$

$$\frac{V_i^k - V_{i-1}^k}{\delta S}$$

$$\text{and} \quad \frac{V_{i+1}^k - V_{i-1}^k}{2\,\delta S}.$$

These are called a **forward difference**, a **backward difference** and a **central difference** respectively.

One of these approximations is better than the others.

From a Taylor series expansion of the option value about the point $S + \delta S$, t we have

$$V(S+\delta S,t) = V(S,t) + \delta S \frac{\partial V}{\partial S}(S,t) + \tfrac{1}{2}\delta S^2 \frac{\partial^2 V}{\partial S^2}(S,t) + O(\delta S^3).$$

Similarly,

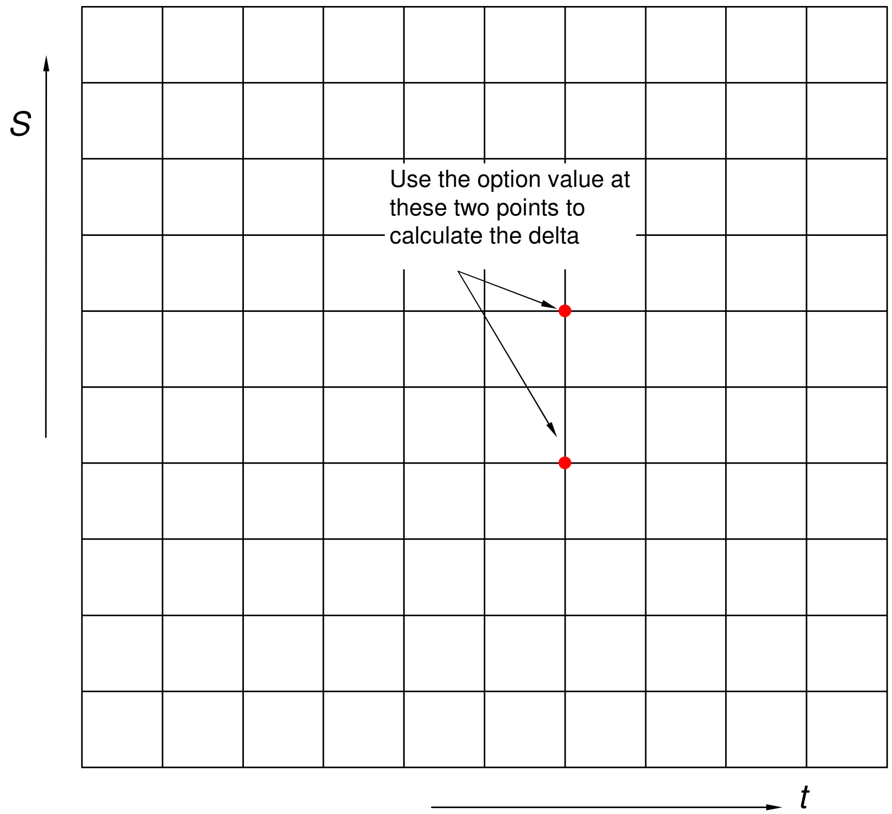
$$V(S-\delta S,t) = V(S,t) - \delta S \frac{\partial V}{\partial S}(S,t) + \tfrac{1}{2}\delta S^2 \frac{\partial^2 V}{\partial S^2}(S,t) + O(\delta S^3).$$

From these we get

$$\frac{\partial V}{\partial S}(S,t) = \frac{V_{i+1}^k - V_{i-1}^k}{2\,\delta S} + O(\delta S^2).$$

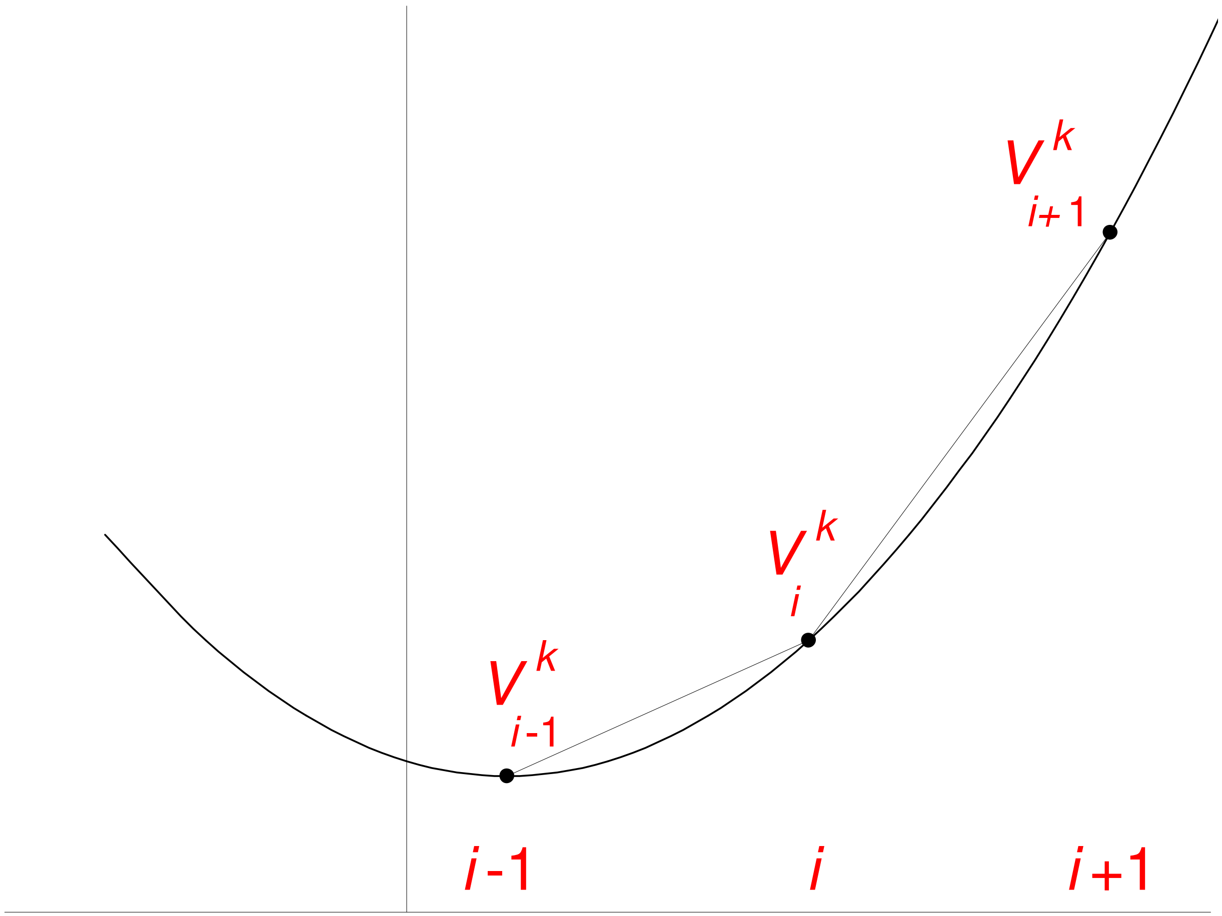
- The central difference has an error of $O(\delta S^2)$, the error in the forward and backward differences are both much larger, $O(\delta S)$.

The central difference calculated at S requires knowledge of the option value at $S + \delta S$ and $S - \delta S$.


S
Use the option value at
these two points to
calculate the delta
t

Approximating Γ

Gamma is the sensitivity of the delta to the underlying.



Calculate the delta half way between i and $i+1$, and the delta half way between $i-1$ and i . . . and difference them!

$$\text{Forward difference } = \frac{V_{i+1}^k - V_i^k}{\delta S}.$$

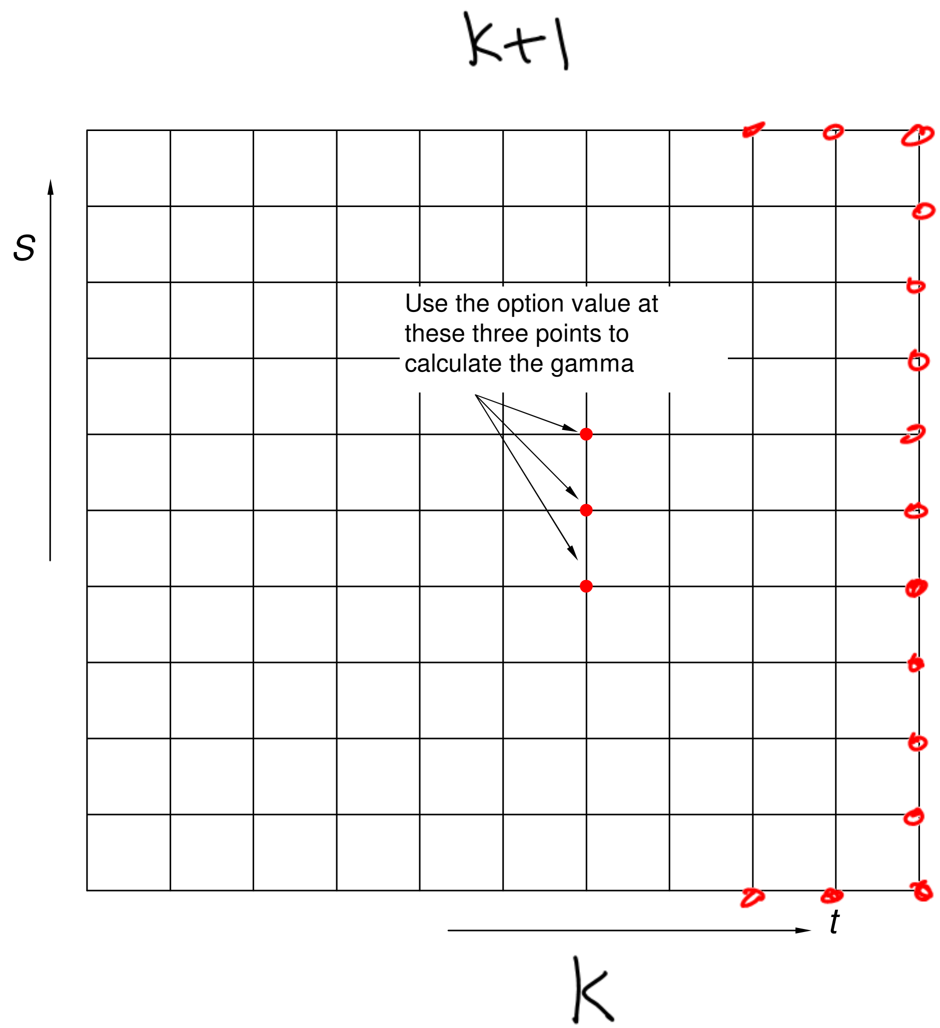
$$\text{Backward difference } = \frac{V_i^k - V_{i-1}^k}{\delta S}.$$

Therefore the natural approximation for the gamma is

$$\frac{\partial^2 V}{\partial S^2}(S,t) \approx \frac{\frac{V_{i+1}^k - V_i^k}{\delta S} - \frac{V_i^k - V_{i-1}^k}{\delta S}}{\delta S}$$

$$= \frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}.$$

The error in this approximation is also $O(\delta S^2)$.


k+1
S
Use the option value at
these three points to
calculate the gamma
t
k

Final conditions and payoffs

We know that at expiry the option value is just the payoff function. At expiry we have

$$V(S,T) = \text{Payoff}(S)$$

or, in our finite-difference notation,

$$V_i^0 = \text{Payoff}(i\,\delta S).$$

The right-hand side is a known function.

For example, if we are pricing a call option we have

$$V_i^0 = \max(i\,\delta S - E, 0).$$

This final condition will get our finite-difference scheme started.

The explicit finite-difference method

The Black–Scholes equation is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Write this as

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

so that we can examine more general problems.

$(r-D)S$

Using the above approximations

$$\frac{V_i^k - V_i^{k+1}}{\delta t} + \cancel{a_i^k}\,\frac{1}{2}\sigma^2 S^2 \left(\frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}\right) + b_i^k \left(\frac{V_{i+1}^k - V_{i-1}^k}{2\delta S}\right)$$

$$+c_i^k V_i^k = O(\delta t, \delta S^2).$$

$-r$

This can be rearranged...

$$\frac{V_n^m - V_n^{m+1}}{\delta t} + \frac{1}{2}\sigma^2 n^2 \left(V_{n-1}^m - 2V_n^m + V_{n+1}^m\right) + \frac{(r-D)n}{2}\left(V_{n+1}^m - V_{n-1}^m\right) - rV_n^m = 0$$

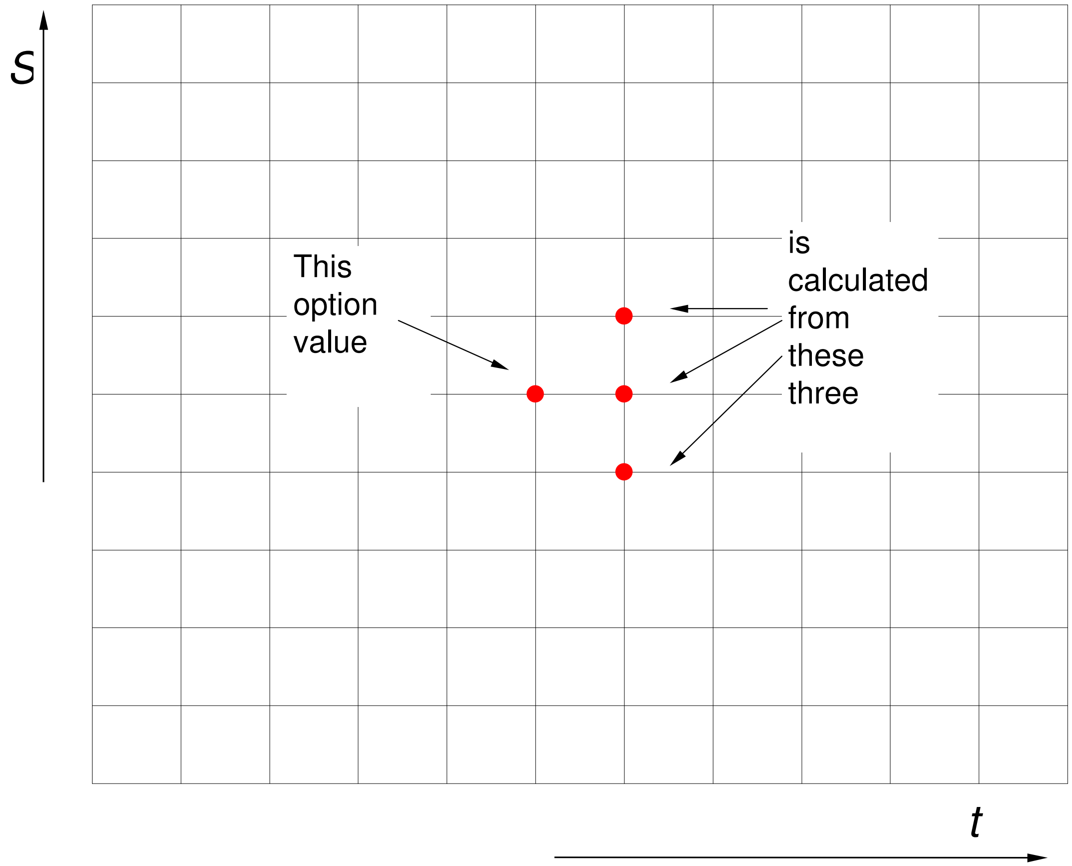
$$V_n^{m+1} = V_n^m + \frac{1}{2}\sigma^2 n^2 \left(V_{n-1}^m - 2V_n^m + V_{n+1}^m\right)\delta t + \frac{(r-D)n}{2}\left(V_{n+1}^m - V_{n-1}^m\right)\delta t - r\,\delta t\,V_n^m$$

$$V_i^{k+1} = \cdots V_{i+1}^k + \cdots V_i^k + \cdots V_{i-1}^k.$$

This is an equation for V_i^{k+1} given three option values at time k.

(That's why this is called the **explicit finite-difference method**.)

The relationship between the option values in the algorithm is shown in the figure below.



Points to note:

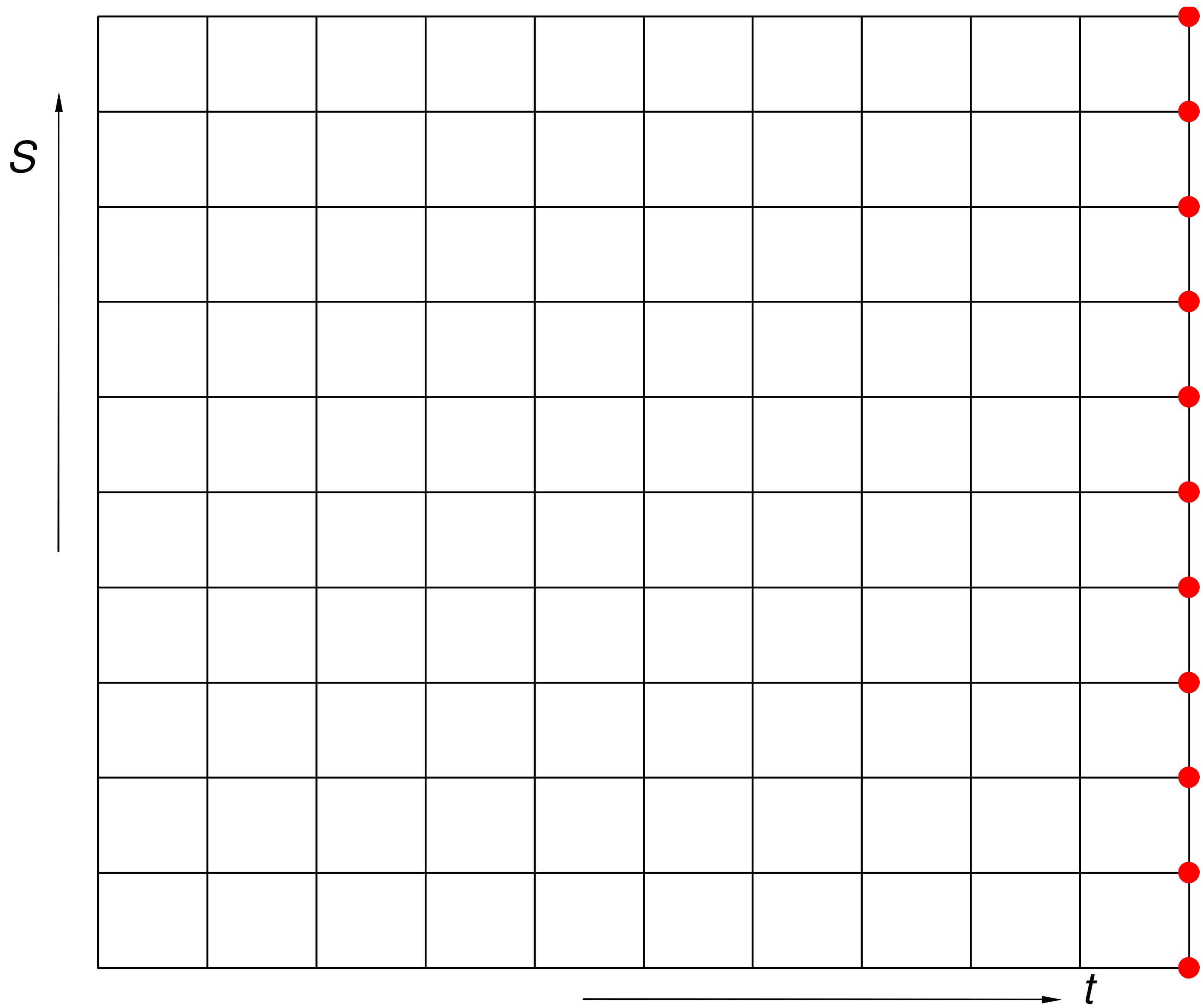
- The time derivative uses the option values at 'times' k and $k+1$, whereas the other terms all use values at k.

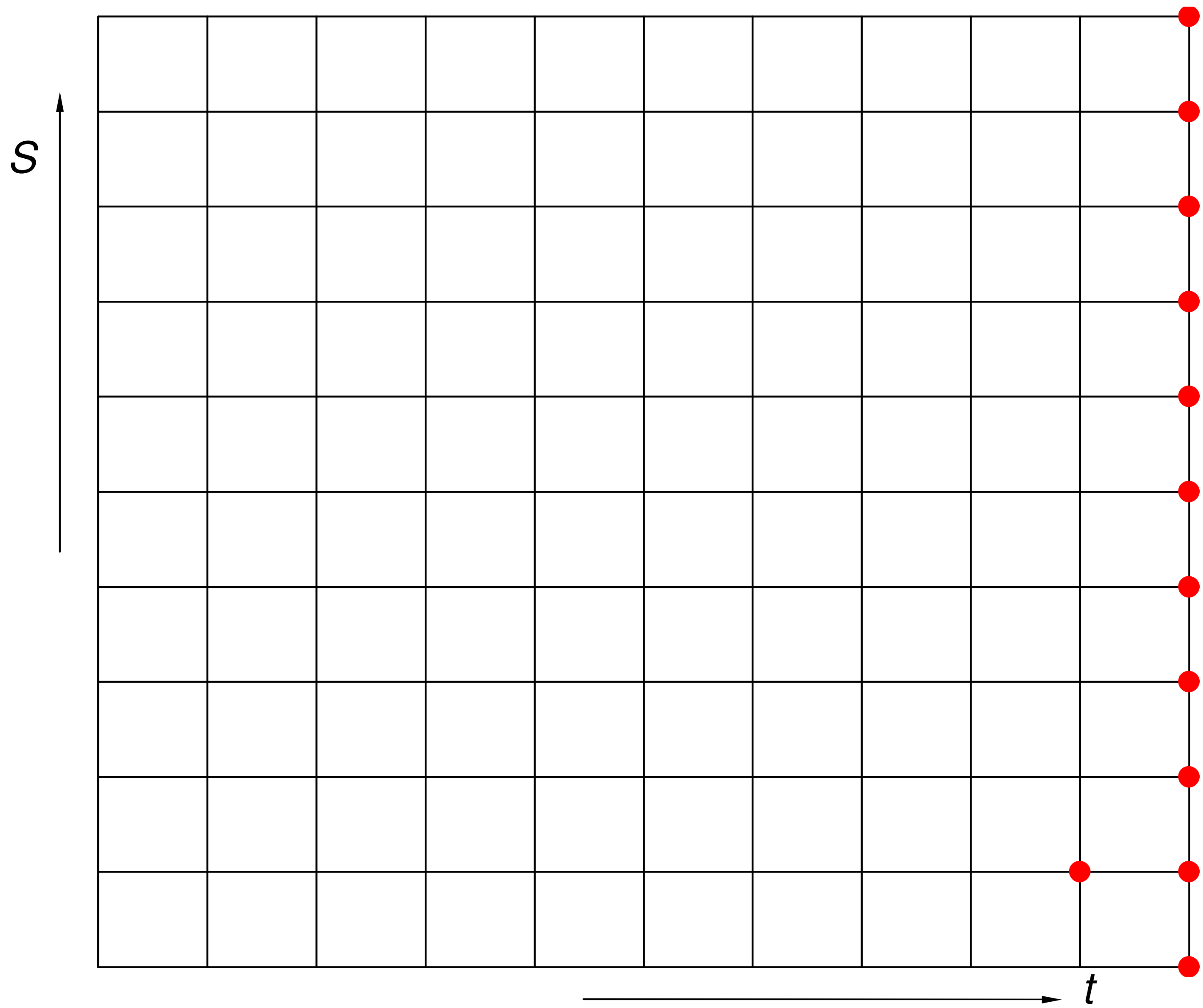
- The gamma term is a central difference, in practice one never uses anything else.

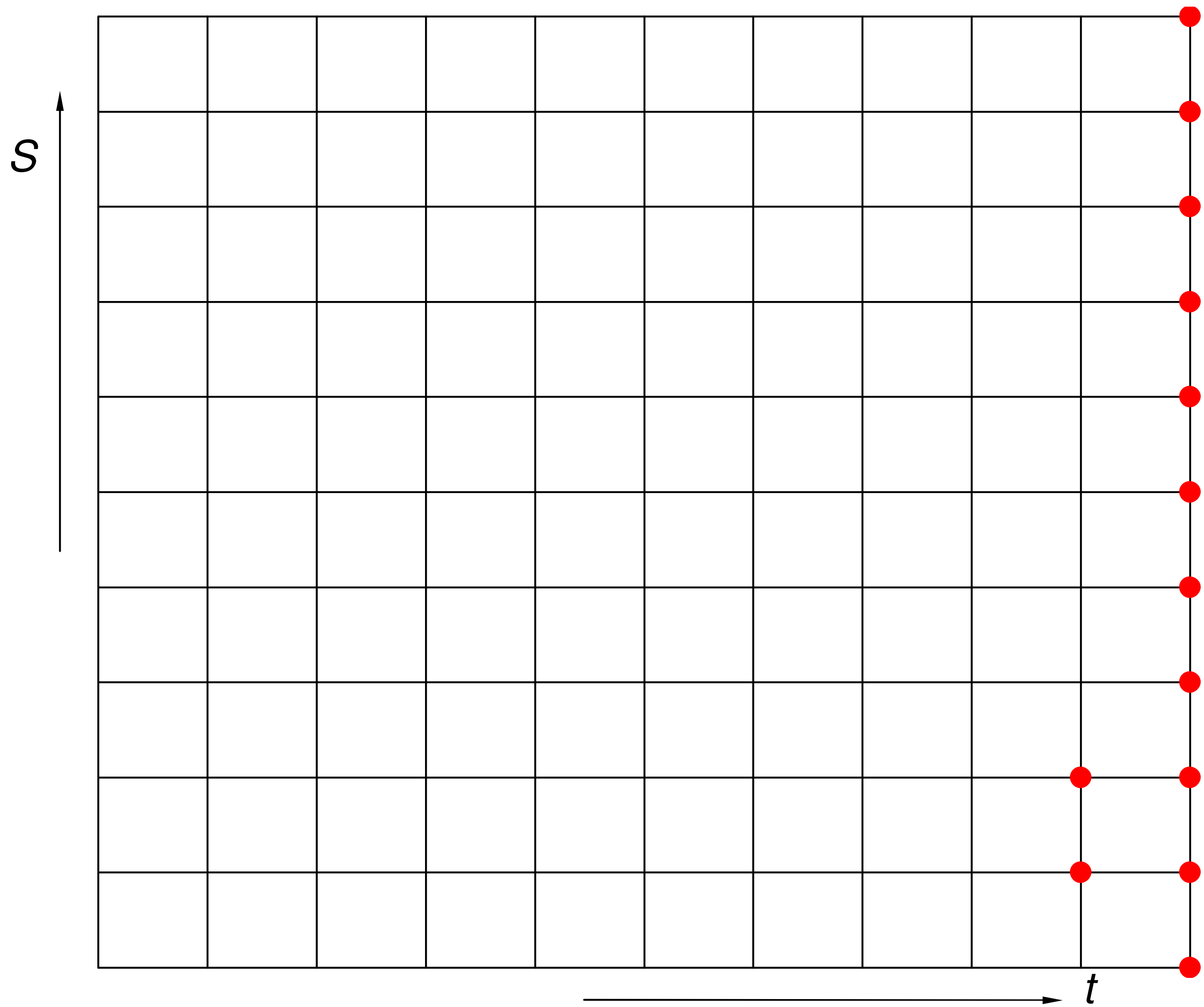
- The delta term uses a central difference. There are often times when a one-sided derivative is better. We'll see examples later.

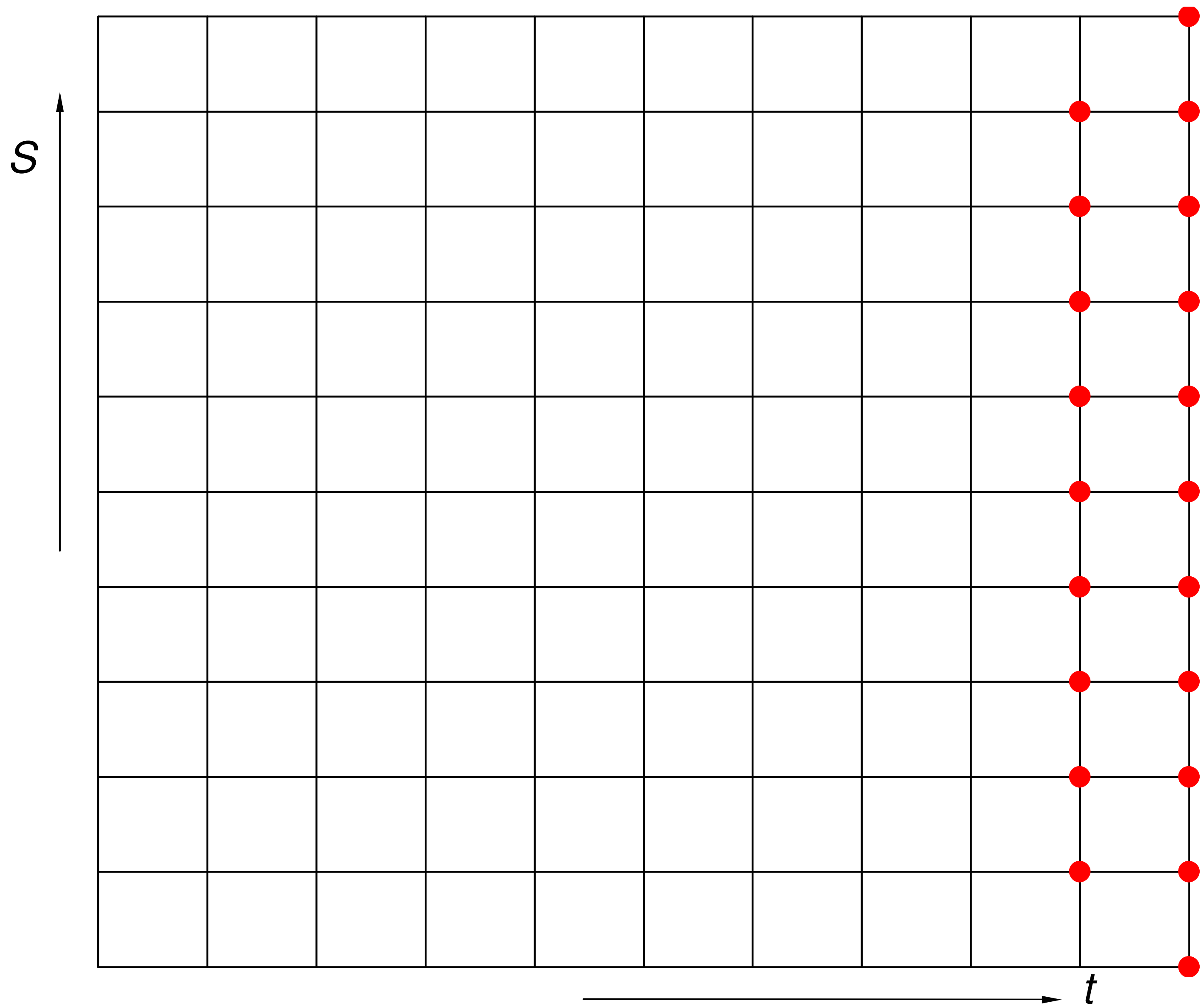
- The asset- and time-dependent functions a, b and c have been valued at $S_i = i\,\delta S$ and $t = T - k\,\delta t$ with the obvious notation.

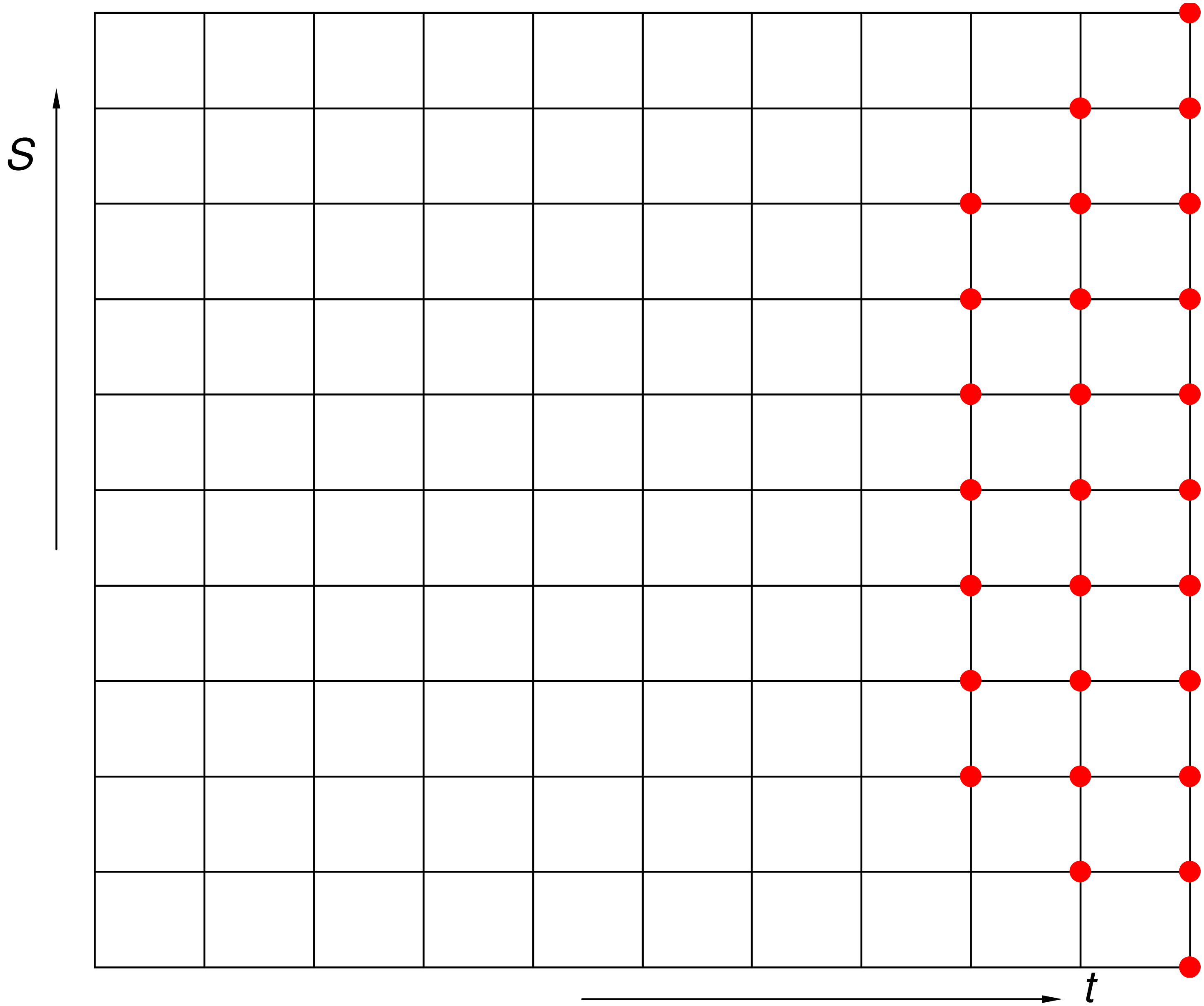
- The error in the equation is $O(\delta t, \delta S^2)$.

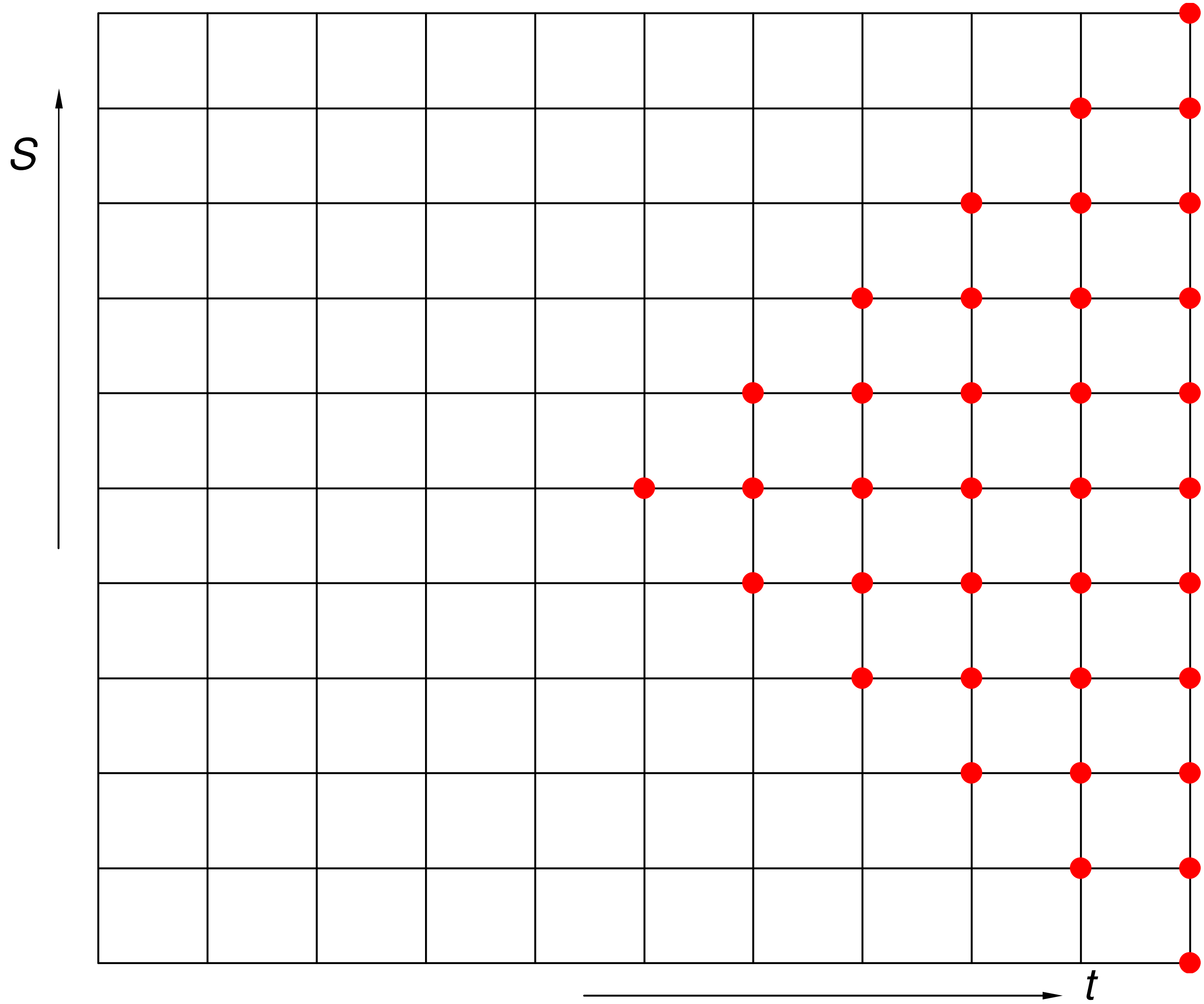

S
t


S
t


S
t


S
t


S
t


S
t

Boundary conditions

We must specify the option values at the extremes of the region, at $S = 0$ and at $S = I\,\delta S$. They will depend on our option.

Example 1: Call option at $S = 0$

At $S = 0$ we know that the value is always zero, therefore

$$V_0^k = 0.$$

Example 2: Call option for large S

For large S the call value asymptotes to $S - Ee^{-r(T-t)}$. Thus

$$V_I^k = I\,\delta S - Ee^{-rk\,\delta t}.$$

Example 3: Put option at $S = 0$

At $S = 0$ $V = Ee^{-r(T-t)}$. I.e.

$$V_0^k = Ee^{-rk\,\delta t}.$$

Example 4: Put option for large S

The put option becomes worthless for large S and so

$$V_I^k = 0.$$

Example 5*: General condition at $S = 0$

A useful boundary condition to apply at $S = 0$ is that the diffusion and drift terms 'switch off.'

$$\frac{\partial V}{\partial t}(0,t) - rV(0,t) = 0 \qquad V_0^k = (1 - r\delta t)V_0^{k-1}.$$

When $S=0$ B.S.E

$$\frac{\partial V}{\partial t} - rV = 0 \qquad \frac{\partial V}{\partial t} = rV$$

$$V_0^k - V_0^{k+1} = r\delta t\, V_0^k$$

Example 6*: General condition at infinity Vanishing Γ.

When the option has a payoff that is linear in the underlying for large S then

$S \to \infty$ $\quad \frac{\partial V}{\partial S} = \Delta(t)$

$$\frac{\partial^2 V}{\partial S^2}(S,t) \to 0 \quad \text{as } S \to \infty.$$

$\Gamma = \frac{\partial}{\partial S}\Delta(t) = 0$

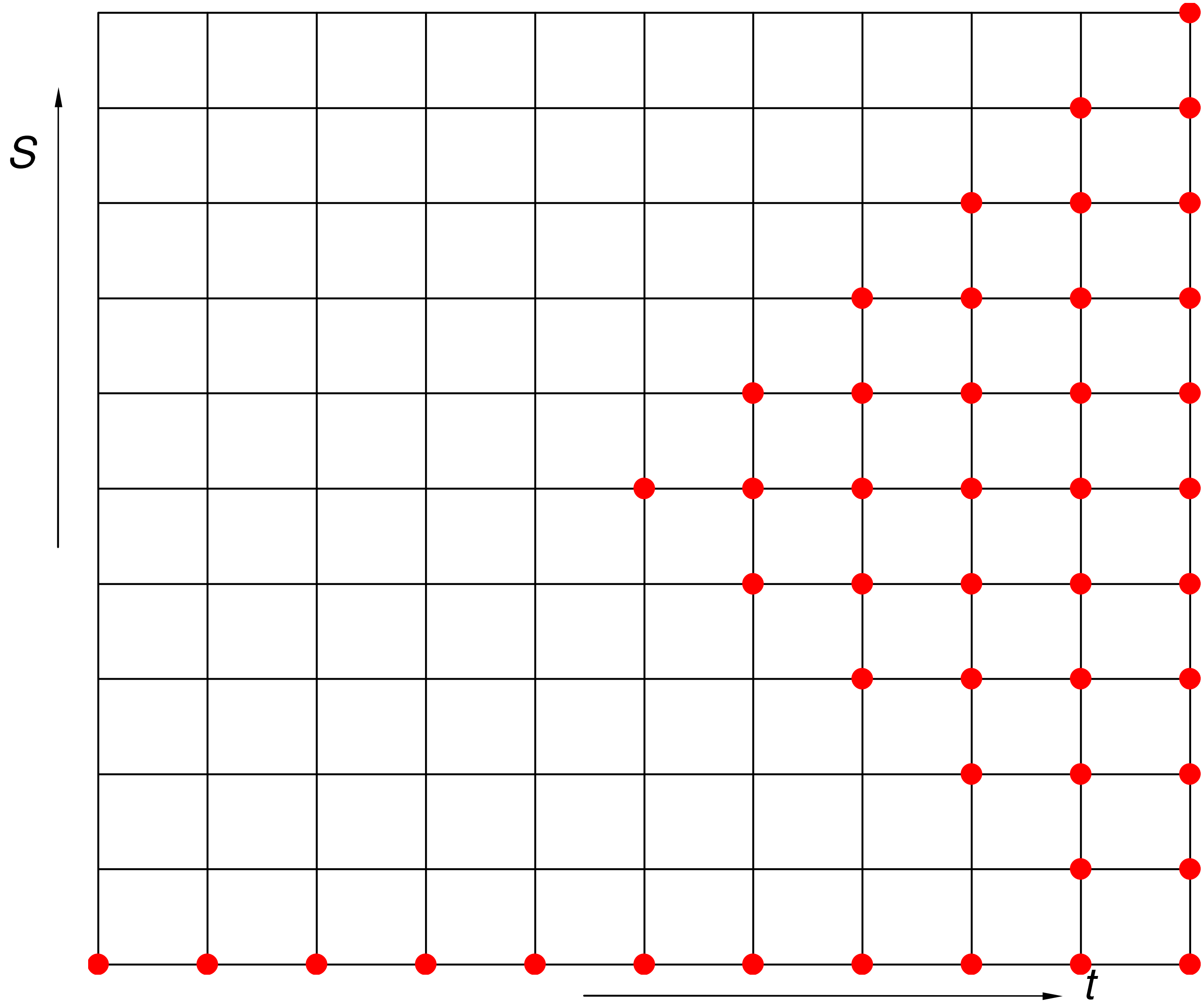
The finite-difference representation is

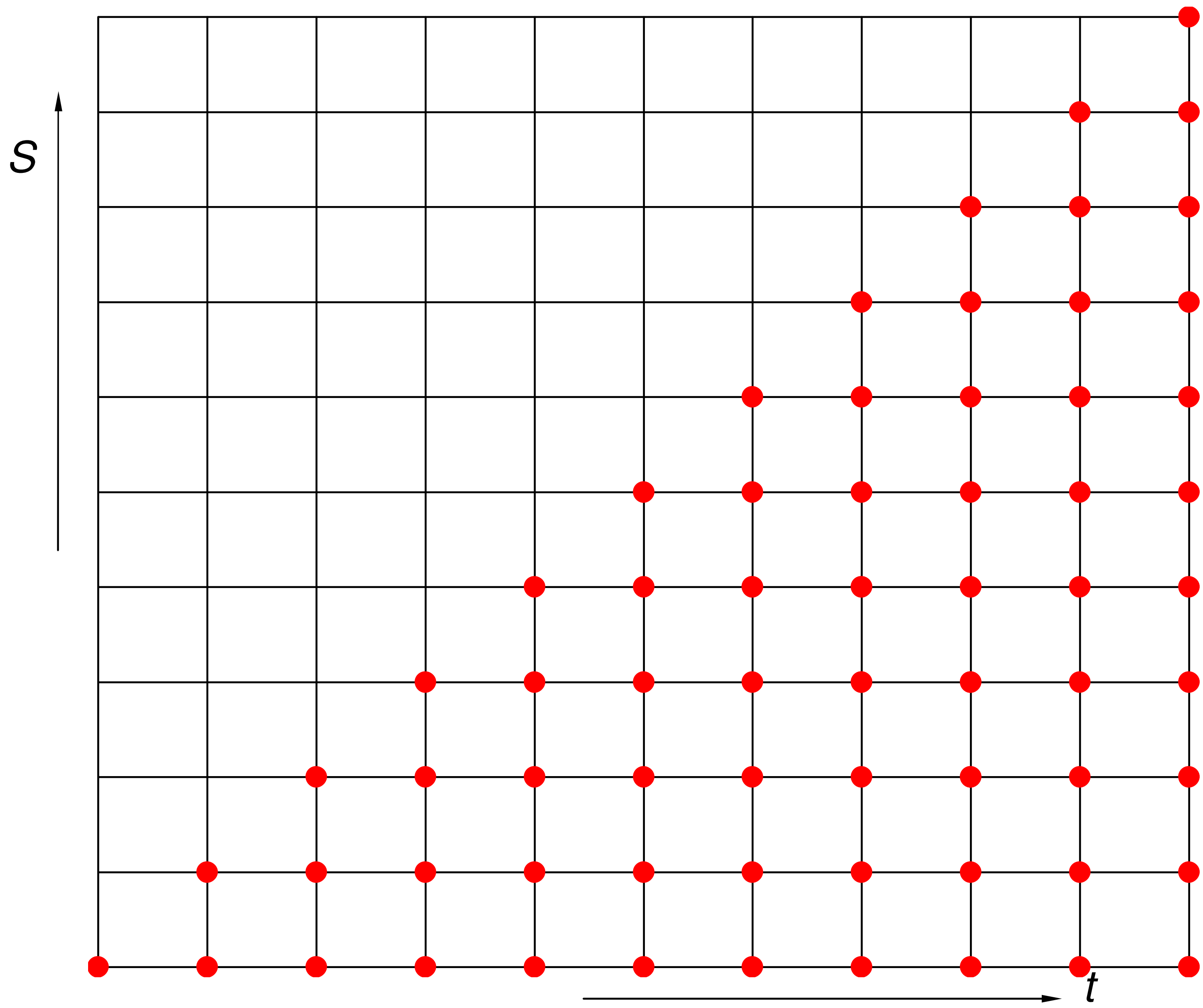
$$V_I^k = 2V_{I-1}^k - V_{I-2}^k.$$

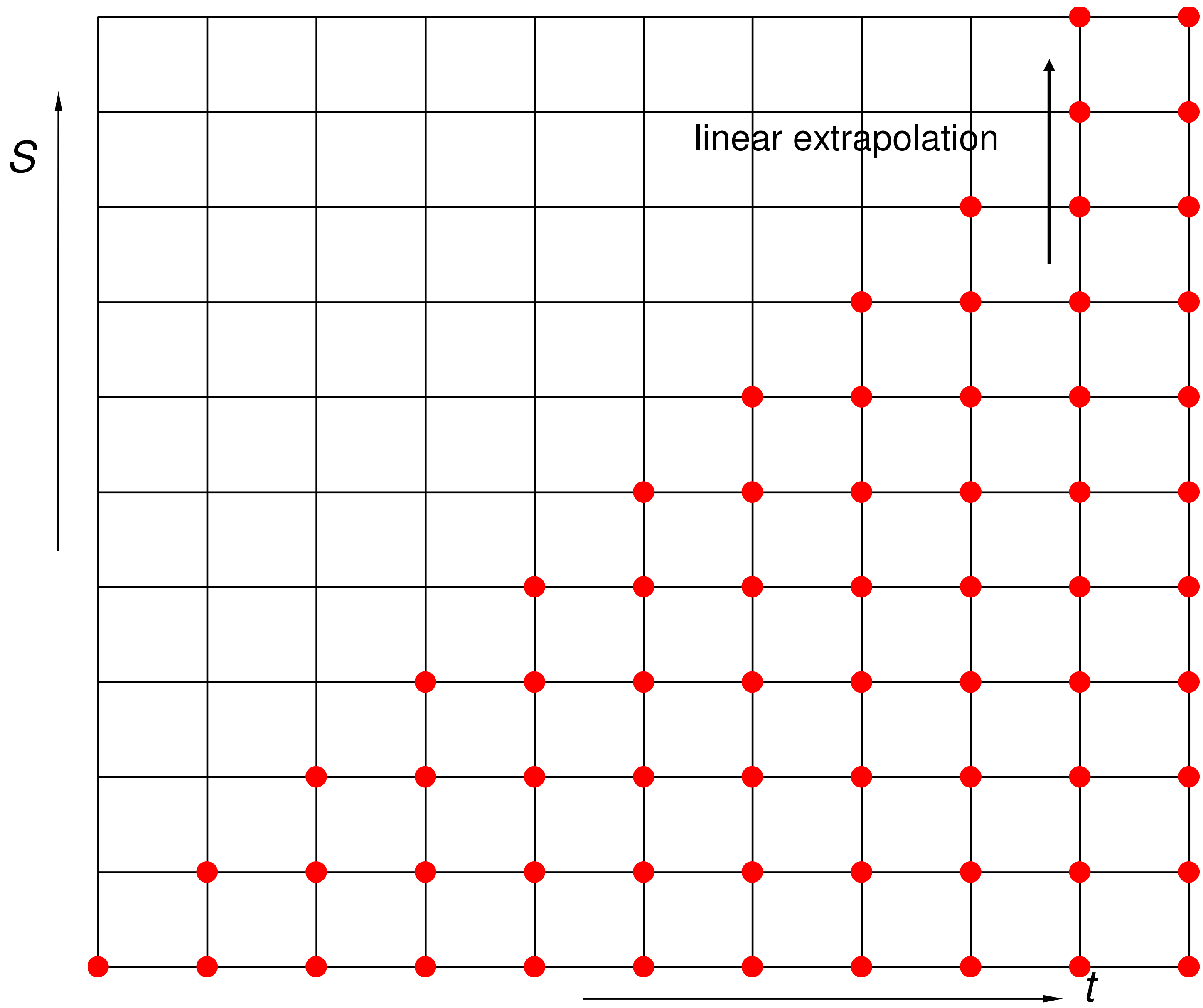
$\Gamma = \frac{\partial^2 V}{\partial S^2} = \frac{V^m_{n-1} - 2V^m_n + V^m_{n+1}}{\delta S^2} \approx 0$ at $n = N$

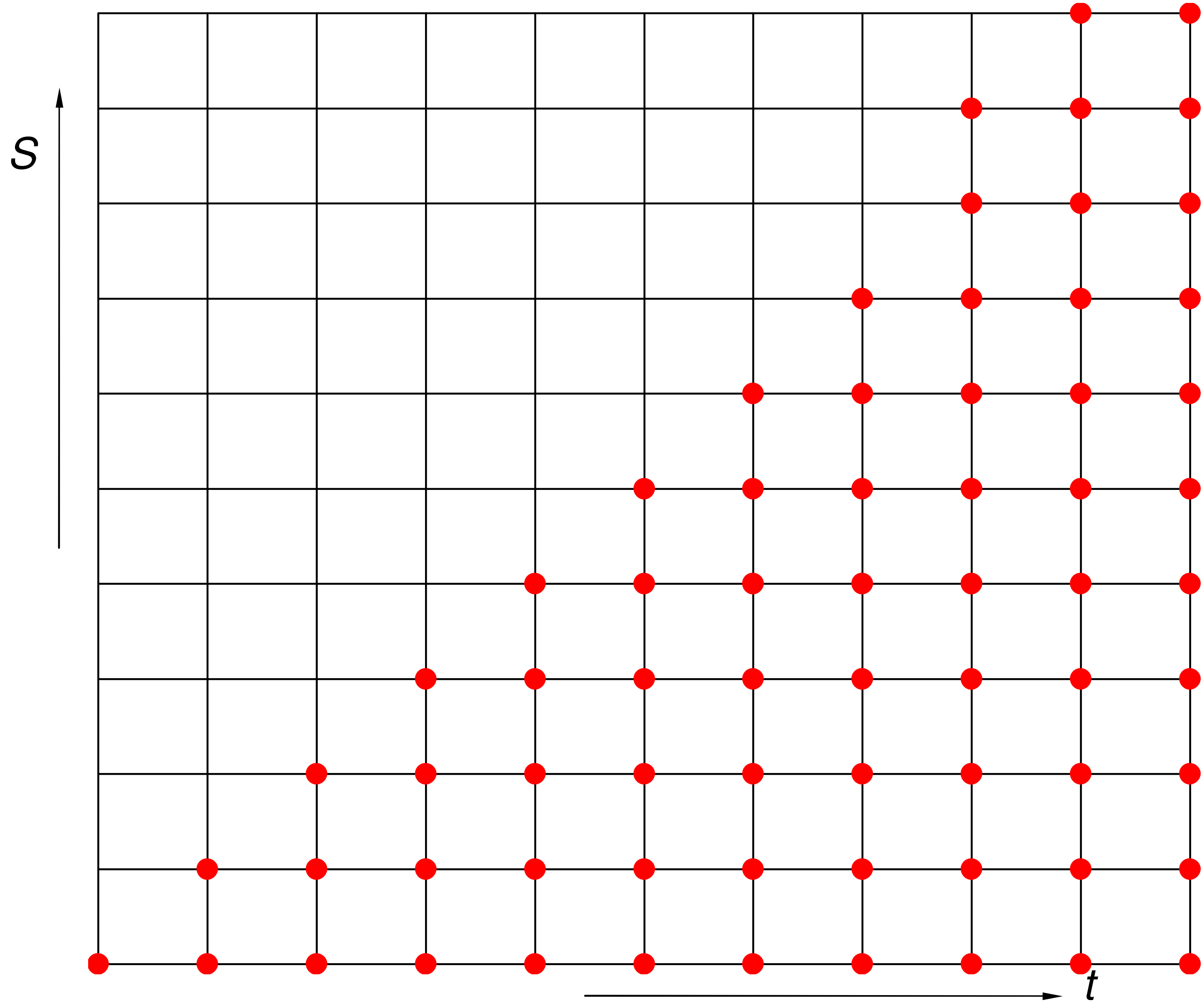
$V^m_{N-1} - 2V^m_N + V^m_{N+1} \approx 0$

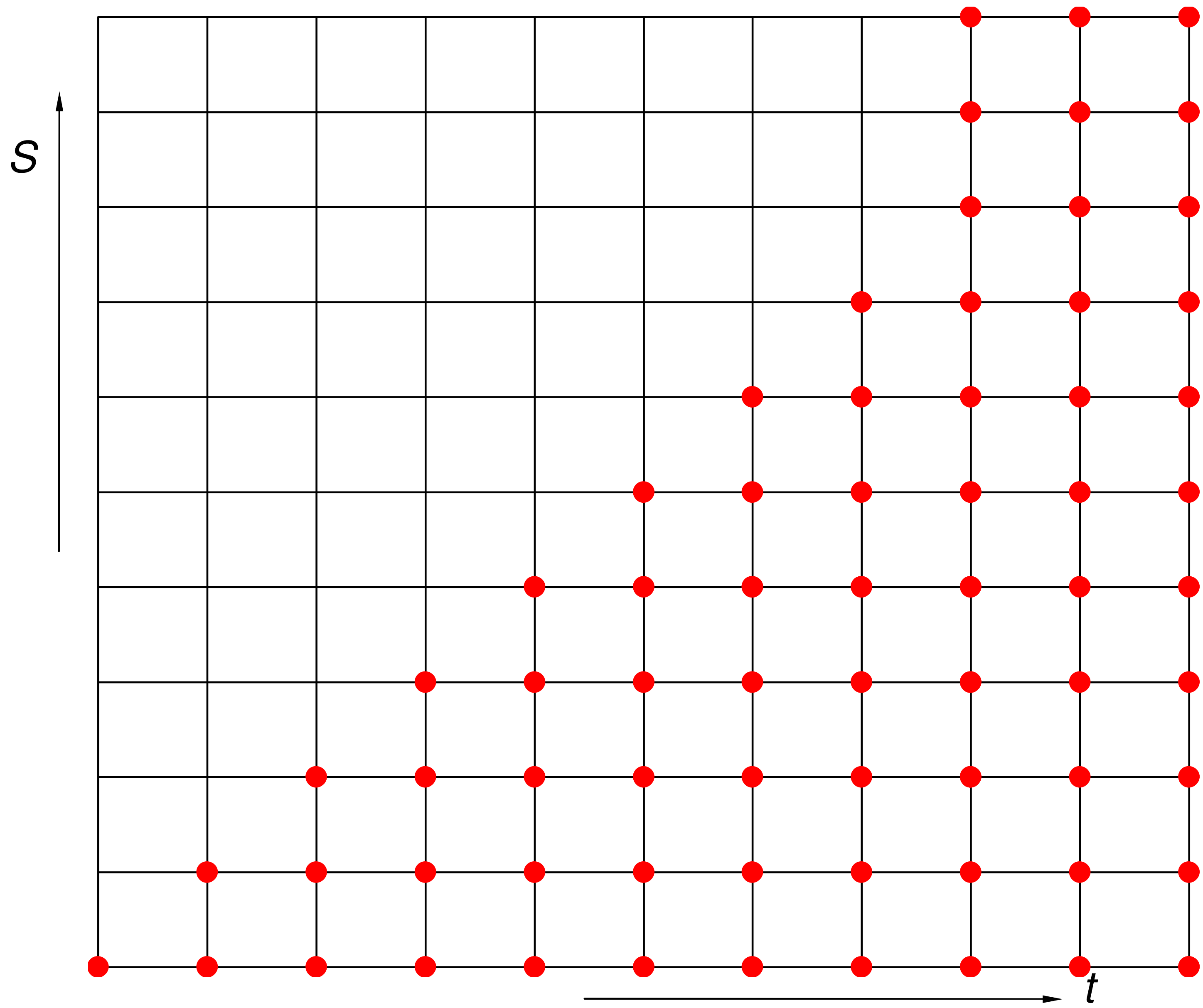
$V^m_{N+1} = 2V^m_N - V^m_{N-1}$

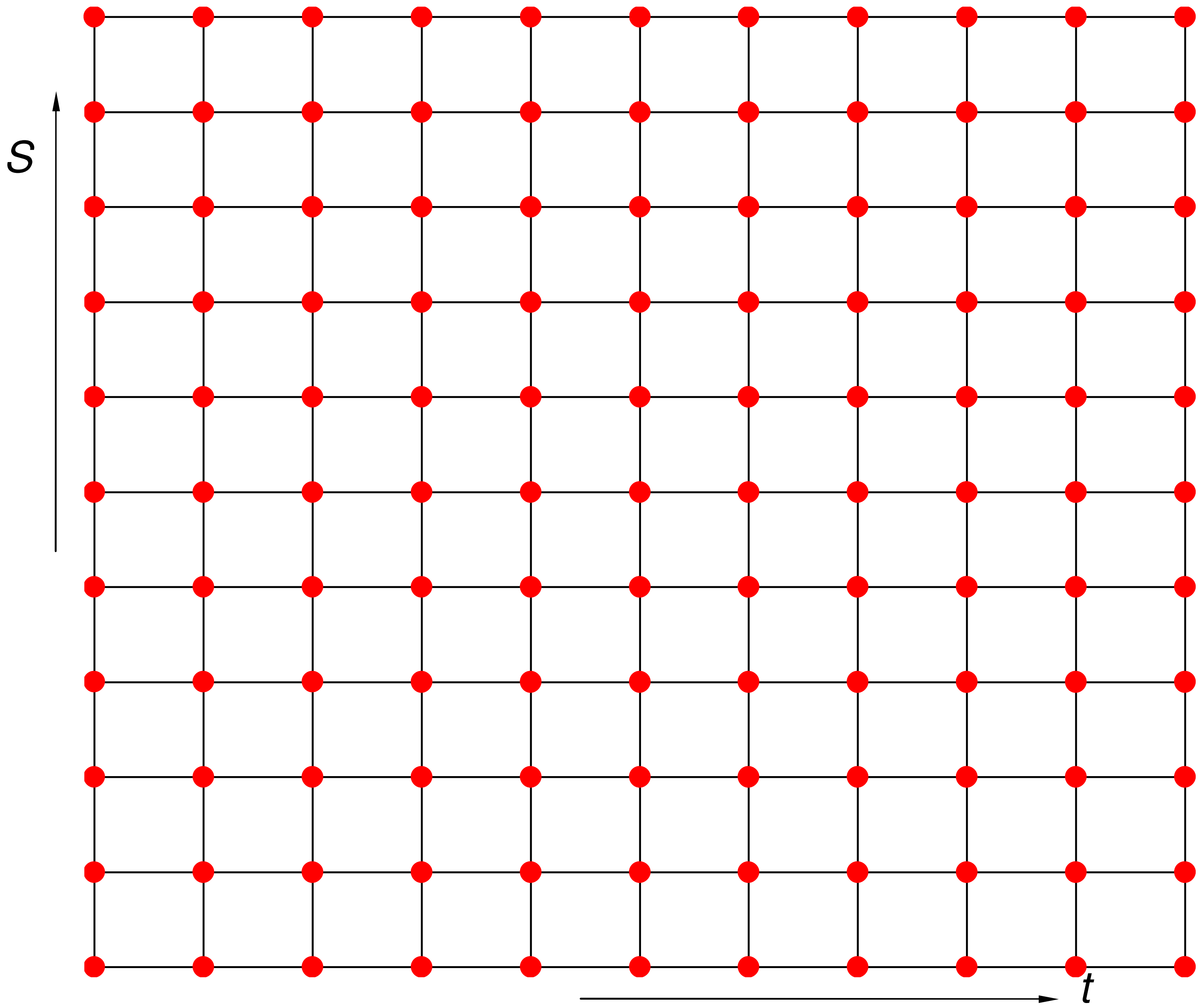

S
t


S
t


linear extrapolation
S
t


S
t


S
t


S
t

Accuracy and computational time

Again let's use ϵ to represent the desired accuracy in a calculation.

We know that errors are $O(\delta t)$ and $O(\delta S^2)$. It makes sense to have errors due to the time step and to the finite number of simulations to be of the same order. So we would choose:

$$\delta t = O(\epsilon) \quad \text{and} \quad \delta S = O(\epsilon^{1/2}).$$

The time taken is then proportional to number of calculations, therefore

$$\text{Time taken} = O(\epsilon^{-3/2}).$$

In higher dimensions...

Suppose you have a basket option with D underlyings. The time taken now becomes

$$\text{Time taken} = O(\epsilon^{-1-D/2}).$$

This is very sensitive to dimension!

Other issues

- Greeks

- Early exercise (and other decisions)

The advantages of the explicit method

- It is very easy to program and hard to make mistakes
- When it does go unstable it is usually obvious
- It copes well with coefficients that are asset and/or time dependent
- it copes very well with early exercise
- It can be used for modern option-pricing models

The disadvantages of the explicit method

- There are restrictions on the time step. The method is conditionally stable.

- It is slower than Monte Carlo in high dimensions

impossible

Summary

Please take away the following important ideas

- There are two main numerical methods for pricing derivatives

- Monte Carlo methods exploit the relationship between option prices and expectations

- The finite-difference method solved a discretized version of the Black–Scholes equation

Exotic Options

In this lecture. . .

- the names and contract details for many basic types of exotic options
- how to classify exotic options according to important features
- how to think about derivatives in a way that makes it easy to compare and contrast different contracts
- pricing exotics using Monte Carlo simulation
- pricing exotics via partial differential equations

By the end of this lecture you will be able to

- characterize most exotic contracts according to a list of important features

- price exotics using Monte Carlo simulations

- interpret the pricing of many exotics in terms of partial differential equations

Introduction

Exotic contracts are traded **over the counter** (**OTC**), meaning that they are designed by the relevant counterparties and are not available as exchange-traded contracts.

Exotic options include contracts with features making them more complex to price and to hedge than vanillas.

Often one takes volatilities ‘implied’ by the market prices of vanillas and put them into the pricing model for exotics.

Important features to look out for

- Time dependence
- Cashflows
- Path dependence
- Dimensionality
- Order

Bermudan options

It is common for contracts that allow early exercise to permit the exercise only at certain specified times, and not at *all* times before expiry.

For example, exercise may only be allowed on Thursdays between certain times. An option with such intermittent exercise opportunities is called a **Bermudan option**.

All that this means mathematically is that the constraint (**??**) is only 'switched on' at these early exercise dates.

1. Time dependence

$V = V(S, t)$

Here we are concerned with time dependence in the option contract.

For example, discrete cashflows necessarily involve time dependence.

Another example, early exercise might only be permitted on certain dates or during certain periods. This intermittent early exercise is a characteristic of **Bermudan options**.

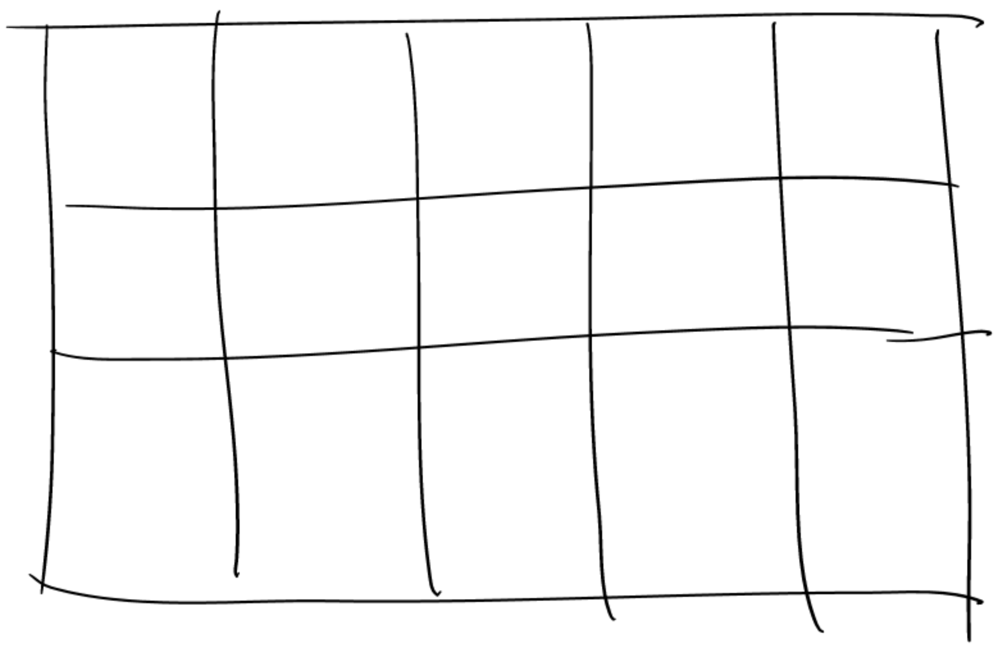
Similarly, the position of the barrier in a knock-out option may change with time. Every month it may be reset at a higher level than the month before.

- These contracts are referred to as **time inhomogeneous**.

When there is time dependence in a contract we might expect

- jumps in option values and/or the greeks

- to have to worry about time discretization in numerical schemes



2. Cashflows

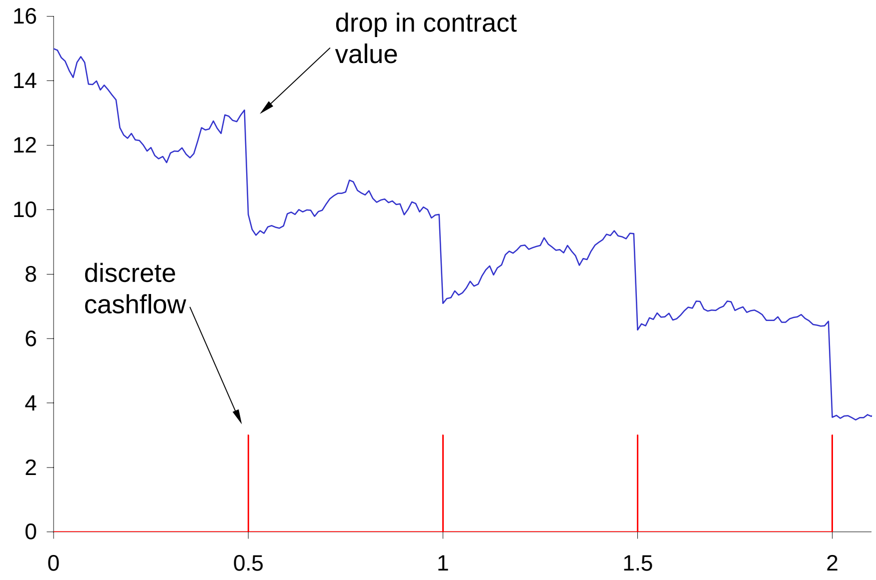
Imagine a contract that pays the holder an amount q at time t_i. The contract could be a bond and the payment a coupon.

If we use $V(t)$ to denote the contract value and t_i^- and t_i^+ to denote just before and just after the cashflow date then simple arbitrage considerations lead to

- $$V(t_i^-) = V(t_i^+) + q.$$

This is a **jump condition**.

The value of the contract jumps by the amount of the cashflow. The behavior of the contract value across the payment date is shown in the figure.



A discrete cashflow and its effect on a contract value.

If the contract is contingent on an underlying variable so that we have $V(S,t)$ then we can accommodate cashflows that depend on the level of the asset S i.e. we could have $q(S)$.

That's an example of a **discrete cashflow**.

Some contracts specify **continuous cashflows**. There may be a payment every day.

When there are cashflows we expect

- option values to jump
- the greeks to jump

3. Path dependence

Path-dependent contracts have payoffs, and therefore values, that depend on the history of the asset price path.

An asset starts at A and ends at Z at expiration. If the contract is path dependent the route taken from A to Z matters. If it is not path dependent then the route does not matter.

Path dependence comes in two main forms:

- Strong path dependence

- Weak path dependence

Strong path dependence

Of particular interest, mathematical and practical, are the **strongly path-dependent contracts**.

These have payoffs that depend on some property of the asset price path in addition to the value of the underlying at the present moment in time; in the equity option language, we cannot write the value as $V(S,t)$.

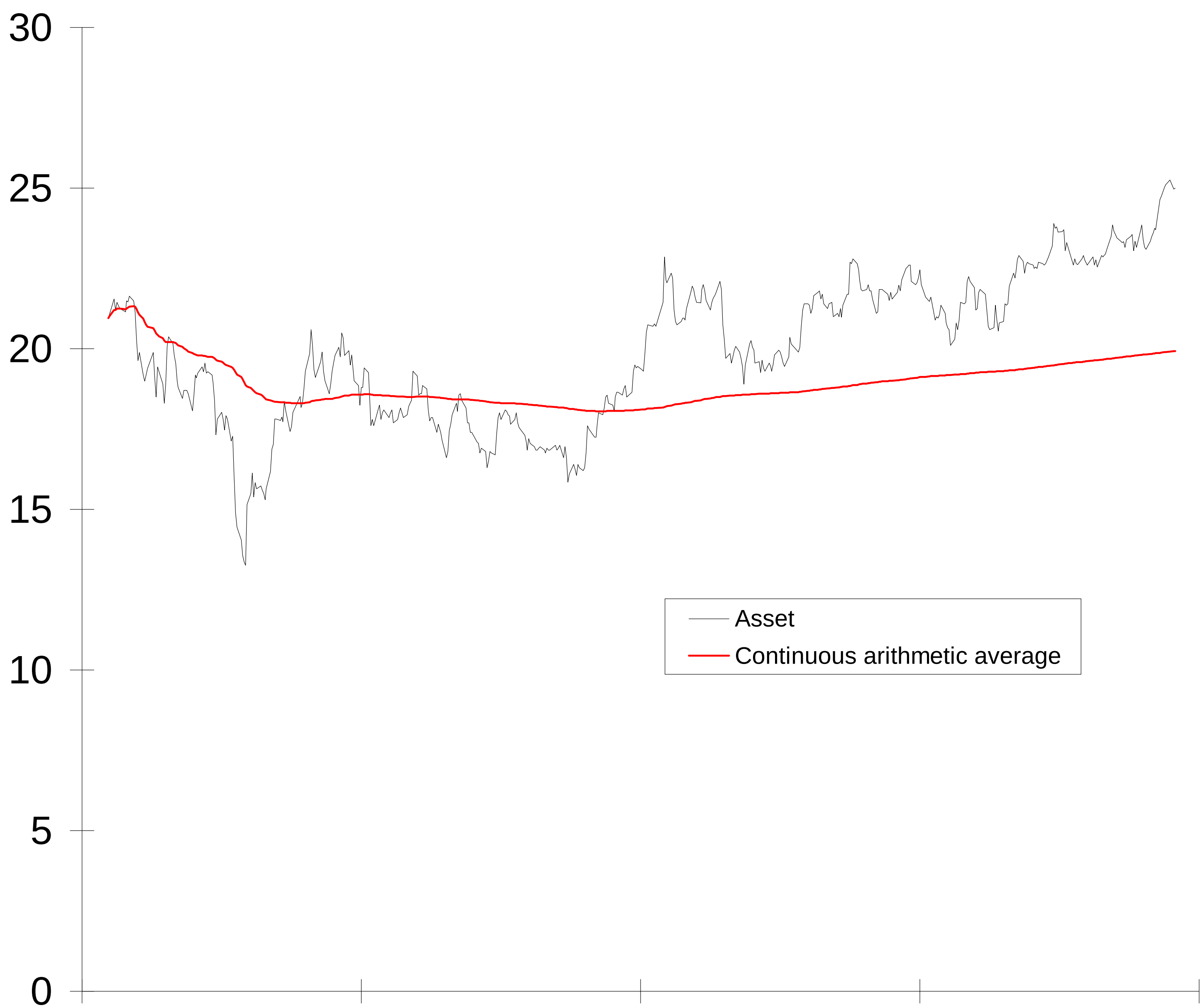
- The contract value is a function of at least one more independent variable.

$V = V(S, I, t)$ I – sampling var.

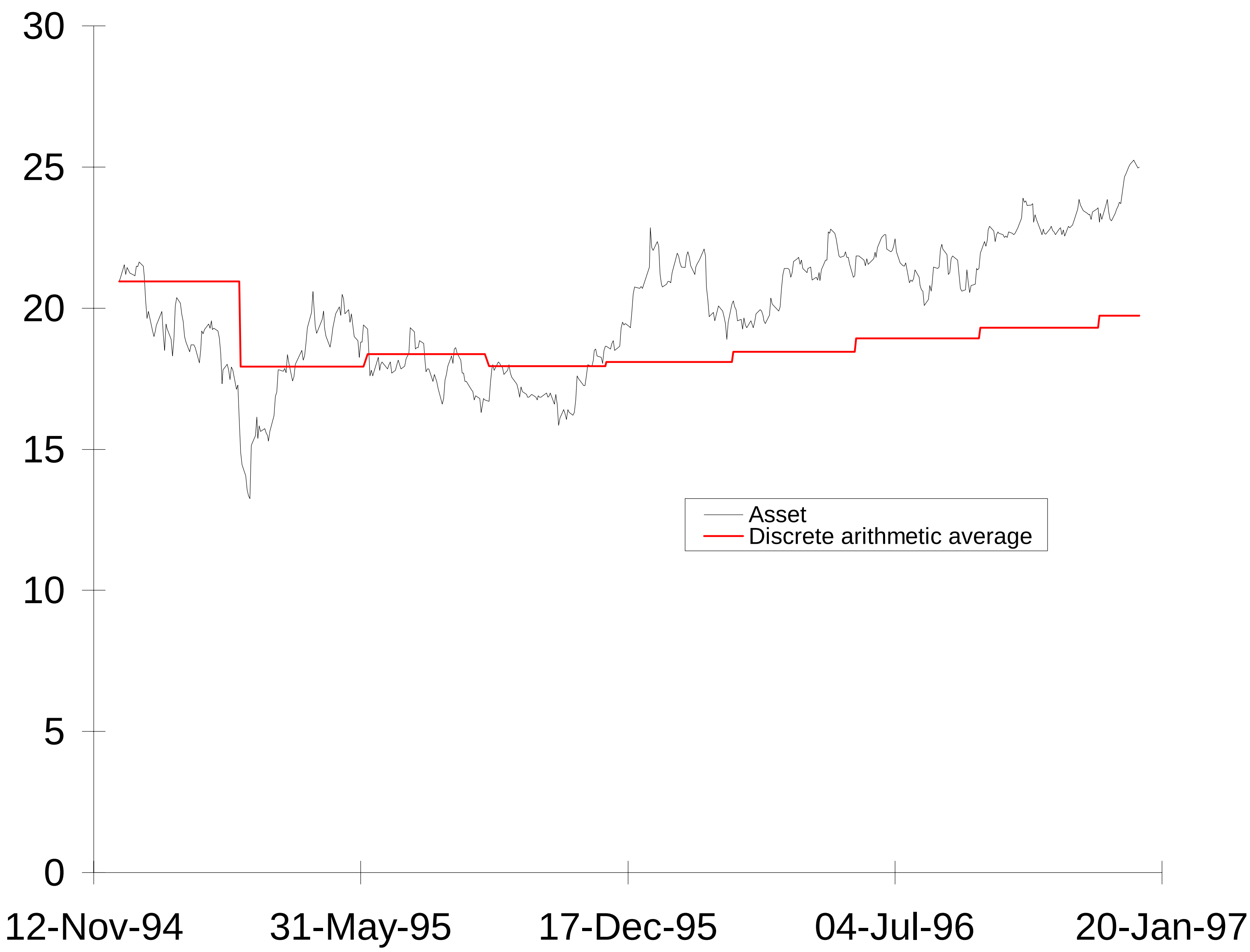
Example:

The Asian option has a payoff that depends on the average value of the underlying asset from inception to expiry. We must keep track of more information about the asset price path than simply its present position.

The extra information that we need is contained in the 'running average.' This is the average of the asset price from inception until the present, when we are valuing the option.


30
25
20
15
10
5
0
Asset
Continuous arithmetic average

Path dependency also comes in **discrete** and **continuous** varieties depending on whether the path-dependent quantity is **sampled** discretely or continuously.



30
25
20
15
10
5
0
12-Nov-94
31-May-95
17-Dec-95
04-Jul-96
20-Jan-97
Asset
Discrete arithmetic average

When there is strong path dependence in a contract we might expect

- to have to solve in higher dimensions

(We have to keep track of a new **state variable** such as the average to date.)

$$I(t) = \text{``}\frac{1}{t}\text{''} \int_0^t f(S, \tau)\, d\tau$$

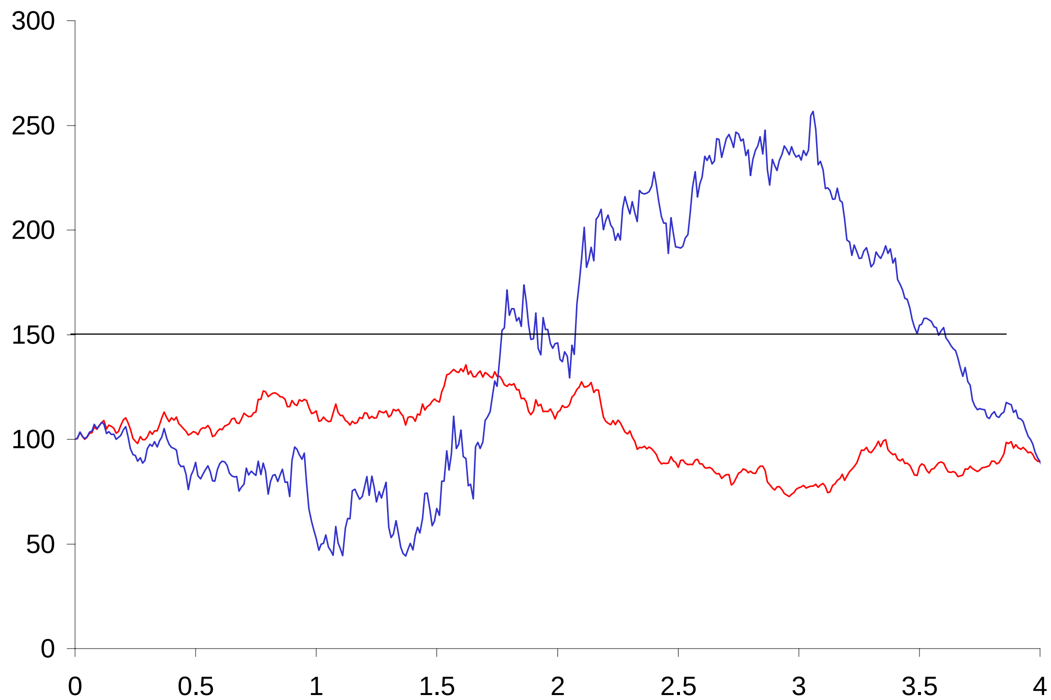
Weak path dependence

- Options whose value depends on the asset history, but can still be written as $V(S,t)$ are said to be **weakly path dependent**.

One of the most common reasons for weak path dependence in a contract is a **barrier**. Barrier (or knock-in, or knock-out) options are triggered by the action of the underlying hitting a prescribed value at some time before expiry.

For example, as long as the asset remains below 150, the contract will have a call payoff at expiry. However, should the asset reach this level before expiry then the option becomes worthless; the option has 'knocked out.'

$S_0 = 100$
$S_T = 150$



Two paths having the same value at expiry but with completely different payoffs.

- Weak path dependency does not add any extra dimensions.

(So a barrier option still only has two dimensions, S and t.)

4. Dimensionality

Dimensionality refers to the number of underlying independent variables.

- The vanilla option has two independent variables, S and t, and is thus two dimensional.

- The weakly path-dependent contracts have the same number of dimensions as their non-path-dependent cousins, i.e. a barrier call option has the same two dimensions as a vanilla call.

But some contracts require us to go in to extra dimensions!

There are two distinct reasons why we need more dimensions. . .

- More sources of randomness

- Strong path dependence

More dimensions caused by more sources of randomness

We will get higher dimensions if we have more sources of randomness

- If we have an option on **10** underlyings ('best of' for example) we will have **11** dimensions (S_1, S_2,. . . ,S_{10} and t)

But we will also get more dimensions if we have other types of randomness, such as volatility.

- If we have an option on **10** underlyings and we use a stochastic volatility model for each asset we will have **21** dimensions (S_1, S_2,. . . ,S_{10} and t, and also σ_1, σ_2,. . . ,σ_{10})

Each new dimensions introduces extra 'diffusion' terms. (What does this mean for the governing PDE?)

More dimensions caused by strong path dependency

We will get higher dimensions if we an option that is strongly path dependent.

- If we have an option that pays off the maximum of the average stock price we will have **4** dimensions (S and t, but also a state variable for the average and another for the maximum of the average!)

We'll see the theory of this later, but the effect on the governing PDE is to sometimes add new terms that are not diffusive! (But no new parameters!)

When the problem is of high dimensions we might expect

- to have restrictions on the kind of numerical solution we employ. The higher the number of dimensions, the more likely we are to want to use Monte Carlo simulations.

5. The order of an option

The basic, vanilla options are of first order. Their payoffs depend only on the underlying asset, the quantity that we are *directly* modeling. Other, path-dependent, contracts can still be of first order if the payoff only depends only on properties of the asset price path.

- **Higher order** refers to options whose payoff, and hence value, is contingent on the value of *another* option.

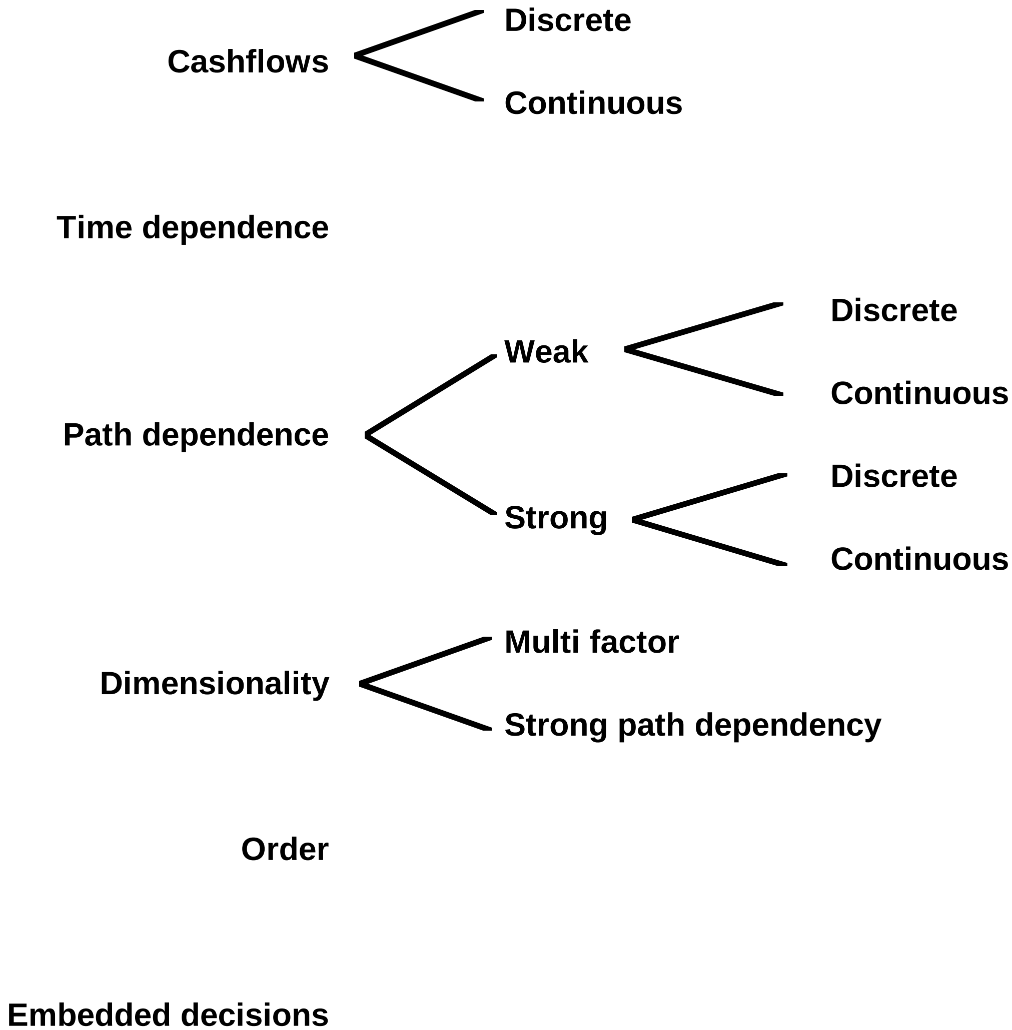
The obvious second-order options are compound options, for example, a call option giving the holder the right to buy a put option. The compound option expires at some date T_1 and the option on which it is contingent, expires at a later time T_2. Technically speaking, such an option is weakly path dependent.

From a practical point of view, the compound option raises some important modeling issues.

- The payoff for the compound option depends on the *market* value of the underlying option, and not on the theoretical price.

If you hold a compound option, and want to exercise the first option then you must take possession of the underlying option. High order option values are very sensitive to the basic pricing model and should be handled with care.

Schematic diagram of exotic option classification system:



Pricing methodologies

Now let's look at the (numerical) pricing of exotic options.

Two main methods:

- Pricing via simulations, Monte Carlo

- Formulating the pricing problem in terms of partial differential equations, for solving by finite-difference methods

Pricing via expectations, Monte Carlo simulation

We can value options in the Black–Scholes world by taking the present value of the expected payoff under a risk-neutral random walk.

Simply simulate the random walk

$$dS = rS\,dt + \sigma S\,dX$$

for many paths, calculate the payoff for each path—and this means calculating the value of the path-dependent quantity which is usually very simple to do—take the average payoff over all the paths and then take the present value of that average.

That is the option's fair value.

When and when not to use MC

This is a very general and powerful technique.

When to use MC:

- Good when there are a large number of dimensions
- Useful for path-dependent contracts (even if low dimensions!) for which a partial differential equation approach is tedious to set up
- Some models (e.g. HJM) are built for MC, not easy (or impossible) to write as PDE

When not to use MC:

- The main disadvantage is that it is hard to value options with embedded decisions using MC simulation

Partial differential equations and finite differences

To be able to turn the valuation of a derivatives contract into the solution of a partial differential equation is a big step forward.

- The partial differential equation approach is one of the best ways to price a contract because of its flexibility and because of the large body of knowledge that has grown up around the fast and accurate numerical solution of these problems

- But there is effort involved in setting up the PDE for numerical solution. (In contrast, Monte Carlo can be used 'straight out of the bag')

Let's look at setting up the PDE approach for two examples, a **barrier** option and an **Asian** option.

Both of these can be priced via Monte Carlo but finite-difference solution of the PDEs will be faster.

- Is it worth the effort? Sometimes you might do initial pricing via MC (just to get a 'number') and then you'll spend a bit of time coding up finite differences before it goes into the bank's 'system.'

After we've looked at these two problems we'll do a general theory of path-dependent options.

Barrier options

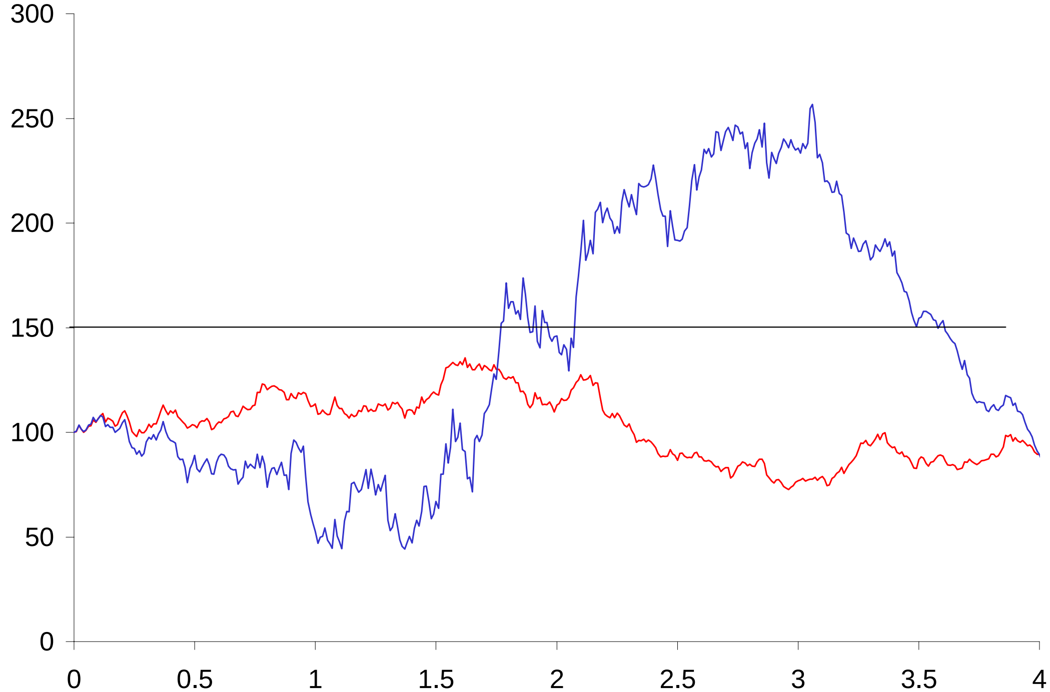
- **Barrier options** have a payoff that is contingent on the underlying asset reaching some specified level before expiry.

The critical level is called the barrier, there may be more than one.

Barrier options come in two main varieties, the ‘in’ barrier option (or **knock-in**) and the ‘out’ barrier option (or **knock-out**). The former only have a payoff if the barrier level is reached before expiry and the latter only have a payoff if the barrier is *not* reached before expiry.

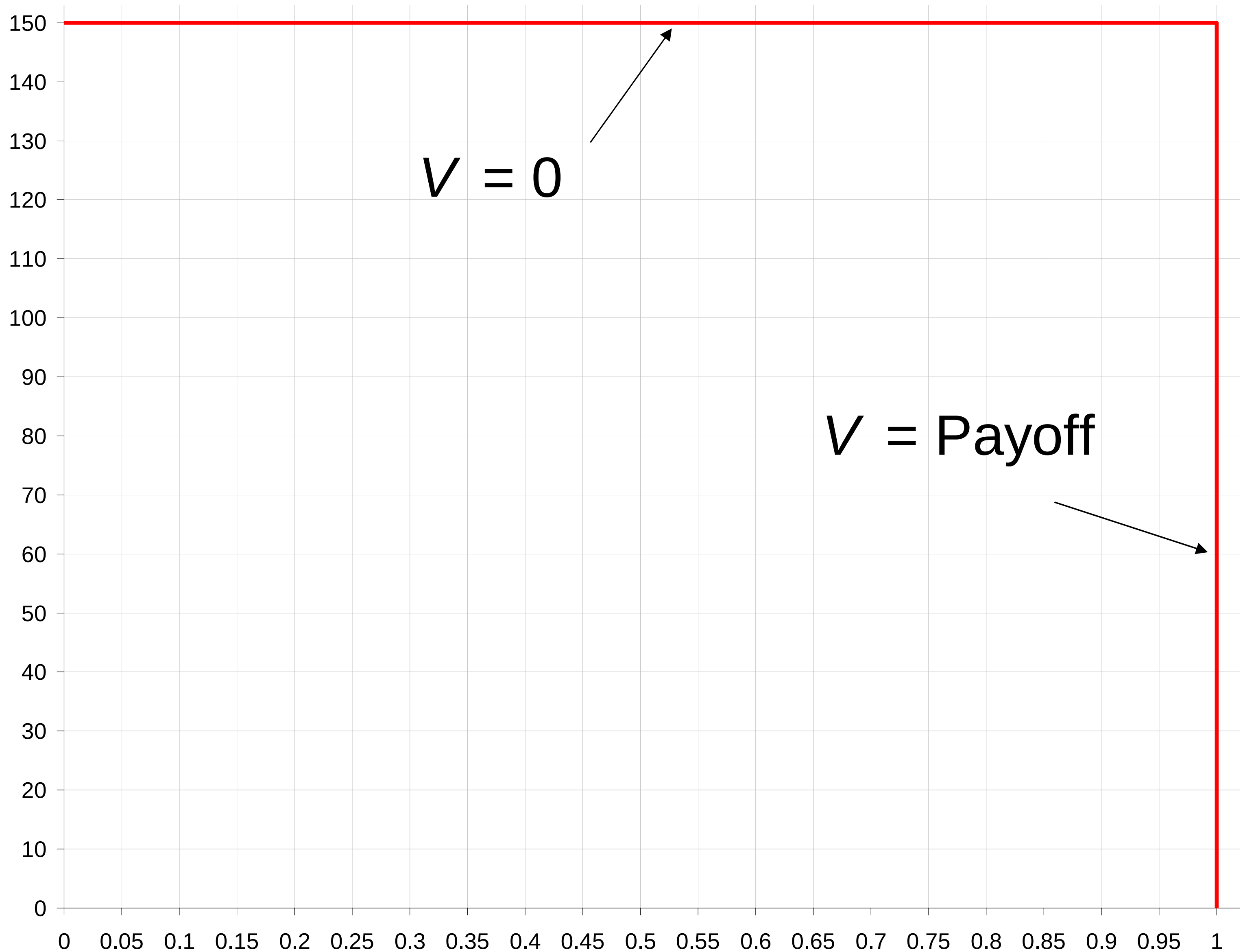
These contracts are weakly path dependent.

Example: An up-and-out call option. This has a call payoff at expiration unless the barrier has been triggered some time before expiration.



This is easily solved by Monte Carlo simulation or by finite-difference methods.

The latter is much preferable. And because the barrier option is weakly path dependent the relevant PDE is exactly the classical Black–Scholes equation! We just have to figure out **initial** and **boundary conditions**.



150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1
$V = 0$
V = Payoff

Asian options

- **Asian options** have a payoff that depends on the average value of the underlying asset over some period before expiry.

They are strongly path dependent. Their value prior to expiry depends on the path taken.

The average used in the calculation of the option's payoff can be defined in many different ways.

It can be an **arithmetic average** or a **geometric average**, for example.

The data could be **continuously sampled**, so that every realized asset price over the given period is used. More commonly, for practical and legal reasons, the data is usually **sampled discretely**.

How is the continuously sampled arithmetic average, for example, defined mathematically?

The final payoff is a function of

$$A = \frac{1}{T} \int_0^T S(\tau)\, d\tau.$$

(The averaging started at time $t = 0$.)

But the running average, and hence our new **state variable** is

$$A = \frac{1}{t} \int_0^t S(\tau)\, d\tau.$$

We need a theory for options with payoff depending on integrals.

A theory for strong path dependence

We will now look at

- pricing many strongly path-dependent contracts in the Black–Scholes partial differential equation framework
- how to handle both continuously sampled and discretely sampled paths
- jump conditions for differential equations

We will now see how to generalize the Black–Scholes analysis, delta hedging and no arbitrage, to the pricing of many more derivative contracts, specifically contracts that are strongly path dependent.

Path-dependent quantities represented by an integral

We start by assuming that the underlying asset follows the log-normal random walk

$$dS = \mu S \, dt + \sigma S \, dX.$$

Imagine a contract that pays off at expiry, T, an amount that is a function of the path taken by the asset between time zero and expiry.

- Let us suppose that this path-dependent quantity can be represented by an integral of some function of the asset over the period zero to T:

$$I(T) = \int_0^T f(S, \tau)\, d\tau.$$

This is not such a strong assumption, many of the path-dependent quantities in exotic derivative contracts, such as averages, can be written in this form with a suitable choice of $f(S, t)$.

You might think that we need to model and remember S at every single moment between now and expiration.

This may look like a problem with an infinite number of variables.

It is much easier than this.

For the basic Asian option it turns out, as we shall see, that not all of the past of the asset matters, only one 'functional' of it.

Prior to expiry we have information about the possible final value of S (at time T) in the present value of S (at time t).

For example, the higher S is today, the higher it will probably end up at expiry.

Similarly, we have information about the possible final value of I in the value of the integral to date:

$$I(t) = \int_0^t f(S,\tau)d\tau. \tag{1}$$

As we get closer to expiry, so we become more confident about the final value of I.

- One can imagine that the value of the option is therefore not only a function of S and t, but also a function of I; I will be our new independent variable, called a **state variable**.

We see in the next section how this observation leads to a pricing equation.

In anticipation of an argument that will use Itô's lemma, we need to know the stochastic differential equation satisfied by I.

This could not be simpler.

- Incrementing t by dt in (1) we find that

$$dI = f(S,t)\,dt + 0 \times dX \tag{2}$$

$$dI^2 = 0$$

Observe that I is a smooth function (except at discontinuities of f) and from (2) we can see that its stochastic differential equation does not contain any stochastic terms.

Continuous sampling: The pricing equation

We will derive the pricing partial differential equation for a contract that pays some function of our new variable I.

- The value of the contract is now a function of the three variables, $V(S, I, t)$.

Set up a portfolio containing one of the path-dependent option and short a number Δ of the underlying asset:

new var.

$$\Pi = V(S, I, t) - \Delta S.$$

$t \to t + dt$

$d\Pi = dV - \Delta dS$

The change in the value of this portfolio is given by

$\sigma^2 S^2 dt$

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \frac{\partial V}{\partial I} dI + \left(\frac{\partial V}{\partial S} - \Delta \right) dS.$$

$$dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial I} dI + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} dS^2$$

GBM

$f dt$

- Choosing

$$\Delta = \frac{\partial V}{\partial S}$$ to eliminate risk

to hedge the risk, and using (2), we find that

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + f(S,t)\frac{\partial V}{\partial I} \right) dt. \quad = r\Pi dt$$

This change is risk free, and thus earns the risk-free rate of interest r, leading to the pricing equation...

Use No-arb. principle

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + f(S,t)\frac{\partial V}{\partial I} + rS\frac{\partial V}{\partial S} - rV = 0$$

new term due to strong path dep.

This is to be solved subject to

$$V(S,I,T) = \text{Payoff}(S,I).$$

This completes the formulation of the valuation problem.

$$I(t) = \int_0^t f(S,\tau)\,d\tau$$

Example:

Continuing with the arithmetic Asian example, we have

$$I = \int_0^t S \, d\tau,$$

so that the equation to be solved is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + S\frac{\partial V}{\partial I} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Similarity reductions

As long as the stochastic differential equation for the path-dependent quantity only contains references to S, t and the path-dependent quantity itself then the value of the option depends on three variables.

Unless we are very lucky, the value of the option must be calculated numerically.

- Some options have a particular structure that permits a reduction in the dimensionality of the problem by use of a similarity variable.

The dimensionality of the continuously sampled arithmetic average strike option can be reduced from three to two.

The payoff for the continuously sampled arithmetic average strike call option is

$$\max\left(S - \frac{1}{T}\int_0^T S(\tau)\,d\tau,\ 0\right) = \max\left(S - \frac{I}{T},\ 0\right)$$

This can be written as

$$I\max\left(\frac{S}{I} - \frac{1}{T},\ 0\right)$$

$$I \max\left(R - \frac{1}{T},\ 0\right)$$

where

$$I = \int_0^t S(\tau)\,d\tau$$

$$V(S, I, t) = I\,W(R, t)$$

and

$$R = \frac{S}{\int_0^t S(\tau)\,d\tau}.$$

In view of the form of the payoff function, it seems plausible that the option value takes the form

$$V(S, I, t) = IW(R, t), \quad \text{with} \quad R = \frac{S}{I}.$$

We find that W satisfies

Laplace Transforms

$$\frac{\partial W}{\partial t} + \tfrac{1}{2}\sigma^2 R^2 \frac{\partial^2 W}{\partial R^2} + R(r - R)\frac{\partial W}{\partial R} - (r - R)W = 0$$

with final condition

$$W(R, T) = \max\left(R - \frac{1}{T}, 0\right).$$

Path-dependent quantities represented by an updating rule

For practical and legal reasons path-dependent quantities are never measured continuously.

There is minimum time step between sampling of the path-dependent quantity.

- From a practical viewpoint it is difficult to incorporate every single traded price into an average, for example. Data can be unreliable and the exact time of a trade may not be known accurately.

If the time between samples is small we can confidently use a continuous-sampling model, the error will be small.

If the time between samples is long, or the time to expiry itself is short we must build this into our model. This is the goal of this section.

When path-dependent quantities are sampled discretely we have summations instead of integrals.

Example: The discretely sampled Asian option

We saw how to use the continuous running integral in the valuation of Asian options.

But what if that integral is replaced by a discrete sum?

In practice, the payoff for an Asian option depends on

$$A_M = \frac{I_M}{M} = \frac{1}{M}\sum_{k=1}^{M} S(t_k), \tag{3}$$

where M is the total number of sampling dates.

This is the discretely sampled average.

This is simply a discrete version of our earlier continuous integral.

Example: The payoff for a discretely sampled arithmetic average strike put is then

$$\max\left(A_M - S, 0\right).$$

Must we remember every single $S(t_k)$ to price the option? That would be an $M + 2$-dimensional problem!

Recall that when valuing the continuously sampled Asian option we only had to remember the value of a single new quantity, the average to date.

Is the same true with the discretely sampled Asian?

Can we write the expression for the running discretely sampled average in a form that does not require us to remember every single $S(t_k)$?

Yes.

$$A_i = \frac{1}{i} \sum_{k=1}^{i} S(t_k)$$

So that

$$\begin{aligned}
A_1 &= S(t_1), \\
A_2 &= \frac{S(t_1) + S(t_2)}{2} = \tfrac{1}{2}A_1 + \tfrac{1}{2}S(t_2), \\
A_3 &= \frac{S(t_1) + S(t_2) + S(t_3)}{3} = \tfrac{2}{3}A_2 + \tfrac{1}{3}S(t_3), \cdots .
\end{aligned}$$

$$A_i = \frac{i-1}{i} A_{i-1} + \frac{1}{i} S_i$$

- This can be expressed as an **updating rule**

$$A_i = \frac{1}{i} S(t_i) + \frac{i-1}{i} A_{i-1}.$$

Generalization

An **updating rule** is an algorithm for defining the path-dependent quantity in terms of the current 'state of the world.'

- The path-dependent quantity is measured on the **sampling dates** t_i, and takes the value I_i for $t_i \leq t < t_{i+1}$.

- At the sampling date t_i the quantity I_{i-1} is updated according to a rule such as

$$I_i = F(S(t_i), I_{i-1}, i).$$

  Note how, in this simplest example, the new value of I is determined by only the old value of I and the value of the underlying on the sampling date, and the sampling date.

Another example: the Lookback option

We will see how to use this for pricing in the next section. But first, another example.

The lookback option has a payoff that depends on the maximum or minimum of the realized asset price.

Realised max $M = \max_{1 \le i \le N} S(t_i)$

Realised min $m = \min_{1 \le i \le N} S(t_i)$

$\max(S_T - E,\ 0)$

$\hookrightarrow \max(S_{t_i})$

If the payoff depends on the maximum sampled at times t_i then we have

$$I_1 = S(t_1), \quad I_2 = \max(S(t_2), I_1), \quad I_3 = \max(S(t_3), I_2) \cdots.$$

- The updating rule is therefore simply

$$I_i = \max(S(t_i), I_{i-1}).$$

How do we use these updating rules in the pricing of derivatives?

Discrete sampling: The pricing equation

- We anticipate that the option value will be a function of three variables, $V(S, I, t)$.

The first step in the derivation is the observation that the stochastic differential equation for I is degenerate:

$$dI = 0.$$

This is because the variable I can only change at the discrete set of dates t_i. This is true if $t \neq t_i$ for any i.

- So provided we are not *on* a sampling date the quantity I is constant, the stochastic differential equation for I reflects this, and the pricing equation is simply the basic Black–Scholes equation:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

How does the equation know about the path dependency?

- Across a sampling date the option value is continuous.

As we get closer and closer to the sampling date we become more and more sure about the value that I will take according to the updating rule.

Since the outcome on the sampling date is known and since *no money changes hands* there cannot be any jump in the value of the option.

This is a simple application of the no arbitrage principle.

We introduce the notation t_i^- to mean the time infinitesimally before the sampling date t_i and t_i^+ to mean infinitesimally just after the sampling date.

Continuity of the option value is represented mathematically by

$$V(S, I_{i-1}, t_i^-) = V(S, I_i, t_i^+).$$

- In terms of the updating rule, we have

$$V(S, I, t_i^-) = V(S, F(S(t_i), I, i), t_i^+)$$

  This is called a **jump condition**.

Examples:

- To price an arithmetic Asian option with the average sampled at times t_i solve the Black–Scholes equation for $V(S,A,t)$ with

$$V(S,A,t_i^-) = V\left(S, \frac{i-1}{i}A + \frac{1}{i}S, t_i^+\right).$$

- To price a lookback depending on the maximum sampled at times t_i solve the Black–Scholes equation for $V(S,M,t)$ with

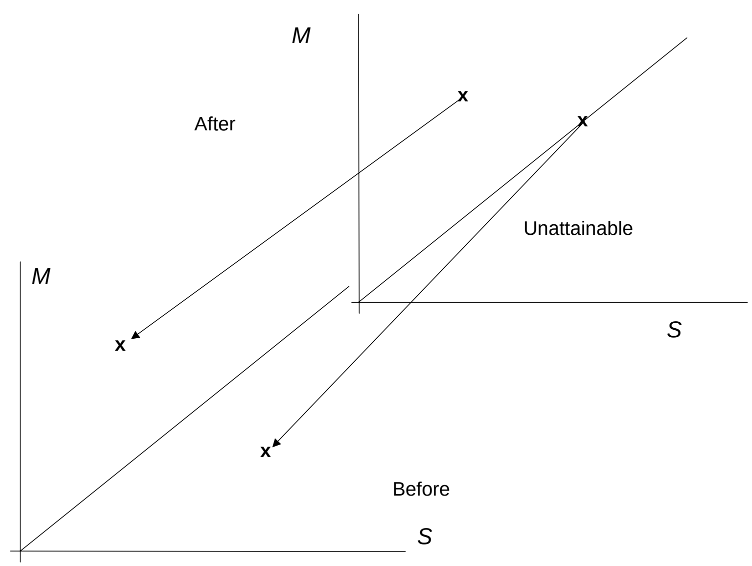
$$V(S,M,t_i^-) = V\left(S, \max(S,M), t_i^+\right).$$

Let's see how this is applied.

The top right-hand plot is the S, M plane just after the sample of the maximum has been taken.

Because the sample has just been taken the region $S > M$ cannot be reached, it is the region labeled 'Unattainable.'

This means that the option value at time t_i^- for $S < M$ is the same as the t_i^+ value. *However*, for $S > M$ the option value comes from the $S = M$ line at time t_i^+ for the same S value.



The algorithm for discrete sampling

The path-dependent quantity, I, is updated discretely and so the partial differential equation for the option value between sampling dates is the Black–Scholes equation.

The algorithm for valuing an option on a discretely sampled quantity is as follows.

- Working backwards from expiry, solve

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

  between sampling dates. Stop when you get to the time step on which the sampling takes place.

- Then apply the appropriate jump condition across the current sampling date to deduce the option value immediately before the present sampling date using the calculated value of the option just after. Use this as your final condition for further time stepping of the Black–Scholes equation.

- Repeat this process to arrive at the current value of the option.

When and when not to use PDEs/FD

PDEs and finite differences take more effort to set up than MC but the reward can be in speed, etc.

When to use PDEs/FD:

- These techniques are best for low dimensions

- They handle embedded decisions exceptionally well

When not to use PDEs/FD:

- When there are high dimensions finite differences will struggle

- Some path dependency can be difficult to model

Volatility Considerations

In this lecture

- The many types of volatility.
- Further greeks - Vega.
- What the market prices of options tells us about volatility.
- The term structure of volatility.
- Numerical methods.
- Stochastic Volatility

Introduction

In this section, we discuss details of volatility. Volatility is one of the most interesting and important aspects of Quantitative Finance. Volatility is the key parameter determining the price of an option, yet it is also the hardest to measure. There are many types of volatility; the precise nature and difference is very important - it is crucial that we know which volatility we are talking about. This adds to the difficulty of volatility considerations.

In the Black-Scholes model, the SDE for the stock has two parameters, μ and σ but later the drift disappears even though the stock depends on it. Some find this counter-intuitive that the value of a call option does not depend on whether the underlying stock is more likely to go up than it is to go down. Recall this is a consequence of hedging. Hence the importance of modelling the volatility 'correctly' if in the business of derivative pricing. If not things become increasingly complex! Suppose we are concerned with stock selection,then the drift comes back in.

Different types of volatility

- Actual/Local
- Historical/Realised
- Implied
- Forward
- Stochastic

The Black-Scholes model is very elegant but it does not perform well in practice. A basic assumption of the framework is a constant geometric Brownian motion for the underlying

$$\frac{dS}{S} = \mu dt + \sigma dW_t$$

and leads to a partial differential equation for which either an analytical solution exists or can be treated numerically.

Thus far the role of σ is that of a parameter. It is the most important parameter when pricing an option, and is also the most difficult to measure. In comparing the solution to reality, the important question arising is "How plausible is a constant volatility?". Recapping the model for a Call option $C(S,t)$, the pricing equation and terminal condition in turn

$$\begin{aligned} \frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC &= 0 \\ C(S,T) &= \max(S-E,0). \end{aligned}$$

The solution is

$$C(S,t) = SN(d_1) - Ee^{-r(T-t)}N(d_2)$$

where

$$d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \text{ and } d_2 = d_1 - \sigma\sqrt{T-t}.$$

This pricing formula relates the option price C to six arguments; the variables S, t and the parameters r, σ, E and T.

Quantity	Observable?	
C_M	yes	quoted market option price
S	yes	today's spot price
t	yes	today's date
T	yes	expiration
E	yes	strike price
r	yes	today's interest rate
σ	No	

So in derivatives the volatility is the most important parameter.

Drift is not important.

It doesn't matter if e.g. it is doubling in price or halving in price. It is the level of noise/randomness that affects the price. But it is very hard to measure - volatility cannot be seen/observed. It is how much randomness there is in a stock price in an instant in time. For these reasons different types of volatility needs to be discussed.

Spot Volatility

- Define logarithmic returns $R_t := \log\left(\frac{S_{t+\delta t}}{S_t}\right)$.

- The mean and variance in turn are

$$m_t = \mathbb{E}\left[R_t\right], \quad v_t = \mathbb{E}\left[(R_t - m_t)^2\right]$$

- The classical Black-Scholes stock price process $S_t,\ t \geq 0$ given by GBM is

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

- Using Itô to calculate $d(\log S_t)$ gives the solution as

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

- Now look at the mean and variance of the logarithmic return $\log\left(\frac{S_t}{S_0}\right)$

$$\mathbb{E}\left[\log\left(\frac{S_t}{S_0}\right)\right] =: \eta t := \left(\mu - \frac{1}{2}\sigma^2\right) t$$

$$\mathbb{V}\left[\log\left(\frac{S_t}{S_0}\right)\right] =: s^2\left(t\right) := \sigma^2 t$$

- This gives one definition of volatility σ. It is a measure of the variance of the logarithmic returns, such that the square of the volatility gives the rate of increase of the log-returns:

$$\frac{d}{dt}s^2\left(t\right) = \sigma^2.$$

Equivalently, if we denote the quadratic variation of a stochastic process Y_t by $[Y]_t$, then we have

$$[\log S]_t = \sigma^2 t,$$

so the squared volatility is the rate of change of the quadratic variation of the log-stock price.

Actual Volatility

Is the most important type and the one which quants refer to. It can't be seen. It is the standard deviation of stock price returns. When we refer to

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

it is the term σ. It is also called spot volatility. Or writing

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0,$$

again it is σ that is actual volatility.

It is the volatility at a particular instant in time - it is always changing. If the stock price exhibits the following behaviour ....

and then a few seconds later does ....

and then ....

this is volatility changing.

We aren't interested in the average of the volatility over some time , but every instant in time, i.e. $\sigma(t)$ or function or some random variable. There is no timescale associated with actual vol because it is at an instant in time. Important to appreciate that it is not over a period.

Example: The actual volatility is now 20%, now it is 22%, now it is 24%, ... because it is not quoted it is very difficult to know what it is from nanosecond to nanosecond.

Historical/Realised Volatility

This is a statistical measure of volatility going back in time. Also an attempt to get a feel for what actual vol is. Compared to actual volatility this is something you can measure. If we have stock prices over a year (say) might look at daily returns and calculate a standard deviation and annualise it - this gives us the volatility over that year. It's realised because it actually happened but it's not today's volatility. So to obtain any sense of volatility now depends on taking a look at volatility in the past (say even last week) but it is not the volatility at this second.

When looking at historical volatility there are two time scales

- long timescale
- short timescale

So consider data over one year (long) and we divide this up in to 52 weeks (short). So this can give us the average over that period **not** a particular point.

An accurate measure of vol requires a lot of data. Small data sets lead to larger sampling errors.

Example

Suppose we have a year's data - 252 days is reasonable but not great either. Now consider having 10 years of stock price data i.e. 2520 - but this means you are going back much further in time, hence very different economic environment which may not be relevant to today.

A different exercise is to reduce the long time scale to 1 week (previous) and the shorter one to hourly increments; or tick data - this gives a very large amount of information. But the pitfall is that the dynamics of what happens over minutes is different to what happens over days; scaling up from minutes to days is difficult.

If hedging options, then you're hedging once a day or perhaps once every two days. What happens minute by minute, you don't care about what happens day by day.

So going too far back takes us to irrelevant economic times while looking at very small timescales gives us totally different dynamics. Scaling up from minutes to days not only difficult but then useless over days.

Takes us back to measure of volatility being very difficult. The best we may get is that say over the last six months the daily volatility was on average 23%. We have this for volatility now and use it to model going forward because we may have an option expiring in a years time.

Forward Volatility

The adjective 'Forward' can be applied to many forms of volatility, and refers to the volatility (whether actual or implied) over some period in the future.

Forward volatility is associated with either a time period, or a future instant.

Deterministic Volatility Models

Simplest generalisation of the Black-Scholes constant volatility paradigm is to allow the volatility to be a deterministic function of time so that the stock price SDE becomes

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

and the modified Black-Scholes equation becomes

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(t) S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

for the option price $V(t, S)$, where $V(T, S) = h(S)$.

By the Feynman-Kac theorem the option pricing function is given by the risk-neutral expectation

$$V(t, S) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)} h(S_T) \middle| S_t = x\right],$$

where under the $\mathbb{Q}$ measure, S_t follows GBM above with μ replaced by r :

$$dS_t = \mu S_t dt + \sigma(t) S_t dW_t^{\mathbb{Q}},$$

where $W_t^{\mathbb{Q}}$ is a $\mathbb{Q}$ Brownian motion. Under $\mathbb{Q}$, the terminal log-stock price is now given by

$$\log S_T = \log S_t + r(T-t) - \frac{1}{2}\int_t^T \sigma^2(s)\,ds + \int_t^T \sigma(s)\,dW_s^{\mathbb{Q}}$$

Hence, under $\mathbb{Q}$, given $S_t = x$, $\log S_T$ is normally distributed:

$$\log S_T \sim N\left(\log x + \left(r - \tfrac{1}{2}\overline{\sigma}^2\right)(T-t), \overline{\sigma}^2(T-t)\right),$$

where $\overline{\sigma}^2$ is the root mean square (RMS) volatility, given by

$$\overline{\sigma}^2(T-t) = \int_t^T \sigma^2(s)\,ds$$

It follows that one simply prices the option using the Black-Scholes formula with the volatility σ replaced by

$$\overline{\sigma}_t^2 = \frac{1}{(T-t)}\int_t^T \sigma^2(s)\,ds$$

Thus, in all BS pricing formulas for European, path-independent options, just replace σ by $\overline{\sigma}_t$.

For example, the price of a vanilla call at time t is given by

$$C\left(\overline{\sigma}_t^2, S, t\right) = SN(d_1) - Ee^{-r(T-t)}N(d_2)$$

where

$$\begin{aligned}
d_1 &= \frac{\log(S/E) + (r + \frac{1}{2}\overline{\sigma}_t^2)(T-t)}{\overline{\sigma}_t\sqrt{T-t}}, \\
d_2 &= d_1 - \overline{\sigma}_t\sqrt{T-t}, \\
\overline{\sigma}_t^2 &= \frac{1}{(T-t)}\int_t^T \sigma^2(s)\,ds.
\end{aligned}$$

Local Volatility

- Further generalisation of Black-Scholes:

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

- The deterministic function

$$(t, S) \to \sigma(t, S_t)$$

  is called *local volatility*.

- Option price $V(t, S_t)$ for terminal payoff $h(S_T)$ satisfies the BSE

$$\begin{aligned} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(t, S) S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV &= 0 \\ V(T, S_T) &= h(S_T). \end{aligned}$$

- The model is also referred to as 'restricted stochastic' volatility model, since the volatility path $\sigma_t = \sigma(t, S_t)$ is stochastic, but the only source of randomness enters through the state variable S.

Special cases are

i. $\sigma = \sigma(t)$ is a function of time alone: There is a term-structure, but the model fails to predict smiles and skews. As discussed above.

ii. $\sigma = \sigma(S)$ is a function of the stock alone: An important family are constant elasticity of variance (CEV) models of the form

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

with an additional parameter α. Again, closed form solutions exist.

Implied Volatility

Comes from looking at option prices. It is what the market's view of the actual volatility over the lifetime of the particular option. It's what the market thinks the volatility is. There is one timescale associated with implied volatility, the expiration. It is the value of the volatility which when substituted into the Black-Scholes equation gives the price of the option in the market.

Scenario: Suppose you see an option on your trading screen priced at £10.45.

The stock is trading at £100 strike is £100, 1 year to expiry, no dividends, interest rate is 5%.

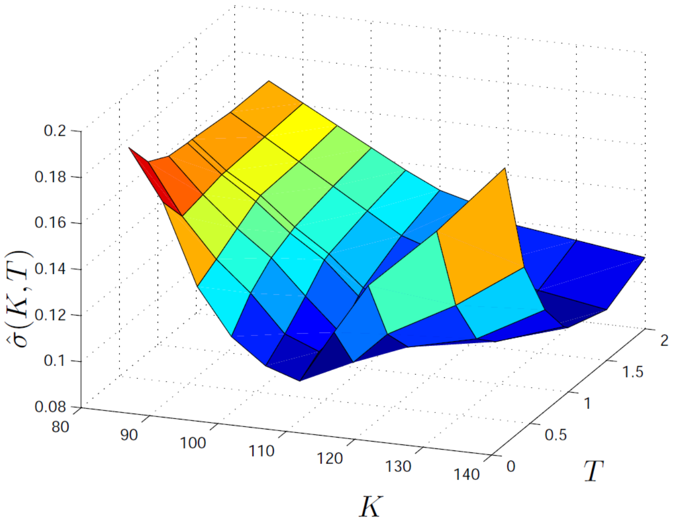
This is all the information you need for the Black-Scholes formula to price an option with the exception of volatility.

If we know the volatility, then the option price can be obtained.

However we know the actual market price – it is £10.45. So we ask what value of the volatility called σ_{i} (or $\widehat{\sigma}$) gives the observed value of 10.45. This is an inverse problem requiring computation - find the volatility given 10.45; $\sigma_{\mathsf{i}} = 20\%$, (left as an exercise for verification).

σ_{i} can be seen as it comes from the observed option price.

Consider a three-dimensional plot with strike and expiry in the $x - y$ plane; volatility along the $z-$axis. If the Black-Scholes model were correct then the surface plot called a *volatility surface*; would be a flat surface, i.e. $\sigma(E, T) = \sigma$. In practice the dynamics exhibited are quite different!



In the previous table we note that C_M, S, t, T, E and r are observables. For a plain vanilla option we define the *implied volatility* denoted σ_i to be that value of the unobservable σ, which gives the market price of the option when substituted into the Black-Scholes option pricing formula. All the other observables are fixed in this process. It is described as the market's view of the future actual volatility during the life of the option.

The implied volatility according to the Black-Scholes model should be independent of both strike and expiration; in reality it depends on both.

Consider the following example: A trader can see on their screen that a certain call option with one year until expiry and a strike of £100 is trading at £10.45 with the underlying at £100 and a short-term interest rate of 5%. Can we use this information in some way?

We can take invert this relationship between volatility and an option price by asking "What volatility must I use to get the correct market price?"

This is called the implied volatility. If $C_{\textsf{BS}}$ denotes the Black-Scholes theoretical price then solving

$$C_M\left(S,t\right)=C_{\textsf{BS}}\left(S,T,r,E,\sigma_i\left(E,T\right)\right)$$

for σ_{i} becomes a root-finding problem. Use of e.g. Newton-Raphson method will work.

For implied volatility to be a useful concept means there should be a unique implied vol. This is only true if the options vega υ, where

$$\upsilon\left(S,t\right)=\frac{\partial V}{\partial\sigma}\left(S,t;\sigma;E,T;r;....\right),$$

does not change sign for any value of S,t or other parameters, besides σ, involved. While this is true for European calls and puts it is true for all options.

Here' an example of data you may see on your screen. Call option prices in the matrix. Strikes along the top; expirations down the side.

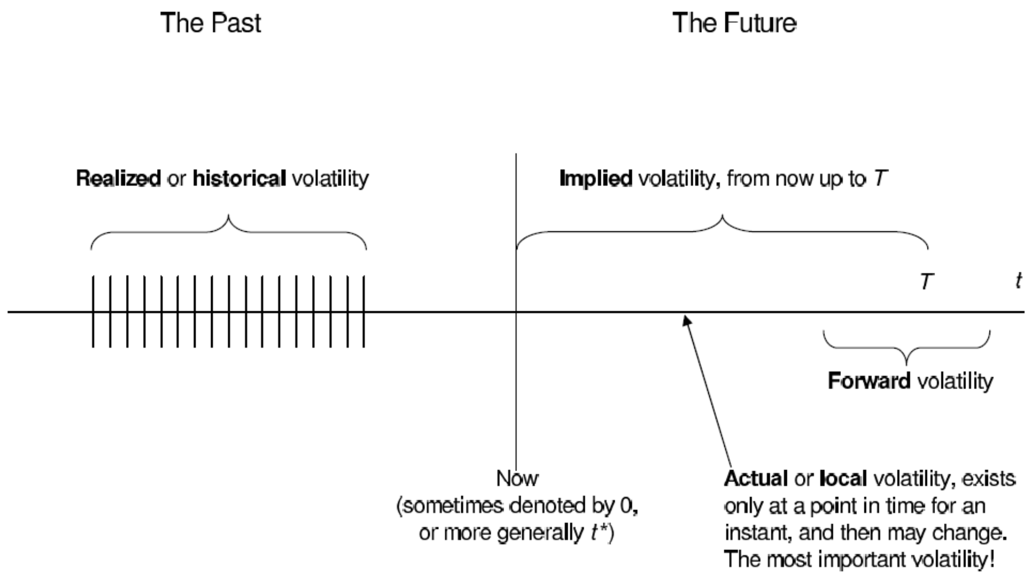
	95	100	105
1 month	11.80	7.24	3.5
3 month	13.33	9.23	5.76
7 month	16.02	12.13	7.97

So as an example, a certain call with strike 100 and expiry of 3 months has a value of 9.23; it is not always easy to interpret these values.

This can be reinterpreted by replacing the option value with implied volatility versus strike and expiry

	95	100	105
1 month	24.1%	22.9%	21.2%
3 month	22.7%	21.5%	20.5%
7 month	21.8%	20.5%	19.4%

So one interpretation of implied vol is the market's best guess of what the vol is going to be in the future. Too good to be true? What doesn't feel quite right about this?



Vega

Vega (also zeta ζ and kappa κ) is a very important but confusing quantity. It is the sensitivity of the option price to volatility.

$$\text{Vega} = \frac{\partial V}{\partial \sigma}$$

This is a completely different from the other Greeks since it is a derivative with respect to a parameter and not a variable.

As with gamma hedging, one can vega hedge to reduce sensitivity to the volatility. This is a major step towards eliminating some model risk, since it reduces dependence on a quantity that is not known very accurately.

An option value can change even when the underlying doesn't move.

Implied volatility can change. A market that is panicking will have a high volatility. Because implied volatility can change we would like to know how sensitive our options are to that change.

Vega is a bastard greek.

A bastard greek is the rate of change of the value of an option to a parameter, as opposed to the classical greeks which are rates of change with respect of variables. Classical greeks are delta, gamma, theta, etc. Bastard greeks are vega, rho, etc.

Bastard greeks are illegitimate because they usually involve differentiating a formula, such as Black-Scholes, with respect to a parameter that has been assumed constant in the derivation.

In the Black-Scholes framework, it is quite simple to find a sufficient condition that the implied volatility is unique. Consider the simplest case

$$\begin{aligned} \frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+rS\frac{\partial V}{\partial S}-rV &= 0 \\ V(S,T) &= P(S). \end{aligned}$$

Differentiating gives

$$\begin{aligned} \frac{\partial v}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 v}{\partial S^2}+rS\frac{\partial v}{\partial S}-rv &= -\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} \\ v(S,T) &= 0. \end{aligned}$$

Provided the forcing term $\frac{\partial^2 V}{\partial S^2}$ is either strictly positive or strictly negative for $t<T$, v will be strictly negative or strictly positive, in turn.

Assuming that the implied volatility is unique, there is a one-to-one correspondence between the price and implied volatility. In this case, the implied volatility can be used as a proxy for the price.

In particular, implied volatility can be used as a proxy for price in the case of vanilla puts and calls. As a result, and for various historical reasons, in many markets it is used instead of the price.

In a sense the Black-Scholes model, in these markets, can be (and probably should be) thought of as more a means of "translating" between price and implied volatility than a model for option pricing.

This convention is unfortunate because not all options have unique implied volatilities. The convention arose because, historically, vanilla calls and puts were the first options to be widely traded.

Smiles

According to the classical Black-Scholes analysis

$$\frac{dS}{S} = \mu dt + \sigma dW$$

so that σ is a property of S alone. For a vanilla call or put option the strike E and the expiry T are properties only of the option. Thus the volatility (which in the Black-Scholes model should be the same thing as the implied volatility) σ should be independent of both strike E and the expiry T or a vanilla put or call.

In practice, we find the implied volatility for a vanilla call (or put) depends on both the strike and the expiry, the so-called *smile*. This implies that there is something wrong with the Black-Scholes model.

The dependence of implied volatility on expiry could imply a term structure for volatility, $\widehat{\sigma}(t)$, rather than a constant volatility σ. This is not a serious problem; we know how to deal with time dependent volatilities; we just replace σ^2 in the Black-Scholes formulae by

$$\frac{1}{T-t}\int_t^T \widehat{\sigma}^2(s)\,ds.$$

The dependence on strike is a serious problem. It is quite inconsistent with the Black-Scholes analysis.

One way of explaining the smile effect is to assume that the volatility is a function of both spot price and time;

$$\sigma = \sigma(S, t).$$

This is not the only possible explanation.

Leibniz Rule

The shortened version of this is

$$\frac{\partial}{\partial x}\int_a^x F(y,x)dy = F(x,x) + \int_a^x \frac{\partial F(y,x)}{\partial x}dy.$$

We do the fitting at time t^*. If we write $\sigma_\mathsf{i}\left(T;t^*\right)$ to mean the implied volatility measured at time t^* of a European option expiring at time T,

$$\sqrt{\frac{1}{(T-t^*)}\int_{t^*}^{T}\sigma^2\left(s\right)ds} = \sigma_\mathsf{i}\left(t^*;T\right).$$

We can do differentiation under the integral sign using the Leibniz rule to obtain a solution of the inverse problem.

$$\sigma\left(t\right) = \sqrt{\sigma_\mathsf{i}\left(t;t^*\right)^2 + 2\left(t-t^*\right)\sigma_\mathsf{i}\left(t;t^*\right)\frac{\partial}{\partial t}\sigma_\mathsf{i}\left(t;t^*\right)}.$$

This is the solution of the integral equation.

$$\frac{\partial}{\partial T}\int_{t^*}^{T}\sigma_a^2(s)\,ds=\frac{\partial}{\partial T}\left((T-t^*)\,\sigma_{\mathsf{i}}(t^*;T)^2\right)$$

Left hand side is

$$\sigma_a^2(T)$$

The right hand side is product and chain rules

$$\sigma_{\mathsf{i}}(t^*;T)^2+2\sigma_{\mathsf{i}}(t^*;T)\,\frac{\partial}{\partial T}\sigma_{\mathsf{i}}(t^*;T)$$

Hence

$$\sigma_a^2(T)=\sigma_{\mathsf{i}}(t^*;T)^2+2\sigma_{\mathsf{i}}(t^*;T)\,\frac{\partial}{\partial T}\sigma_{\mathsf{i}}(t^*;T)$$

and the solution is obtained by replacing T by t

$$\sigma_a(t)=\sqrt{\sigma_{\mathsf{i}}(t^*;t)^2+2\,(t-t^*)\,\sigma_{\mathsf{i}}(t^*;t)\,\frac{\partial}{\partial T}\sigma_{\mathsf{i}}(t^*;T)\Big|_{T=t}}$$

Volatility Surfaces

One means of implementing a no-arbitrage model is to assume that $\sigma = \sigma(S_t, t)$. The stock price process dynamics follow

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma(S_t, t)\, dW_t$$

The Black-Scholes equation then becomes

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S, t)\, S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

and its solution can be written in the form

$$V(S, t) = e^{-r(T-t)} \int_0^\infty p\left(S, t; S', T\right) V\left(S', T\right) dS'.$$

$V(S', T)$ is the payoff and $p(S, t; S', T)$ is the risk-neutral probability density associated with the Kolmogorov equations

$$\begin{aligned}
\frac{\partial p}{\partial T} &= \frac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S', t\right) S'^2 p\right) - \frac{\partial}{\partial S'}\left(rS'p\right), \\
-\frac{\partial p}{\partial T} &= \frac{1}{2}\sigma^2 (S, t)\, S^2 \frac{\partial^2 p}{\partial S^2} + rS\frac{\partial p}{\partial S}.
\end{aligned}$$

That is $p(S, t; S', T)$ can be viewed in two ways (given that the above analysis only makes sense if $t < T$):

If S and t are fixed (today's spot price and date) then we can regard $p(S, t; S', T)$ as the probability density that at time $T > t$ the spot price will be S'. This is a conditional probability density for future values, S' and T, given that the present values are S and $t < T$.

If S' and T are fixed (some given value of the spot price and date, say) then $p(S, t; S', T)$ is the probability density that at time $t < T$ the spot price was S; again this is a conditional probability function; the probability that the spot price was S at time t given that the spot price is S' at time T.

Dupire's method

Suppose now that we want to find $\sigma(S,t)$ from market data. In fact we shall find $\sigma(E,T)$. More correctly, we find $\sigma(S,t;E,T)$ with the usual notation of the function arguments.

Using

$$\frac{\partial p}{\partial T}=\frac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S',t\right)S'^2p\right)-\frac{\partial}{\partial S'}\left(rS'p\right)$$

and that

$$C(S,t)=e^{-r(T-t)}\int_E^\infty p\left(S,t;S',T\right)\left(S'-E\right)dS'.$$

Then

$$\frac{\partial C}{\partial T}=-rC+e^{-r(T-t)}\int_E^\infty\frac{\partial p}{\partial T}\left(S'-E\right)dS'.$$

Also (from Leibniz)

$$\frac{\partial C}{\partial E} = -e^{-r(T-t)}\int_E^\infty p\left(S,t;S',T\right)dS'$$

$$\frac{\partial^2 C}{\partial E^2} = e^{-r(T-t)}p\left(S,t;E,T\right)$$

Now use the fact that

$$\frac{\partial p}{\partial T} = \frac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S',t\right)S'^2 p\right) - \frac{\partial}{\partial S'}\left(rS'p\right)$$

where as earlier,

$$p\left(S,t;S',T\right)dS'$$

denotes the risk-neutral probability of a spot price in $(S', S'+dS')$ at time $T > t$, contingent on the spot price being S at t.

Recall that p is contingent on today's information, (S,t); in general

$$p\left(S_1,t_1;S',T\right) = p\left(S_2,t_2;S',T\right)$$

for say today, (S_1,t_1), and tomorrow (S_2,t_2).

Substituting for $\frac{\partial p}{\partial T}$ and using integration by parts we arrive, after lengthy calculation, at

$$\frac{\partial C}{\partial T} = -rC + e^{-r(T-t)} \left(\sigma\left(E, T\right)^2 E^2 p\left(S, t; E, T\right) + r \int_E^\infty S' p dS' \right).$$

Later we will use

$$\int_E^\infty S' p dS' = \int_E^\infty \left(S' - E\right) p dS' + E \int_E^\infty p dS'$$

and from earlier we note that

$$\frac{\partial C}{\partial E} = e^{-r(T-t)} \int_E^\infty p dS', \quad \frac{\partial^2 C}{\partial E^2} = e^{-r(T-t)} p\left(S, t; E, T\right),$$

where the second expression gives

$$p\left(S, t; E, T\right) = e^{r(T-t)} \frac{\partial^2 C}{\partial E^2}.$$

Now use the Kolmogorov equation to express $\frac{\partial p}{\partial T}$ in $\frac{\partial C}{\partial T}$

$$\begin{aligned}\frac{\partial C}{\partial T} &= -rC + e^{-r(T-t)}\int_E^\infty \frac{\partial p}{\partial T}\left(S' - E\right) dS' \\ &= -rC + \\ & \quad e^{-r(T-t)}\int_E^\infty \left(\frac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S', t\right) S'^2 p\right) - \frac{\partial}{\partial S'}\left(rS'p\right)\right)\left(S' - E\right) dS'\end{aligned}$$

The integral

$$\int_E^\infty \left(\frac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S', t\right) S'^2 p\right) - \frac{\partial}{\partial S'}\left(rS'p\right)\right)\left(S' - E\right) dS' =$$

$$\int_E^\infty \frac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S', t\right) S'^2 p\right)\left(S' - E\right) dS'$$

$$-\int_E^\infty \frac{\partial}{\partial S'}\left(rS'p\right)\left(S' - E\right) dS'$$

In what follows, we assume p decays sufficiently fast.

$$\frac{1}{2}\int_E^\infty \frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S',T\right)S'^2 p\right)\left(S'-E\right)dS' :$$

$$\begin{array}{ll} v = S'-E & u' = \frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S',T\right)S'^2 p\right) \\ v' = 1 & u = \frac{\partial}{\partial S'}\left(\sigma^2\left(S',T\right)S'^2 p\right) \end{array}$$

$$\begin{aligned} &\frac{1}{2}\int_E^\infty \frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S',T\right)S'^2 p\right)\left(S'-E\right)dS' \\ = \; &\underbrace{\frac{1}{2}\frac{\partial^2}{\partial S'^2}\left(\sigma^2\left(S',T\right)S'^2 p\right)\left(S'-E\right)\Big|_E^\infty}_{=0} - \frac{1}{2}\int_E^\infty \frac{\partial}{\partial S'}\left(\sigma^2\left(S',T\right)S'^2 p\right)dS' \end{aligned}$$

$$= \frac{1}{2}\sigma^2\left(E,T\right)E^2 p\left(S,t;E,T\right) = \frac{1}{2}\sigma^2\left(E,T\right)E^2 e^{r(T-t)}\frac{\partial^2 C}{\partial E^2}$$

Similarly $\int_E^{\infty} \frac{\partial}{\partial S'} (rS'p)\,(S' - E)\,dS'$:

$$\begin{array}{ll} v = S' - E & u' = \frac{\partial}{\partial S'} (rS'p) \\ v' = 1 & u = rS'p \end{array}$$

$$\begin{aligned} & \int_E^{\infty} \frac{\partial}{\partial S'} \left(rS'p\right) \left(S' - E\right) dS' \\ = \ & \underbrace{rS'p\left(S' - E\right)\Big|_E^{\infty}}_{=0} - r\int_E^{\infty} S'pdS' \end{aligned}$$

Now using

$$\int_E^{\infty} S'pdS' = \int_E^{\infty} \left(S' - E\right) pdS' + E\int_E^{\infty} pdS'$$

- For the first integral term it is the expected payoff (i.e. option price without discount factor), i.e. $e^{r(T-t)}C$.

- The second integral term from earlier $\frac{\partial C}{\partial E} = -e^{-r(T-t)}\int_E^\infty p dS'$ gives $-e^{r(T-t)}\frac{\partial C}{\partial E}$

Hence

$$\begin{aligned}
\int_E^\infty \frac{\partial}{\partial S'}\left(rS'p\right)\left(S'-E\right)dS' &= re^{r(T-t)}C - rEe^{r(T-t)}\frac{\partial C}{\partial E}\\
&= -re^{r(T-t)}\left(C - E\frac{\partial C}{\partial E}\right)
\end{aligned}$$

Putting everything together

$$\begin{aligned}\frac{\partial C}{\partial T} &= -rC + e^{-r(T-t)}\left(\tfrac{1}{2}\sigma^2(E,T)\,E^2 e^{r(T-t)}\frac{\partial^2 C}{\partial E^2} + re^{r(T-t)}\left(C - E\frac{\partial C}{\partial E}\right)\right)\\ &= -rC + \tfrac{1}{2}\sigma^2(E,T)\,E^2\frac{\partial^2 C}{\partial E^2} + rC - rE\frac{\partial C}{\partial E}\end{aligned}$$

Hence

$$\frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2(E,T)\,E^2\frac{\partial^2 C}{\partial E^2} - rE\frac{\partial C}{\partial E}.$$

We can now, in principle solve for

$$\sigma^2(E,T) = \frac{\frac{\partial C}{\partial T} + rE\frac{\partial C}{\partial E}}{\frac{1}{2}E^2\frac{\partial^2 C}{\partial E^2}}.$$

If we now retrace our calculations we find that, because the call value C is a function of today's spot, S, today's date, t, the call's strike E and the call's maturity, T, $C = C\left(S, t; E, T\right)$, what we have called $\sigma\left(E, T\right)$ is actually

$$\sigma^2\left(S, t; E, T\right) = \frac{\frac{\partial C(S,t;E,T)}{\partial T} + rE\frac{\partial C(S,t;E,T)}{\partial E}}{\frac{1}{2}E^2\frac{\partial^2 C(S,t;E,T)}{\partial E^2}}.$$

Recall that, in practice, when we compute $\sigma\left(S, t; E, T\right)$, today's spot price S and date t are fixed. We can vary only the strike E and the maturity T.That is, we have found a *local volatility surface* $\sigma\left(E, T\right)$, or more correctly $\sigma\left(S, t; E, T\right)$, as it is conditional on today's spot price S and date t.

Practical problems with this approach

- requires continuum of strikes and maturities (interpolation, extrapolation)

- numerical differentiation is ill conditioned

- the denominator $\frac{\partial^2 C}{\partial E^2}$ tends to zero for $E \longrightarrow \infty$

The last problem can be circumvented to some extent by switching from quoted prices to implied volatilities.

Implied and local volatility

If we use implied volatilities σ_i, repeated application of the implicit function theorem gives

$$\sigma^2(E,T) =$$

$$\frac{\sigma_i^2 + 2\sigma_i(T-t)\frac{\partial\sigma_i}{\partial T} + 2r\sigma_i E(T-t)\frac{\partial\sigma_i}{\partial E}}{\left(1 + Ed_1\sqrt{T-t}\frac{\partial\sigma_i}{\partial E}\right)^2 + \sigma_i(T-t)E^2\left(\frac{\partial^2\sigma_i}{\partial E^2} - d_1\left(\frac{\partial\sigma_i}{\partial E}\right)^2\sqrt{T-t}\right)}$$

where, as usual,

$$d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma_i^2)(T-t)}{\sigma_i\sqrt{T-t}}$$

Finding Roots

A fundamental problem in numerical analysis consists of obtaining the zero of a function. Given a function $y = f(x)$ obtain the root of $f(x) = 0$, i.e. find the value of $x = c$ which satisfies $f(c) = 0$. e.g.

$$f(x) = x - \sin x.$$

Four broad categories of root finding:

(i) Methods which do not use derivatives of the function

(ii) Methods which do use $f'(x)$

(iii) Methods for polynomials

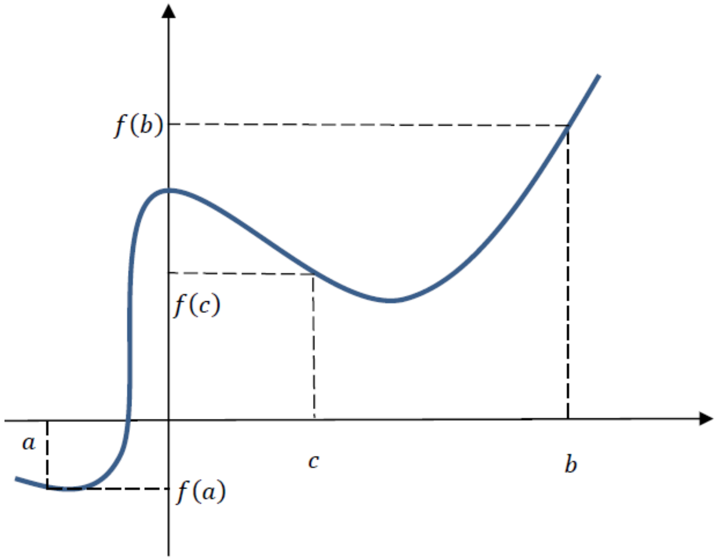
(iv) Methods which deal with complex roots

Bisection

The simplest method is that of bisection. The following theorem, from calculus class, insures the success of the method.

Intermediate Value Theorem

Suppose $f(x)$ is continuous on $[a,\ b]$ then for any y s.t y is between $f(a)$ and $f(b)$ there $\exists\ c \in [a,\ b]$ s.t $f(c) = y$.



Example 1

The function $f(x) = \dfrac{1}{x}$ is not continuous at 0. Thus if $0 \in [a,\ b]$, we *cannot* apply the IVT. In particular, if $0 \in [a,\ b]$ it happens to be the case that for *every* y between $f(a),\ f(b)$ there is no $c \in [a,\ b]$ such that $f(c) = y$.

In particular, the IVT tells us that if $f(x)$ is continuous and we know $a,\ b$ such that $f(a),\ f(b)$ have different sign, then there is some root in $[a,\ b]$. This is a fundamental test we can apply.

Example 2 Show that the function $g(x) = x^3 + 2x^2 + 5x - 1$ has a root lying between 0 and 1.

We note $f(0) = -1;\ f(1) = 7$. The sign change confirms that $\exists$ a root α s.t. $\alpha \in (0, 1)$.

Once location of a root α is established then a reasonable estimate of α is $c = \dfrac{a+b}{2}$. We can check whether $f(c) = 0$. If this does not hold then one and only one of the two following options holds:

1. $f(a)\,,\ f(c)$ have different signs.

2. $f(c)\,,\ f(b)$ have different signs.

We now choose to recursively apply bisection to either $[a,\ c]$ or $[c,\ b]$, respectively, depending on which of these two options hold.

Whichever an interval is chosen,the new interval containing the root can be further subdivided. If the first interval is $|b-a|$, then the second is half the length and so on.

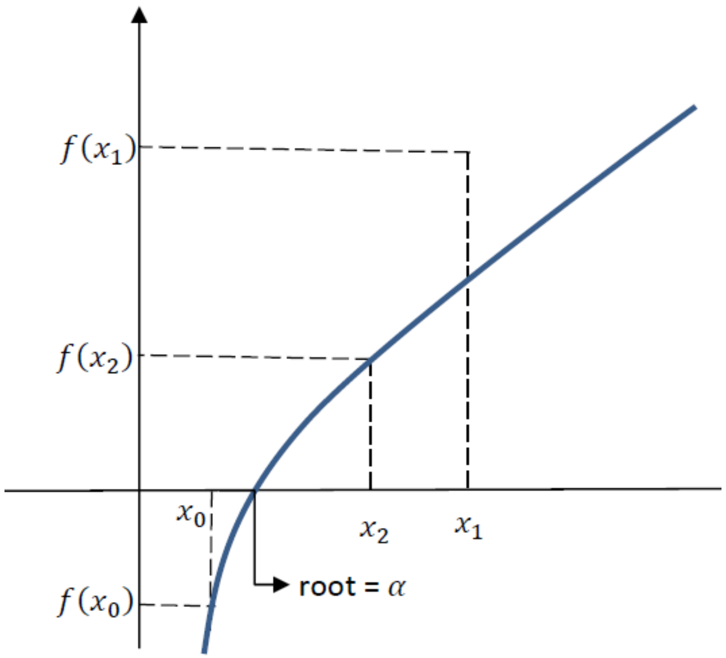
After n steps of bisection the interval containing the root will be reduced in size to

$$\frac{|b-a|}{2^n}$$

where in the earlier example the value $b = 1$ and $a = 0$.

If the size of the interval becomes smaller than some specified tolerance, t, then the calculation stops and convergence has been attained.

Theorem 2 (Bisection Method Theorem)

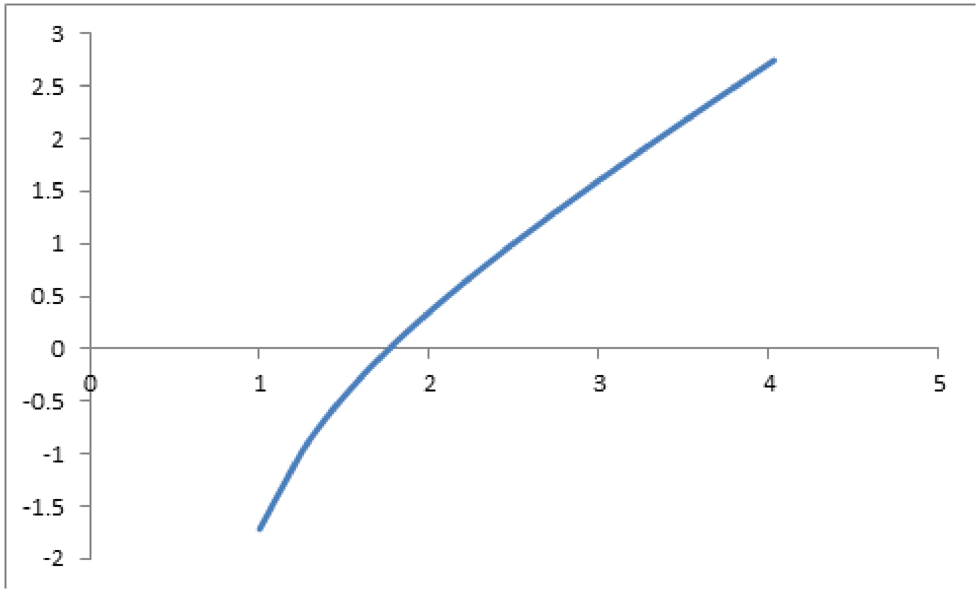


If $f(x)$ is a continuous function on $[a,\ b]$ such that $f(a)f(b) < 0$, then after n steps, the method will return c such that

$$|c - \alpha| \leq \frac{|b - a|}{2^n}$$

where α is some approximate root of f.

Example Consider $f(x) = x - e^{1/x}$. There is a root in $[1,\ 2]$.



Use the bisection method to show that the root of $f(x) = x - e^{1/x}$ in the interval $[1,\ 2]$ is 1.763 (correct to 3 decimal places).

$$\left.\begin{array}{r} f(x_0) = f(1) = 1 - e^1 < 0 \\ f(x_1) = f(2) > 0 \end{array}\right| \quad \rightarrow \quad x_2 = \frac{x_0 + x_1}{2} = 1.5$$

$$\left.\begin{array}{r} f(1.5) = 1.5 - e^{2/3} = f(x_2) = -0.4477 \\ f(x_0)\, f(x_2) > 0 \end{array}\right| \quad \therefore \quad \text{root in } [x_2,\ x_1] \text{ i.e. in } [1.5,\ 2]$$

$$
\begin{aligned}
x_3 &= \frac{x_2 + x_1}{2} = \frac{1.5 + 2}{2} = 1.75 \\
f(x_3) &= 1.75 - e^{1/1.75} = -0.0208 \\
f(x_3)\, f(x_2) &> 0 \therefore \text{root in } [x_3,\ x_1] \text{ i.e. in } [1.75,\ 2]
\end{aligned}
$$

$$
\begin{aligned}
x_4 &= \frac{1.75 + 2}{2} = \frac{3.75}{2} = 1.875 \\
f(x_4) &= f(1.875) = 0.1704 \\
f(x_3)\, f(x_4) &< 0 \therefore \text{root in } [x_3,\ x_4] \text{ i.e. in } [1.75,\ 1.875]
\end{aligned}
$$

$$
\begin{aligned}
x_5 &= \frac{x_3 + x_4}{2} = \frac{1.75 + 1.875}{2} = 1.8125 \\
f(x_5) &= f(1.8125) = 0.0763 \\
f(x_3)\, f(x_5) &< 0
\end{aligned}
$$

Continuing in this way we have we find the root in $[x_9,\ x_{10}]$ i.e. in $[1.761718,\ 1.7636715]$

$$
\begin{aligned}
x_{11} &= \frac{x_9 + x_{10}}{2} = 1.76269 \\
&= 1.763 \text{ to 3 decimal places}
\end{aligned}
$$

Newton's Method

Newton's method is an *iterative* method for root finding. That is, starting from some guess at the root, x_0, one iteration of the algorithm produces a number x_1, which is supposed to be closer to a root; guesses x_2, x_3, ..., x_n follow identically.

We know from Taylor that

$$f(x+h) = f(x) + f'(x)\,h + O\left(h^2\right).$$

This approximation is better when $f''(.)$ is "well-behaved" between x and $x+h$. Newton's method attempts to find some h such that

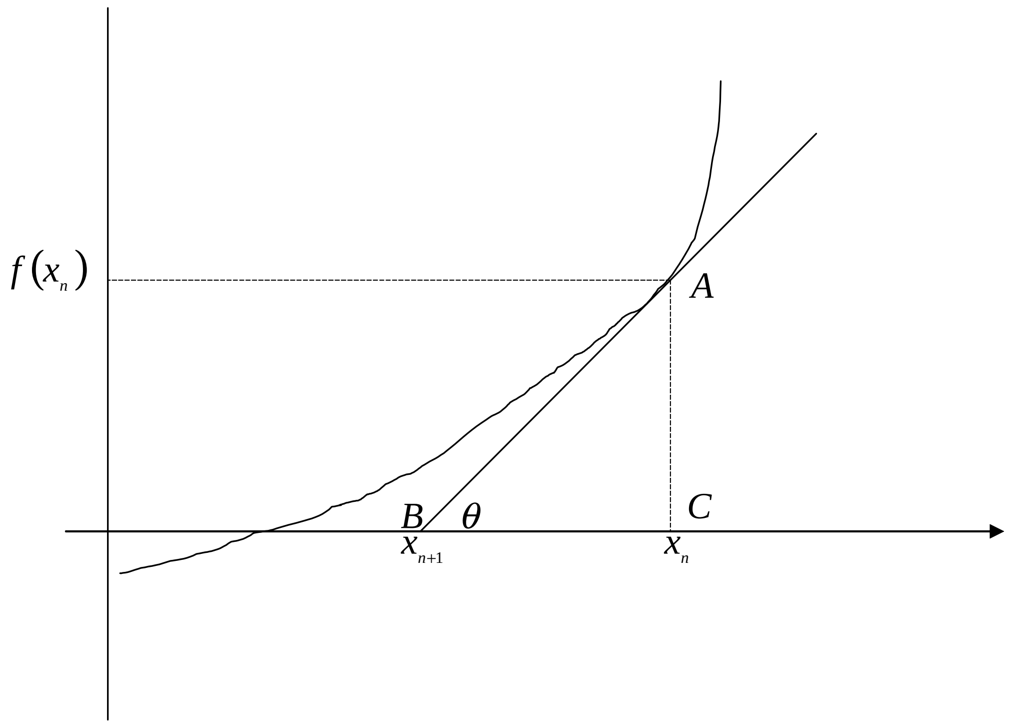
$$0 = f(x+h) = f(x) + f'(x)\,h.$$

This is easily solved as

$$h = \frac{-f(x)}{f'(x)}$$

An iteration of Newton's method, then, takes some guess x_n and returns x_{n+1} defined by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$



From above we see that $\tan\theta = \dfrac{AC}{BC} = \dfrac{f(x_n)}{(x_n - x_{n+1})}$

But

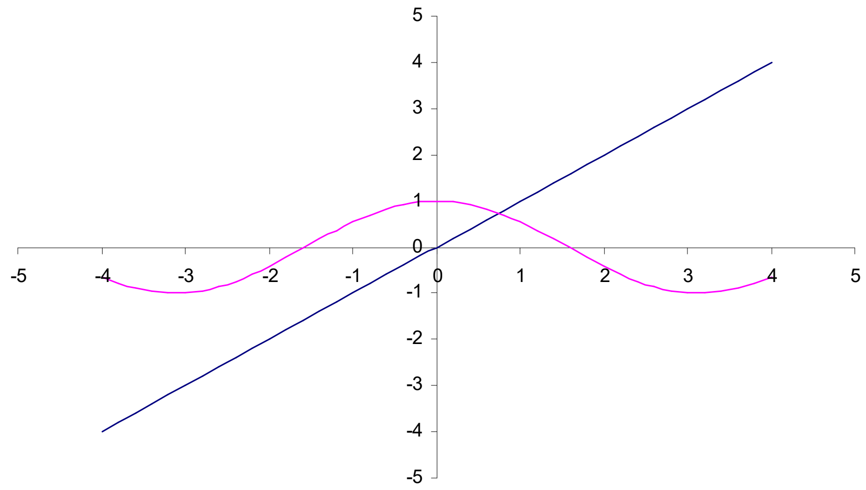
$$\begin{aligned}\tan\theta &= f'(x_n)\\ f'(x_n) &= \frac{f(x_n)}{(x_n - x_{n+1})}\end{aligned}$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

This is the *Newton-Raphson Technique.*

Example: Solve for roots, the function $f(x) = x - \cos x$.

Start by considering $x = \cos x$. That is draw $y = x$ and $y = \cos x$ to obtain an initial guess for the root(s).



Clearly the diagram above shows that there is only one root $\alpha \in (0,1)$.

We use the Newton formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \ldots.$$

where $n = 0$ is the initial guess. $f(x_n) = x_n - \cos x_n \longrightarrow f'(x_n) = 1 + \sin x_n$.

x	0	1
$f(x)$	-1	0.75

so numerically we also see that $f(0)\,f(1) < 0 \Longrightarrow \alpha \in (0,1)$. NR formula for this function becomes

$$x_{n+1} = x_n - \frac{x_n - \cos x_n}{1 + \sin x_n}, \quad x_0 = 1$$

$$\begin{aligned}
x_1 &= x_0 - \frac{x_0 - \cos x_0}{1 + \sin x_0} = 0.75036 \\
x_2 &= 0.75036 - \frac{0.75036 - \cos 0.75036}{1 + \sin 0.75036} = 0.73911 \\
x_3 &= 0.73911 - \frac{0.73911 - \cos 0.73911}{1 + \sin 0.73911} = 0.73909 \\
x_4 &= 0.73909 - \frac{00.73909 - \cos 0.73909}{1 + \sin 0.73909} = 0.73909
\end{aligned}$$

which gives the root $\alpha \approx 0.73909$.

Problems

As mentioned above, convergence is dependent on $f(x)$, and the initial estimate x_0. A number of conceivable problems might come up. We illustrate them here.

Example Consider Newton's method applied to the function $f(x) = \dfrac{\ln x}{x}$, with initial estimate $x_0 = 3$. Note that $f(x)$ is continuous on $\mathbb{R}^+$. It has a single root at $x = 1$. Our initial guess is not too far from this root. However, consider the derivative:

$$f'(x) = \frac{1 - \ln x}{x^2}$$

If $x > e^1$, then $1 - \ln x < 0$, and so $f'(x) < 0$. However, for $x > 1$, we know $f(x) > 0$. Thus taking

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} > x_n$$

The estimates will diverge from the root $x = 1$.

Stochastic Volatility

An observation when pricing derivatives is the fact that volatility of an asset price is anything but constant. We have seen in the much celebrated Black-Scholes framework that the assumptions do not consider these market features. Volatility does not behave how the Black–Scholes equation would like it to behave; it is not constant, it is not predictable, it's not even directly observable. Volatility is difficult to forecast - although not impossible.

This makes it a prime candidate for modelling as a random (stochastic) variable. There are many economic, empirical, and mathematical reasons for choosing a model with such a form. Empirical studies have shown that an asset's log-return distribution is non-Gaussian. It is characterised by heavy tails and high peaks (leptokurtic). There is also empirical evidence and economic arguments that suggest that equity returns and implied volatility are negatively correlated (also termed 'the leverage effect').

These reasons has been cited as evidence for non-constant volatility.

Stochastic volatility models were first introduced by Hull and White (1987), Scott (1987) and Wiggins (1987) to overcome the drawbacks of the Black, Scholes and Merton (1973). So it seems plausible to model volatility as a stochastic process. The method gives more parameters to fit, hence popular for calibration purposes.

These are systems of bi-variate SDEs. We continue to assume that S satisfies GBM

$$dS = \mu S dt + \sigma S dW_1,$$

but we further assume that volatility σ satisfies an arbitrary SDE

$$d\sigma = p(S, \sigma, t)dt + q(S, \sigma, t)dW_2.$$

Here both drift and diffusion are arbitrary, with $q(S, \sigma, t)$ being volatility of the volatility (vol of vol).

The two increments dW_1 and dW_2 have a correlation of ρ

$$\mathbb{E}^{\mathbb{P}}\left[dW_1 dW_2\right] = \rho dt.$$

Here $\mathbb{P}$ represents the physical measure. The choice of functions $p(S, \sigma, t)$ and $q(S, \sigma, t)$ is crucial to the evolution of the volatility, and thus to the pricing of derivatives.

The value of an option with stochastic volatility is a function of three variables, $V(S, \sigma, t)$.

Let's do the general theory first and then think about specific forms for p and q.

The pricing equation

The new stochastic quantity that we are modelling, the volatility, is not a traded asset. So as with the spot rate we cannot hold volatility. Thus, when volatility is stochastic we are faced with the problem of having a source of randomness that cannot be easily hedged away.

Because we have two sources of randomness we must hedge our option with two other contracts, one being the underlying asset as usual, but now we also need another option to hedge the volatility risk.

We therefore must set up a portfolio containing one option, with value denoted by $V(S, \sigma, t)$, a quantity $-\Delta$ of the asset and a quantity $-\Delta_1$ of another option with value $V_1(S, \sigma, t)$.

We have

$$\Pi = V - \Delta S - \Delta_1 V_1$$

The change in this portfolio in a time dt is given by

$$\begin{aligned} d\Pi &= \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma q S \frac{\partial^2 V}{\partial S \partial \sigma} + \tfrac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2} \right) dt \\ & -\Delta_1 \left(\frac{\partial V_1}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 V_1}{\partial S \partial \sigma} + \tfrac{1}{2} q^2 \frac{\partial^2 V_1}{\partial \sigma^2} \right) dt \\ & + \left(\frac{\partial V}{\partial S} - \Delta_1 \frac{\partial V_1}{\partial S} - \Delta \right) dS \\ & + \left(\frac{\partial V}{\partial \sigma} - \Delta_1 \frac{\partial V_1}{\partial \sigma} \right) d\sigma. \end{aligned}$$

where a higher dimensional form of Itô has been used on functions of S, σ and t.

To eliminate all randomness from the portfolio we must choose

$$\frac{\partial V}{\partial S} - \Delta_1 \frac{\partial V_1}{\partial S} - \Delta = 0,$$

to eliminate the dS terms, which are the sources of randomness, and

$$\frac{\partial V}{\partial \sigma} - \Delta_1 \frac{\partial V_1}{\partial \sigma} = 0,$$

to get rid off $d\sigma$ terms.

Therefore our choice of delta terms to make the portfolio risk free become

$$\Delta_1 = \frac{\frac{\partial V}{\partial \sigma}}{\frac{\partial V_1}{\partial \sigma}}$$

and

$$\Delta = \frac{\partial V}{\partial S} - \frac{\frac{\partial V}{\partial \sigma}}{\frac{\partial V_1}{\partial \sigma}} \frac{\partial V_1}{\partial S}.$$

This leaves us with

$$
\begin{aligned}
d\Pi &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma q S \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2} \right) dt \\
&\quad - \Delta_1 \left(\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 V_1}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V_1}{\partial \sigma^2} \right) dt \\
&= r\Pi dt = r\left(V - \Delta S - \Delta_1 V_1\right) dt,
\end{aligned}
$$

where we have used arbitrage arguments to set the return on the portfolio equal to the risk-free rate.

As it stands this is one equation in the two unknowns V and V_1.

This contrasts with the earlier Black–Scholes case with one equation in the one unknown - but presents the same type of problem when deriving the bond pricing equation.

Collecting all V terms on the left-hand side and all V_1 terms on the right-hand side we find that

$$\frac{\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma qS\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2}+rS\frac{\partial V}{\partial S}-rV}{\frac{\partial V}{\partial\sigma}}$$

$$=\quad\frac{\frac{\partial V_1}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V_1}{\partial S^2}+\rho\sigma qS\frac{\partial^2 V_1}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V_1}{\partial\sigma^2}+rS\frac{\partial V_1}{\partial S}-rV_1}{\frac{\partial V_1}{\partial\sigma}}$$

We are lucky that the left-hand side is a functional of V but not V_1 and the right-hand side is a function of V_1 but not V.

Therefore both sides can only be functions of the independent variables, S,σ and t. So set both sides equal to

$$f\left(S,\sigma,t\right).$$

Thus we have

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma Sq\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2}+rS\frac{\partial V}{\partial S}-rV=-\left(p-\lambda q\right)\frac{\partial V}{\partial\sigma},$$

for some function $\lambda\left(S, \sigma, t\right)$.

Reordering this equation, we usually write

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma S q\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2}+rS\frac{\partial V}{\partial S}+(p-\lambda q)\frac{\partial V}{\partial\sigma}-rV=0.$$

The function $\lambda(S, \sigma, t)$ is called the *market price of* (*volatility*) *risk*.

The market price of volatility risk

If we can solve the pricing equation on the previous slide then we have found the value of the option, and the hedge ratios.

But note that we find two hedge ratios, $\frac{\partial V}{\partial S}$ and $\frac{\partial V}{\partial \sigma}$.

• We have two hedge ratios because we have two sources of randomness that we must hedge away.

Because one of the modelled quantities, the volatility, is not traded we find that the pricing equation contains a market price of risk term.

What does this term mean?

Let's see what happens if we only hedge to remove the stock risk.

Suppose we hold one of the option with value V , and satisfying the pricing equation, delta hedged with the underlying asset only i.e. we have

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

The change in this portfolio value is

$$\begin{aligned} d\Pi &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma q S \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2} \right) dt \\ &\quad + \left(\frac{\partial V}{\partial S} - \Delta \right) dS + \frac{\partial V}{\partial \sigma} d\sigma. \end{aligned}$$

Because we are delta hedging the coefficient of dS is zero, leaving

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma q S \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2} \right) dt + \frac{\partial V}{\partial \sigma} d\sigma.$$

Now from the pricing PDE we have

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma S q \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2} = -rS\frac{\partial V}{\partial S} - (p - \lambda q)\frac{\partial V}{\partial \sigma} + rV.$$

We find that

$$d\Pi - r\Pi dt =$$

$$\begin{aligned} & \left(-rS\frac{\partial V}{\partial S} - (p - \lambda q)\frac{\partial V}{\partial \sigma} + rV \right) dt + \tfrac{\partial V}{\partial \sigma} d\sigma - r\left(V - \frac{\partial V}{\partial S} S \right) dt \\ = \quad & -(p - \lambda q)\frac{\partial V}{\partial \sigma} dt + \frac{\partial V}{\partial \sigma}(p dt + q dW_2) \end{aligned}$$

Now simplifying this last term gives

$$\lambda q \frac{\partial V}{\partial \sigma} dt + \frac{\partial V}{\partial \sigma} q dW_2$$

Observe that for every unit of volatility risk, represented by dW_2, there are λ units of extra return, represented by dt. Hence the name 'market price of risk.'

The return on this partially hedged portfolio in excess of the risk-free return is

$$q\frac{\partial V}{\partial \sigma}\left(\lambda dt + dW_2\right)$$

Returning to the pricing equation

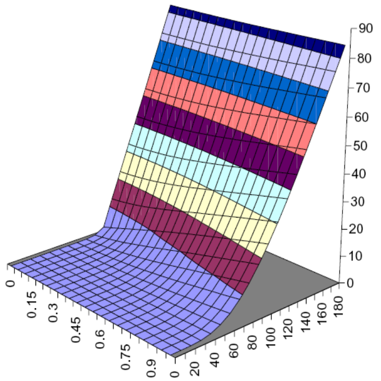
$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma S q\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2}+rS\frac{\partial V}{\partial S}+(p-\lambda q)\frac{\partial V}{\partial\sigma}-rV=0.$$

The quantity $p-\lambda q$ is called the risk-neutral drift rate of the volatility.

Recall that the risk-neutral drift of the underlying asset is r and not μ.

When it comes to pricing derivatives, it is the risk-neutral drift that matters and not the real drift, whether it is the drift of the asset or of the volatility.

stochastic volatility: an example for particular value of $p, q,\ \rho$



The option price is shown for varying stock and volatility.

This is a snapshot at a fixed point in time. We notice it looks like a typical European option.

Note for larger σ we have greater curvature (i.e. larger diffusion).

In addition to the model for GBM we have SDE for volatility, where $v = \sigma^2$.

The equations look nicer expressed in terms of the variance (important quantity).

Many volatility models are of the form

$$dv = A(v)\, dt + cv^{\gamma} dW_2,$$

for some value γ and mean reverting drift $A(v)$, where the variance $v = \sigma^2$.

In the presence of a continuous dividend yield, the earlier PDE can be written as

$$\frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho \sqrt{v} S q \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2} + (r - D)\, S \frac{\partial V}{\partial S} + cv^{\gamma} \frac{\partial V}{\partial \sigma} - rV = 0.$$

Popular Models

GARCH - diffusion: Generalized Autoregressive Conditional Heteroskedasticity.

A commonly used popular discrete time model in econometrics. It can be turned into the continuous time limit of many GARCH-processes by the following SDE

$$dv = (a - bv)\,dt + cvdW_2.$$

The popularity lies in the ease with which the positive valued parameters a, b and c can be estimated, hence allowing the pricing of options.

There is a mean reverting drift with speed b and mean rate a/b. c is the *vol of vol* which sets the scale for the random nature of volatility. In the case $a = b = 0$, the GARCH diffusion model reduces to the log-normal process without drift in the Hull and White (1987) model.

Given $v(0) > 0$,

$$v(t) = v(0)\, e^{-\left(b+\frac{1}{2}c^2\right)t+cW_t} + a\int_0^t e^{-\left(b+\frac{1}{2}c^2\right)(s-t)+c(W_t-W_s)} ds.$$

Heston:

Takes

$$dv = \gamma\,(m - v)\,dt + \xi\sqrt{v}dW_2$$

Also called the square root model because of the term in the diffusion - which gives a closed form solution, hence the popularity. This means it is easier to calibrate. Heston takes $\rho \neq 0$. In this model the the process is proportional to the square root of its level.

Must be comfortable with Complex Analysis Methods, as it requires the use of Fourier Transforms.

3/2 model:

Pronounced the *three-halves model* because of the $3/2$ power in the diffusion.

$$dv = v\,(a - bv)\,dt + cv^{3/2} dW_2$$

Again mean reverting - the existence of a Closed-form solution makes it a popular model. But note the mean reverting and volatility parameters are now stochastic.

See Alan Lewis' book on *Option valuation under stochastic volatility*, where he presents analytical solutions for this model.

This does a supposedly better job of calibrating than Heston, although Heston is more popular.

Hull & White

$$\frac{dv}{v} = \mu dt + \xi dW_2$$

No mean reversion. They take $\rho = 0$. Note the lognormal structure hence it can grow indefinitely.

Stein & Stein

$$d\sigma = -\theta(\sigma - m)\,dt + \xi dW_2$$

The model allows mean-reversion but σ can become negative. They take $\rho = 0$.

Ornstein-Uhlenbeck process:

This model is expressed in terms of the log of the variance.

Writing $y = \log x$

$$dy = (a - by)\, dt + c dW_2$$

Already seen the O-U-P interest rate model (looks very similar). This model matches data well.

This has a steady state distribution which is lognormal.

A closed form solution does not exist so requires numerical treatment.

The Heston Model

In his model the variance follows a mean-reverting square root process, first used by Cox-Ingersoll-Ross in 1985 to capture the dynamics of the spot rate where the mean reversion rate $m > 0$, and the speed $\gamma > 0$. The vol of vol $\xi > 0$.

$$dS = (\mu - D)\, S dt + \sqrt{v} S dW_1,$$

$$dv = \gamma (m - v)\, dt + \xi \sqrt{v} dW_2$$

Solving problems numerically is simple (FDM or Monte Carlo). In the case of MC take the stock drift as $(r - D)$.

In order for the mean-reverting square root dynamics for the variance to remain positive, there are a number of analytical results available. In particular is the Feller condition, i.e. if

$$\gamma m \geq \xi^2$$

then the variance process cannot become negative. If this condition is not satisfied then the origin is attainable and strongly rejecting so that the variance process may attain zero in finite time, without spending time at this point.

In deriving the PDE for Heston, he takes

$$f (S, v, t) = -\gamma (m - v) + \Lambda (S, v, t)\, \xi \sqrt{v}$$

giving the following pricing PDE for the option $U(S, v, t)$

$$\frac{\partial U}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 U}{\partial S^2} + \rho \xi v S \frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2} \xi^2 v \frac{\partial^2 U}{\partial v^2} +$$

$$\left(\gamma (m - v) - \Lambda (S, v, t) \sigma \sqrt{v}\right) \frac{\partial U}{\partial v} + rS \frac{\partial U}{\partial S} - rU = 0$$

Consider the pricing of a call option subject to the final condition $C(S, v, T) = \max(S_T - E, 0)$ with the following boundary conditions

$$\begin{aligned} C(0, v, t) &= 0 \\ \lim_{S \to \infty} \frac{\partial C}{\partial S}(S, v, t) &= 1 \\ \frac{\partial C}{\partial t} + rS \frac{\partial C}{\partial S} + \gamma m \frac{\partial C}{\partial v} &= rC \\ \lim_{v \to \infty} C(S, v, t) &= S \end{aligned}$$

How to use Heston

There are <u>four</u> parameters in the model, speed of mean reversion, level of mean reversion, volatility of volatility, correlation. That is b, $a/b, c, \rho$ respectively.

And also potentially a market price of volatility risk parameter.

The main four parameters can be chosen by matching data or by calibration.

Experience suggests that calibrated parameters are very unstable, and often unreasonable. (For example, the best fit to market prices might result in a correlation of exactly -1.)

Consider calibrating. Suppose

Parameters	Today	Next week
$a =$	14	-487
$b =$	29	$\sqrt{-12}$
$c =$	0.01	1000
$\rho =$	-0.6	-3

so a somewhat exaggerated sarcastic example, but nevertheless shows that when recalibrating it hasn't worked - the parameters which were fixed are totally different!

The Heston model with jumps

Increasingly popular are stochastic volatility with jumps models (SVJ).

Jump models require a parameter to measure probability of a jump (a Poisson process) and a distribution for the jumps.

Also have SVJJ - jumps in the stock and jumps in the volatility.

Pros: More parameters allow better fitting. The jump component of the model has most impact over short time scales.

Therefore use longer-dated options to fit the stochastic volatility parameters and the shorter-dated options to fit the jump component.

Cons: Mathematics slightly more complicated (and again we must work in the transform domain).

Hedging is even harder when the underlying stock process is potentially discontinuous.

People also looking at stochastic correlation models.

Whilst there is no such thing as the perfect model, you can always pretend to have the ideal one by introducing more parameters.

More parameters means more quantities to calibrate.

Case Study: The REGARCH model and its diffusion limit

REGARCH = Range-based Exponential GARCH

Although a closed form solution does not exist, a fairly nice model which looks very plausible.

'Range-based' refers to the use of the daily range, defined as the difference between the highest and lowest log asset price recorded throughout the day, rather than simply the closing prices.

'Exponential' refers to modelling the logarithm of the variance.

Diffusion limits exist for all GARCH-type of processes. That is, they can be expressed in continuous time using stochastic differential equations.

(This is achieved via ‘moment matching.’ The statistical properties of the discrete-time GARCH processes are recreated with the continuous-time SDEs.)

REGARCH is another econometrics discrete time model, but can be turned into the following three-factor model:

$$
\begin{aligned}
dS &= \mu S dt + \sigma_1 S dW_0 && \text{(a)} \\
d\left(\log \sigma_1\right) &= a_1 \left(\log \sigma_2 - \log \sigma_1\right) dt + b_1 dW_1 && \text{(b)} \\
d\left(\log \sigma_2\right) &= a_2 \left(c_2 - \log \sigma_2\right) dt + b_2 dW_2. && \text{(c)}
\end{aligned}
$$

This is a three-factor (higher dimensional) model, with two volatilities.

σ_1 represents the actual (short term) volatility of the asset returns, which is stochastic.

The σ_2 represents the (longer term) level to which σ_1 reverts, and is itself stochastic.

What are the dynamics of this model?

We have the usual GBM random walk for the stock given by (a) which has actual volatility σ_1. This is short term.

Note from (b) that the **log** of σ_1 mean reverts to $\log \sigma_2$. So rather than σ_2 being constant, it is fluctuating and σ_1 is chasing that.

From (c) we observe that σ_2 reverts to a constant mean c_2.

For pricing options we must replace these SDEs with the risk-neutral versions:

$$\begin{aligned} dS &= rS dt + \sigma_1 S dW_0 \\ d(\log \sigma_1) &= a_1 (\log \sigma_2 - \log \sigma_1 - \lambda b_1/a_1)\, dt + b_1 dW_1 \\ d(\log \sigma_2) &= a_2 (c_2 - \log \sigma_2 - \lambda b_2/a_2)\, dt + b_2 dW_2. \end{aligned}$$

The λ terms represent the market prices of risk.

The a and b coefficients and the correlations between the three sources of randomness give this system seven parameters.

These parameters are related to the parameters of the original REGARCH model and can be estimated from asset data.

Example: let's look at some parameters.

$a_1 = 56.6$, $b_1 = 1.138$, $a_2 = 2.82$, $b_2 = 0.388$, $c_2 = -1.25$. ($\lambda_1 = \lambda_2 = 0$.)

σ_1 is very rapidly mean reverting to the level of σ_2. This is a 'short-term' volatility. The time scale for mean reversion is about one week.

σ_2, the 'long-term volatility, reverts more slowly, over a period of about six months.

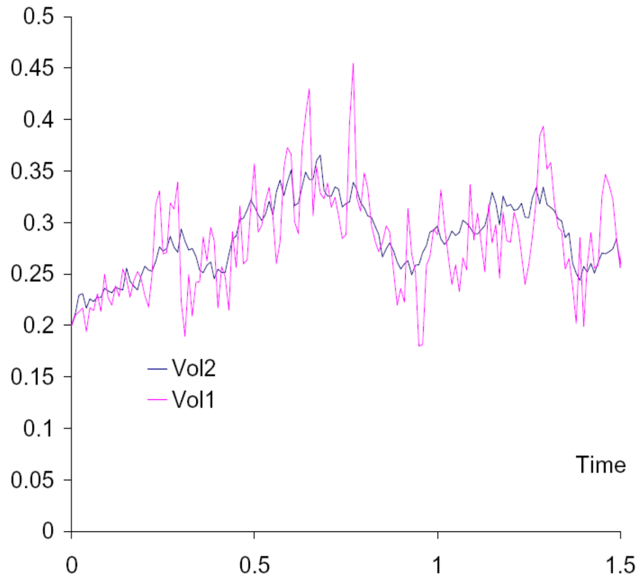
a_1 is the speed for $\log \sigma_1$. The bigger this is, the faster the reversion to $\log \sigma_2$.

$a_1 dt$ is non-dimensional therefore a_1 has dimensions of $1/\text{time} \Longrightarrow 1/a_1$ has dimensions of time.

So a time scale of approximately 1 week, for $\log \sigma_1$ to mean revert.

$b_1 \gg b_2$, volatility of $d(\log \sigma_1)$ much greater than $d(\log \sigma_2) -$ which it is chasing.

$1/a_2$ is approximately 0.5 years, so it takes $\log \sigma_2$ 6 months to revert back to its (long term) mean.



How do you solve these equations?

- Monte Carlo: The solutions of the two-factor partial differential equations you get with stochastic volatility models can still be interpreted as 'the present value of the expected payoff.' So all you have to do is to simulate the relevant random walks for the underlying and volatility (risk neutral) many times, calculate the average payoff and then present value it.

- Finite differences: The partial differential equations can still be solved by finite differences but you will need to work with a three-dimensional grid.

Pros and cons of stochastic volatility models

Pros:

- Evidence (and common sense) suggests that volatility changes, possibly randomly

- More parameters means that calibration can be 'better'

Cons:

- As with any incomplete-market model hedging is only possible if you believe in the market price of (volatility) risk

Further Finite-difference Methods for One-factor Models

In this lecture. . . Implicit methods

fully
- implicit finite-difference methods including Crank–Nicolson → θ-method
- Douglas schemes
- Richardson extrapolation
- American-style exercise and exotic options
- 2 factor model

By the end of this lecture you will

- know several more ways of solving parabolic partial differential equations numerically

$$\underbrace{\frac{\partial V}{\partial t} \qquad \frac{\partial^2 V}{\partial S^2}}_{\text{parabolic PDE}}$$

Introduction

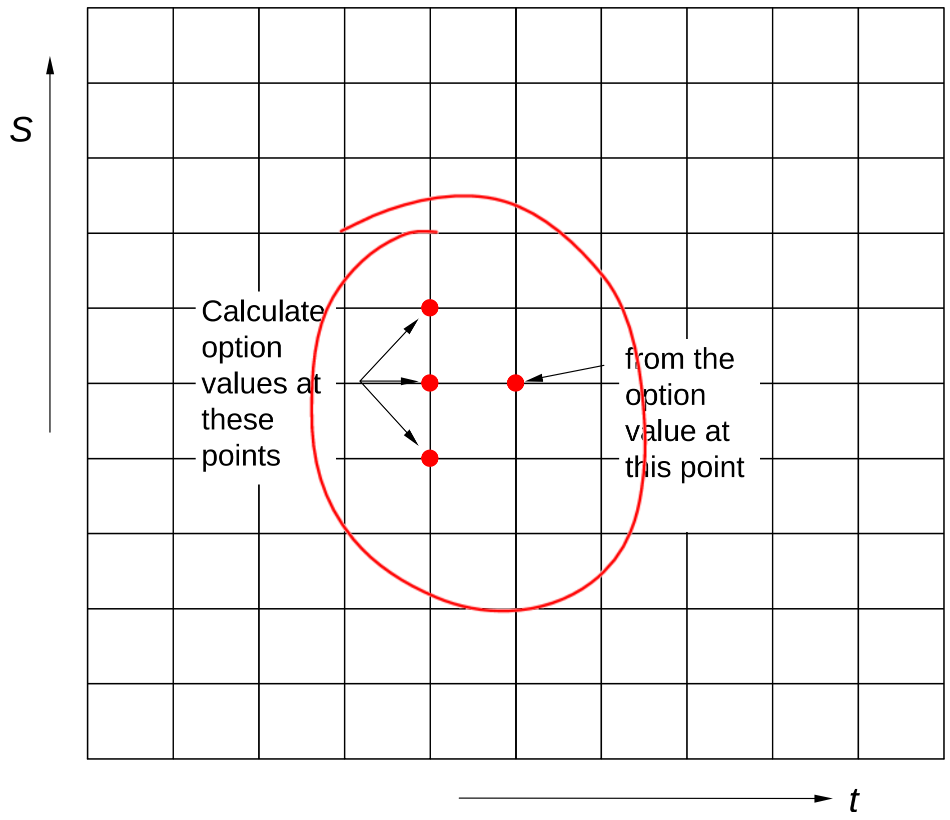
There are many more ways of solving parabolic partial differential equations than the explicit method.

The more advanced methods are usually more complicated to program but have advantages in terms of stability and speed.

& order of accuracy

Implicit finite-difference methods

The **fully implicit method** uses the points shown in the figure below to calculate the option value.



$\tau = T - t$ $\quad T - k\,\delta t$

The relationship between the option values on the mesh is simply

$\frac{1}{2}\sigma^2 i^2 \cancel{\delta S^2}$

Known

$$\frac{\boxed{V_i^k} - V_i^{k+1}}{\delta t} + a_i^{k+1}\left(\frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\cancel{\delta S^2}}\right)$$

$\frac{\partial V}{\partial t} \sim \frac{V_i^k - V_i^{k+1}}{\delta t}$ $\quad (r-D)\,i\,\cancel{\delta S}$ $\quad -r$

$$+ b_i^{k+1}\left(\frac{V_{i+1}^{k+1} - V_{i-1}^{k+1}}{2\cancel{\delta S}}\right) + c_i^{k+1} \underbrace{V_i^{k+1}} = 0.$$

The method is accurate to $O(\delta t, \delta S^2)$.

$\frac{\partial^2 V}{\partial S^2} \sim \frac{V_{i-1}^{k+1} - 2V_i^{k+1} + V_{i+1}^{k+1}}{\delta S^2}$

$\frac{\partial V}{\partial S} \sim \frac{V_{i+1}^{k+1} - V_{i-1}^{k+1}}{2\,\delta S}$

This can be written as

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

where

$$A_i^{k+1} = -\nu_1 a_i^{k+1} + \tfrac{1}{2}\nu_2 b_i^{k+1},$$

$$B_i^{k+1} = 2\nu_1 a_i^{k+1} - \delta t\, c_i^{k+1}$$

and

$$C_i^{k+1} = -\nu_1 a_i^{k+1} - \tfrac{1}{2}\nu_2 b_i^{k+1}$$

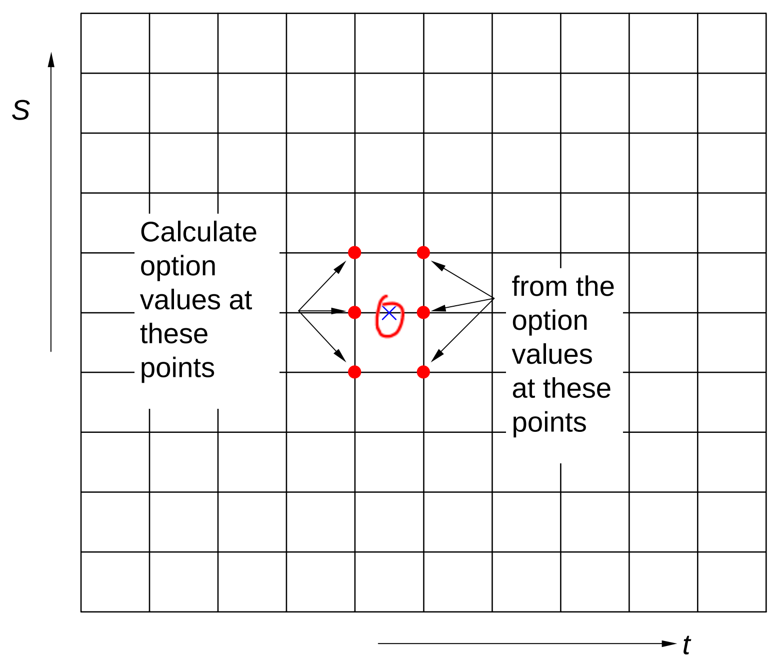
where

$$\nu_1 = \frac{\delta t}{\delta S^2} \quad \text{and} \quad \nu_2 = \frac{\delta t}{\delta S}.$$

Equation (1) does not hold for $i = 0$ or $i = I$, the boundary conditions supply the two remaining equations.

The Crank–Nicolson method

The **Crank–Nicolson method** can be thought of as an average of the explicit method and the fully implicit method. It uses the six points shown in the figure below.



The Crank–Nicolson scheme is

H

Imp Exp

$$\frac{V_i^k - V_i^{k+1}}{\delta t} + \frac{a_i^{k+1}}{2}\left(\frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\delta S^2}\right) + \frac{a_i^k}{2}\left(\frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}\right)$$

$\frac{\partial V}{\partial t}$ Γ

$$+\frac{b_i^{k+1}}{2}\left(\frac{V_{i+1}^{k+1} - V_{i-1}^{k+1}}{2\,\delta S}\right) + \frac{b_i^k}{2}\left(\frac{V_{i+1}^k - V_{i-1}^k}{2\,\delta S}\right)$$

Imp Exp

Δ

$$+\tfrac{1}{2}c_i^{k+1}V_i^{k+1} + \tfrac{1}{2}c_i^k V_i^k = O(\delta t^2, \delta S^2).$$

Imp Exp

$-rV$

k , $k+1$

This can be written as

Implicit unknown

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

$$= \underbrace{A_i^k V_{i-1}^k + (1 + B_i^k)V_i^k + C_i^k V_{i+1}^k}_{\text{Known.}},$$

where

$$A_i^k = \tfrac{1}{2}\nu_1 a_i^k - \tfrac{1}{4}\nu_2 b_i^k,\; B_i^k = -\nu_1 a_i^k + \tfrac{1}{2}\delta t\, c_i^k,\; C_i^k = \tfrac{1}{2}\nu_1 a_i^k + \tfrac{1}{4}\nu_2 b_i^k.$$

The beauty of this method lies in its stability and accuracy. The error in the method is $O(\delta t^2, \delta S^2)$.

The Crank–Nicolson method can be written in the matrix form

coeff matrix

$$\begin{pmatrix} -A_1^{k+1} & 1-B_1^{k+1} & -C_1^{k+1} & 0 & \cdot & \cdot & \cdot \\ 0 & -A_2^{k+1} & 1-B_2^{k+1} & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1-B_{I-2}^{k+1} & -C_{I-2}^{k+1} & 0 \\ \cdot & \cdot & \cdot & 0 & -A_{I-1}^{k+1} & 1-B_{I-1}^{k+1} & -C_{I-1}^{k+1} \end{pmatrix} \begin{pmatrix} V_0^{k+1} \\ V_1^{k+1} \\ \cdot \\ \cdot \\ \cdot \\ V_{I-1}^{k+1} \\ V_I^{k+1} \end{pmatrix}$$

LHS

$$\underbrace{= \begin{pmatrix} A_1^{k} & 1+B_1^{k} & C_1^{k} & 0 & \cdot & \cdot & \cdot \\ 0 & A_2^{k} & 1+B_2^{k} & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1+B_{I-2}^{k} & C_{I-2}^{k} & 0 \\ \cdot & \cdot & \cdot & 0 & A_{I-1}^{k} & 1+B_{I-1}^{k} & C_{I-1}^{k} \end{pmatrix} \begin{pmatrix} V_0^{k} \\ V_1^{k} \\ \cdot \\ \cdot \\ \cdot \\ V_{I-1}^{k} \\ V_I^{k} \end{pmatrix}}_{\text{RHS}}$$

coeff matrix

$$\boxed{A \underline{V}^{k+1} = \underline{V}^k}$$

vector

The two matrices have $I-1$ rows and $I+1$ columns. This is a representation of $I-1$ equations in $I+1$ unknowns. Our aim is to write this as

$$\mathbf{M}_L^{k+1}\mathbf{v}^{k+1} + \mathbf{r}^k = \mathbf{M}_R^k\mathbf{v}^k,$$

for known *square* matrices $\mathbf{M}_L^{k+1}$ and $\mathbf{M}_R^k$, and a known vector $\mathbf{r}^k$ and where details of the boundary conditions have been fully incorporated.

$S=0$ lower bdy $\rightarrow$ $S=0$

$S=S_\infty$ upper bdy.

$\leftarrow$ $S=S_\infty$

Example:

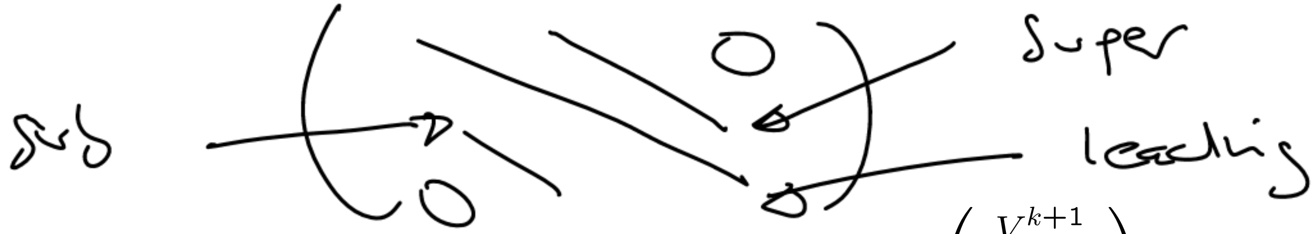
Sometimes we know that our option has a particular value on the boundary $i = 0$, or on $i = I$.

$S = 0$ $S = S_\infty$

For example, if we have a European put we know that $V(0, t) = Ee^{-r(T-t)}$.

This translates to knowing that $V_0^{k+1} = Ee^{-r(k+1)\delta t}$.

We can write



$$\begin{pmatrix} -A_1^{k+1} & 1-B_1^{k+1} & -C_1^{k+1} & 0 & \cdot & \cdot & \cdot \\ 0 & -A_2^{k+1} & 1-B_2^{k+1} & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1-B_{I-2}^{k+1} & -C_{I-2}^{k+1} & 0 \\ \cdot & \cdot & \cdot & 0 & -A_{I-1}^{k+1} & 1-B_{I-1}^{k+1} & -C_{I-1}^{k+1} \end{pmatrix} \begin{pmatrix} V_0^{k+1} \\ V_1^{k+1} \\ \cdot \\ \cdot \\ \cdot \\ V_{I-1}^{k+1} \\ V_I^{k+1} \end{pmatrix}$$

as

$$\begin{pmatrix} 1-B_1^{k+1} & -C_1^{k+1} & 0 & \cdot & \cdot & \\ -A_2^{k+1} & 1-B_2^{k+1} & \cdot & \cdot & \cdot & \\ & 0 & \cdot & \cdot & \cdot & 0 \\ & \cdot & \cdot & \cdot & 1-B_{I-2}^{k+1} & -C_{I-2}^{k+1} \\ & \cdot & \cdot & 0 & -A_{I-1}^{k+1} & 1-B_{I-1}^{k+1} \end{pmatrix} \begin{pmatrix} V_1^{k+1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ V_{I-1}^{k+1} \end{pmatrix} + \begin{pmatrix} A_1^{k+1} V_0^{k+1} \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ \cdot \end{pmatrix}$$

$$= \mathbf{M}_L^{k+1} \mathbf{v}^{k+1} + \mathbf{r}^k.$$

B.C.s are held in

The matrix $\mathbf{M}_L$ is square and of size $I-1$.

$A\underline{V}^{k+1} = \underline{V}^k$ $\qquad$ A^{-1} exists, r^k.

$[\underline{V}^{k+1} = A^{-1}\underline{V}^k]$

The matrix equation

Remembering that we know $\mathbf{v}^k$, the matrix multiplication and vector addition on the right-hand side is simple enough to do.

But how do we then find $\mathbf{v}^{k+1}$? In principle, the matrix $\mathbf{M}_L^{k+1}$ could be inverted to give

$$\mathbf{v}^{k+1} = (\mathbf{M}_L^{k+1})^{-1}(\mathbf{M}_R^k\mathbf{v}^k - \mathbf{r}^k),$$

except that matrix inversion is very time consuming, and from a computational point of view extremely inefficient.

$$\boxed{A\underline{x} = \underline{b}}$$

There are two much better ways for solving this system, called **LU decomposition** and **successive over-relaxation**.

Jacobi / Gauss-Seidel

Direct vs. Indirect

Iteration.

LU decomposition

The matrix $\mathbf{M}_L^{k+1}$ is special in that it is **tridiagonal**.

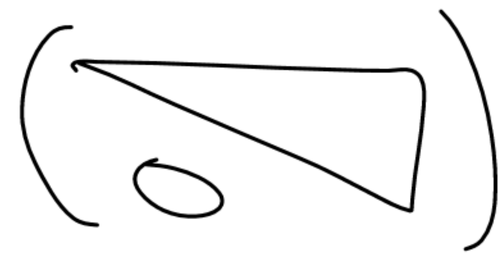
We can decompose the matrix into the product of two other matrices, one having non-zero elements along the diagonal and the subdiagonal and the other having non-zero elements along the diagonal and the superdiagonal.

Call these two matrices **L** and **U** respectively.

$$M\underline{v}^{k+1} = \underline{v}^{k}$$

$$M = LU$$

$L =$ 

$U =$ 

M

$$
\begin{pmatrix}
1+B_1 & C_1 & 0 & \cdot & \cdot & \cdot & 0 \\
A_2 & 1+B_2 & C_2 & 0 & \cdot & \cdot & \cdot \\
0 & A_3 & 1+B_3 & \cdot & \cdot & \cdot & \cdot \\
\cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & 1+B_{I-3} & C_{I-3} & 0 \\
\cdot & \cdot & \cdot & 0 & A_{I-2} & 1+B_{I-2} & C_{I-2} \\
\cdot & \cdot & \cdot & \cdot & 0 & A_{I-1} & 1+B_{I-1}
\end{pmatrix}
$$

$$
= \begin{pmatrix}
1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\
l_2 & 1 & 0 & 0 & \cdot & \cdot & \cdot \\
0 & l_3 & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & 1 & 0 & 0 \\
\cdot & \cdot & \cdot & 0 & l_{I-2} & 1 & 0 \\
\cdot & \cdot & \cdot & \cdot & 0 & l_{I-1} & 1
\end{pmatrix}
\begin{pmatrix}
d_1 & u_1 & 0 & \cdot & \cdot & \cdot & 0 \\
0 & d_2 & u_2 & 0 & \cdot & \cdot & \cdot \\
0 & 0 & d_3 & \cdot & \cdot & \cdot & \cdot \\
\cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & d_{I-3} & u_{I-3} & 0 \\
\cdot & \cdot & \cdot & 0 & 0 & d_{I-2} & u_{I-2} \\
\cdot & \cdot & \cdot & \cdot & 0 & 0 & d_{I-1}
\end{pmatrix}
$$

L U

ones

i.e.

$l_{ii} = 1$

$$\mathbf{M} = \mathbf{LU}$$

l, d, u unknowns.

$d_1 = 1 + B_1$

$u_1 = C_1$

It is not difficult to show that

$$d_1 = 1 + B_1$$

and then

$$l_i d_{i-1} = A_i, \quad u_{i-1} = C_{i-1} \quad \text{and} \quad d_i = 1 + B_i - l_i u_{i-1} \quad \text{for} \quad 2 \le i \le I - 1.$$

Now we exploit the decomposition to solve the original matrix equation. This equation is of the form

$$\mathbf{Mv} = \mathbf{q}$$

where we want to find $\mathbf{v}$.

We can write

$$\mathbf{LUv} = \mathbf{q}. \tag{2}$$

The vector $\mathbf{q}$ contains both the old option value array, at time step k, and details of the boundary conditions.

Solve Equation (2) in two steps. First find $\mathbf{w}$ such that

$$\mathbf{L}\mathbf{w} = \mathbf{q}$$

and then $\mathbf{v}$ such that

$$\mathbf{U}\mathbf{v} = \mathbf{w}.$$

And then we are done.

The first step gives

$$w_1 = q_1$$

and

$$w_i = q_i - l_i w_{i-1} \quad \text{for} \quad 2 \leq i \leq I - 1.$$

Finally

$$v_{I-1} = \frac{w_{I-1}}{d_{I-1}}$$

and

$$v_i = \frac{w_i - u_i v_{i+1}}{d_i} \quad \text{for} \quad I - 2 \geq i \geq 1.$$

$V(S, r, t)$

The advantages of the LU decomposition method

- It is quick

done once

- The decomposition need only be done once if the matrix **M** is independent of time

$V(S, t)$ $V(S, \sigma, t)$

The disadvantages of the LU decomposition method

- It is not immediately applicable to American options
- The decomposition must be done each time step if the matrix **M** is time dependent

higher dim^{n} problem

Successive over-relaxation, SOR

We now come to an example of an ‘indirect method.’

In this method we solve the equations iteratively.

Suppose that the matrix **M** in the matrix equation

$$\mathbf{M}\mathbf{v} = \mathbf{q}$$

Matrix Inversion problem

has entries M_{ij} then the system of equations can be written as

$$\begin{aligned}
&① \quad M_{11}v_1 + M_{12}v_2 + \cdots + M_{1N}v_N = q_1 \\
&② \quad M_{21}v_1 + M_{22}v_2 + \cdots + M_{2N}v_N = q_2 \\
&\qquad \cdots \\
&Ⓝ \quad M_{N1}v_1 + M_{N2}v_2 + \cdots + M_{NN}v_N = q_N
\end{aligned}$$

N x N system

where N is the number of equations, the size of the matrix.

÷ by M_{ii}

Rewrite this as

$$
\begin{aligned}
\cancel{M_{11}}v_1 &= \left[q_1 - (M_{12}v_2 + \cdots + M_{1N}v_N)\right]/M_{11} \leftarrow \\
M_{22}v_2 &= \left[q_2 - (M_{21}v_1 + \cdots + M_{2N}v_N)\right]/M_{22} \leftarrow \\
&\cdots \\
\cancel{M_{NN}}v_N &= \left(q_N - (M_{N1}v_1 + \cdots)\right)/M_{NN} \leftarrow
\end{aligned}
$$

Assuming we know everything on the RHS.

This system is easily solved *iteratively* using

$$
\begin{aligned}
v_1^{n+1} &= \frac{1}{M_{11}}\left(q_1 - (M_{12}v_2^n + \cdots + M_{1N}v_N^n)\right)\\
v_2^{n+1} &= \frac{1}{M_{22}}\left(q_2 - (M_{21}v_1^{n+1} + \cdots + M_{2N}v_N^n)\right)\\
&\cdots\\
v_N^{n+1} &= \frac{1}{M_{NN}}\left(q_N - (M_{N1}v_1^{n+1} + \cdots)\right)
\end{aligned}
$$

where the superscript n denotes the level of the iteration *and not the time step*.

This iteration is started from some initial guess $\mathbf{v}^0$.

This iterative method is called the **Jacobi method**.

n – iteration no. Start with n=0

Which is a guess of values

$\underline{V}^{(0)} = (0, 0, ---, 0)$

Using $n=0$ $n+1 \rightarrow V^1 \rightarrow V^2 \rightarrow V^3$

When we implement the Jacobi method in practice we find some of the values of v_i^{n+1} before others.

In the **Gauss–Seidel** method we use the updated values as soon as they are calculated.

This method can be written as

current iteration

previous iteration

$$\underbrace{v_i^{n+1} = \frac{1}{M_{ii}}\left(q_i - \sum_{j=1}^{i-1} M_{ij} v_j^{n+1} - \sum_{j=i}^{N} M_{ij} v_j^{n}\right).}$$

Observe that there are some terms on the right-hand side with the superscript $n+1$.

These are values of v that were calculated earlier but at the same level of iteration.

When the matrix **M** has come from a finite-difference discretization of a parabolic equation (and that includes almost all finance problems) the above iterative methods usually converge to the correct solution *from one side*.

This means that the corrections $v_i^{n+1} - v_i^n$ stay of the same sign as n increases.

$\boxed{\omega}$

This is exploited in the **successive over-relaxation** or **SOR** method to speed up convergence.

$M\underline{v} = \underline{g}$ $\qquad$ $M\underline{v} - \underline{g} = \underline{0}$ exact

$\lim_{n\to\infty} \underline{r}^{(n)} \to \underline{0}$ $\qquad$ $M\underline{v} - \underline{g} = \underline{r}$ vector of residuals.

The method can be written as

$$\boxed{v_i^{n+1} = v_i^n + \frac{\omega}{M_{ii}}\left(q_i - \sum_{j=1}^{i-1} M_{ij} v_j^{n+1} - \sum_{j=i}^{N} M_{ij} v_j^n\right).}$$

Again, the new values for v_i are used as soon as they are obtained.

But now the factor ω, called the **acceleration** or **over-relaxation parameter** is included.

This parameter, which must lie between 1 and 2, speeds up the convergence to the true solution.

Under relaxed $\omega < 1$ method does not cge

over relaxed $\omega > 1$ speed up cgce.

The advantages of the SOR method

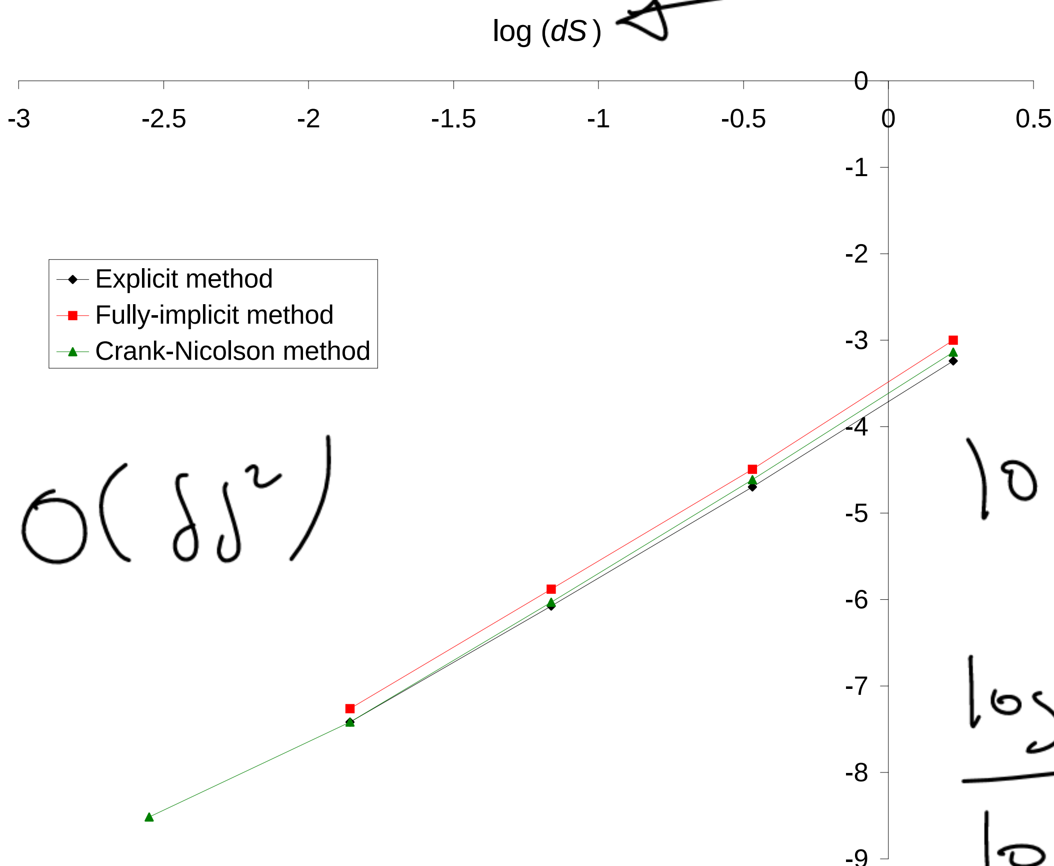
- It is easier to program than LU decomposition
- It is easily applied to American options

The disadvantage of the SOR method

- It is slightly slower than LU decomposition for European options

Comparison of finite-difference methods

Example: An at-the-money European call option, strike 20, three months to expiry, a volatility of 20% and an interest rate of 5%.



$2 \rightarrow O(\delta S^2)$

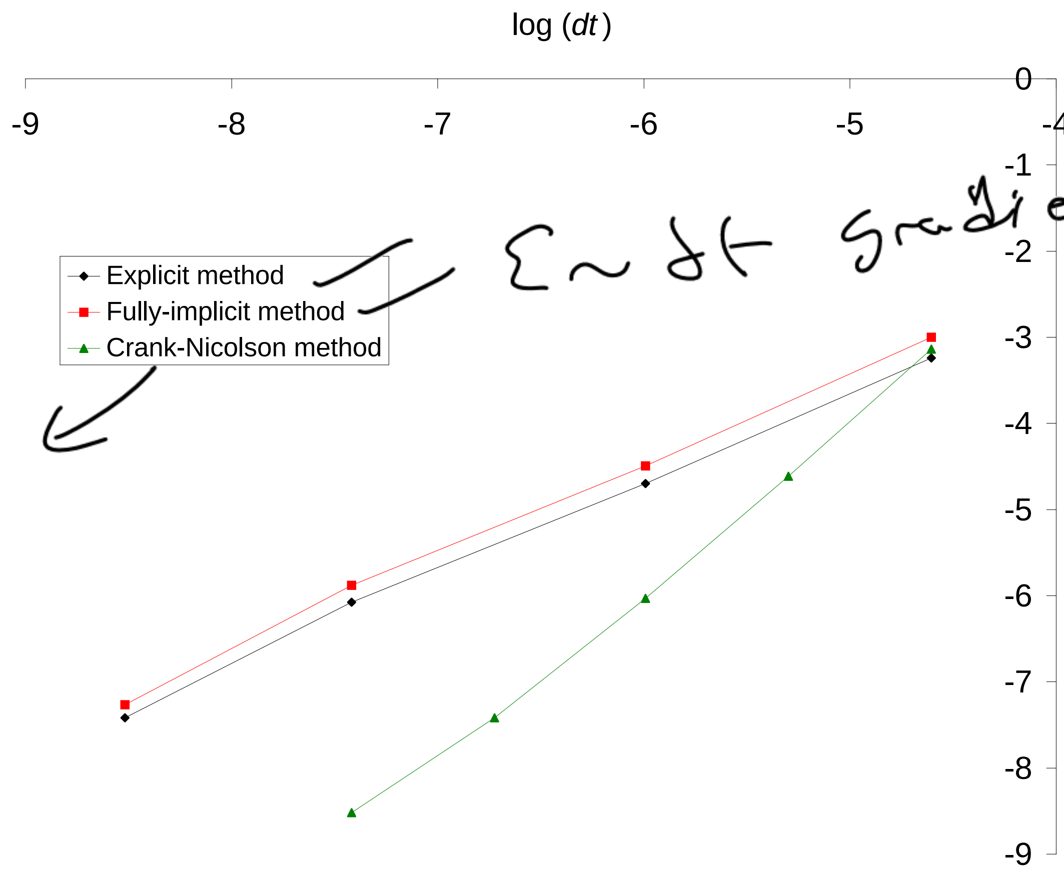
$\log \epsilon$

$$\frac{\log \epsilon}{\log dS} \approx 2$$

The logarithm of the error against the logarithm of the asset step size.

$\cancel{\log} \epsilon = \cancel{\log} dS^2$

$\epsilon \sim dS^2$



$\varepsilon \sim \delta t$ gradient of one

$\varepsilon \sim \delta t^2$ gradient of 2

The logarithm of the error against the logarithm of the time step.

Other methods

The questions that arise in any method are

- What is the error in the method in terms of δt and δS?
- What are the restrictions on the time step and/or asset step?

  Stability

- Can we solve the resulting difference equations quickly?

  Comp. efficiency.

- Is the method flexible enough to cope with changes in coefficients, boundary conditions etc.?

Douglas schemes

This is a method that manages to have a local truncation error of $O(\delta S^4, \delta t^2)$ for the same computational effort as the Crank–Nicolson scheme.

θ – method.

The basic diffusion equation is

$$\frac{\partial V}{\partial t} + \frac{\partial^2 V}{\partial S^2} = 0.$$

The explicit method applied to this equation is just

$$\frac{V_i^{k+1} - V_i^k}{\delta t} = \frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}.$$

and the fully implicit is similarly

$$\frac{V_i^{k+1} - V_i^k}{\delta t} = \frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\delta S^2}.$$

The Crank–Nicolson scheme is an average of these two methods. What about a *weighted* average?

This leads to the θ method.

Take a weighted average of the explicit and implicit methods to get...

$0 \leq \theta \leq 1$

$\frac{\partial V}{\partial t}$ Imp. Exp.

$$\frac{V_i^{k+1} - V_i^k}{\delta t} = \theta \left(\frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\delta S^2} \right) + (1-\theta) \left(\frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2} \right).$$

When $\theta = \frac{1}{2}$ we are back to the Crank–Nicolson method.

$\theta = 0$ Exp.

$\theta = \frac{1}{2}$ C–N

$\theta = 1$ F.I.

For a general value of θ the local truncation error is

$$O\left(\tfrac{1}{2}\delta t + \tfrac{1}{12}\delta S^2 - \theta\delta t, \delta S^4, \delta t^2\right).$$

When $\theta = 0$, $\frac{1}{2}$ or 1 we get the results we have seen so far.

But if

$$\theta = \frac{1}{2} - \frac{\delta S^2}{12\,\delta t}$$

then the local truncation error is improved.

The implementation of the method is no harder than the Crank–Nicolson scheme.

$$O(\delta S^4, \delta t^2)$$

Three time-level methods

Numerical schemes are not restricted to the use of just two time levels.

We can construct many algorithms using three or more time levels.

Again, we would do this if it gave us a better local truncation error or had better convergence properties.

$$(\delta t, \delta S^2) \rightarrow (\delta t^2, \delta S^4)$$

$$F(V^k, V^{k+1})$$

$$F(V^{k-1}, V^k, V^{k+1})$$

Trunc^n error analysis.

G. D. Smith

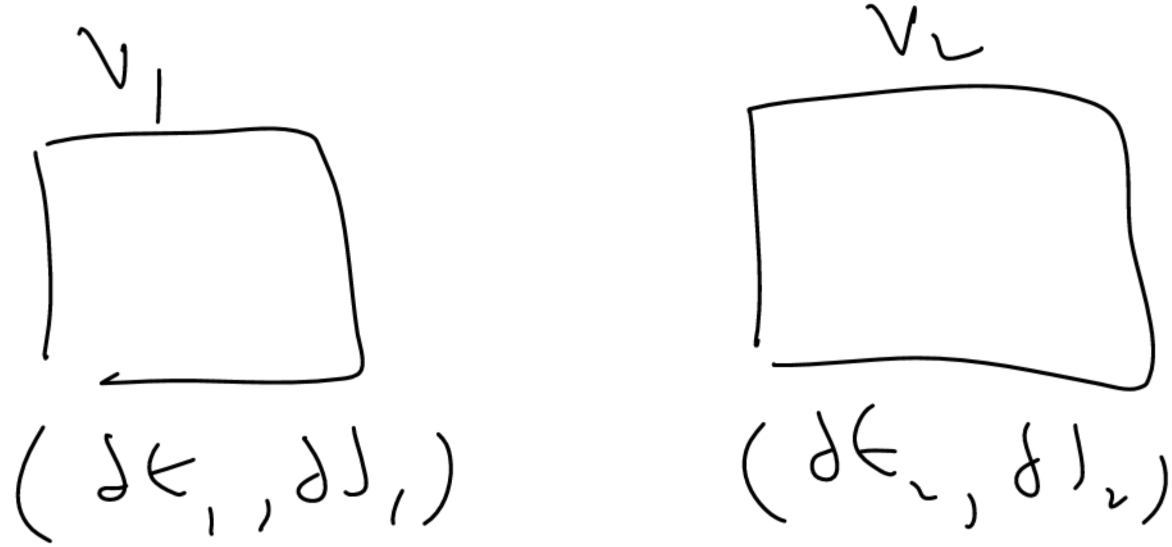
Richardson extrapolation

In the explicit method the error is $O(\delta t, \delta S^2)$.

If we assume that the approach to the correct solution as the time step and asset step tend to zero is in a sense 'regular' then we could postulate that

$$\text{approximate solution} \;=\; \text{exact solution} + \epsilon_1\,\delta t + \epsilon_2\,\delta S^2 + \epsilon_3\,\delta t^2 + \cdots$$

for some coefficients ϵ_i.



Suppose that we have *two* approximate solution V_1 and V_2 using different grid sizes, we can write

$$
\begin{aligned}
V_1 = & \text{ exact solution } + \epsilon_1\, \delta t_1 + \epsilon_2\, \delta S_1^2 + \epsilon_3\, \delta t_1^2 + \cdots \\
= & \text{ exact solution } + \delta S_1^2 \left(\epsilon_1 \frac{\delta t_1}{\delta S_1^2} + \epsilon_2 \right) + \cdots
\end{aligned}
$$

and

$$
\begin{aligned}
V_2 = & \text{ exact solution } + \epsilon_1\, \delta t_2 + \epsilon_2\, \delta S_2^2 + \epsilon_3\, \delta t_2^2 + \cdots \\
= & \text{ exact solution } + \delta S_2^2 \left(\epsilon_1 \frac{\delta t_2}{\delta S_2^2} + \epsilon_2 \right) + \cdots .
\end{aligned}
$$

If we choose

$$\frac{\delta t_1}{\delta S_1^2} = \frac{\delta t_2}{\delta S_2^2},$$

i.e. ν_1 constant, where the subscripts denote the step sizes used in finding the solutions V_1 and V_2, then we can find a *better* solution than both V_1 and V_2 by eliminating the leading-order error terms in the above two equations.

This better approximation is given by

$$\boxed{\frac{\delta S_2^2\, V_1 - \delta S_1^2\, V_2}{\delta S_2^2 - \delta S_1^2}}.$$

The accuracy of the method is now $O(\delta t^2, \delta S^3)$.

C.F.D.

Example: Call option with a strike of 100, expiry of one year, underlying asset at 100 with a volatility of 20% and an interest rate of 10%. The exact Black–Scholes value is 13.269.

Solve using an explicit finite-difference scheme.

V_1

V_2

Method	BS	Numerical	Error	Time taken
20 asset steps	13.269	13.067	0.202	8,000
30 asset steps	13.269	13.183	0.086	27,000
Richardson	13.269	13.275	0.006	**35,000**
100 asset steps	13.269	13.276	0.007	**1,000,000**

$O(\underline{\underline{dt}}, \delta S^2)$

10^{-8}

10^{-2}

10^{-4}

Free boundary problems and American options

The value of American options must always be greater than the payoff, otherwise there will be an arbitrage opportunity

$$V(S,t) \geq \text{Payoff}(S).$$

American options are examples of 'free boundary problems' we must solve a partial differential equation with an unknown boundary, the position of which is determined by having one more boundary condition than if the boundary were prescribed.

In the American option problem we know that both the option value and its delta are continuous with the payoff function, the so called **smooth pasting condition**.

Early exercise and the explicit method

Suppose that we have found V_i^k for all i at the time step k, time step to find the option value at $k+1$ using the finite-difference scheme

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

Don't worry about whether or not you have violated the American option constraint until you have found the option values V_i^{k+1} for all i.

Now let's check whether the new option values are greater or less than the payoff.

If they are less than the payoff then we have arbitrage.

We can't allow that to happen so at every value of i for which the option value has allowed arbitrage, replace that value by the payoff at that asset value.

Early exercise and Crank–Nicolson

The Crank–Nicolson method is implicit, and every value of the option at the $k+1$ time step is linked to every other value at that time step.

It is therefore not good enough to just replace the option value with the payoff after the values have all been calculated, the replacement must be done at the same time as the values are found:

< / >

$$v_i^{n+1} = \max\left(v_i^n + \frac{\omega}{M_{ii}}\left(q_i - \sum_{j=1}^{i-1} M_{ij} v_j^{n+1} - \sum_{j=i}^{N} M_{ij} v_j^n \right), \text{Payoff} \right).$$

The payoff is evaluated at i and $k+1$ in the obvious manner. This method is called **projected SOR**.

computed American price

Two-factor models $V(t, S, r)$

We are going to refer to the general two-factor equation

equity option with stoch. i.r.

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

$$+d(S,r,t)\frac{\partial^2 V}{\partial r^2} + e(S,r,t)\frac{\partial^2 V}{\partial S \partial r} + f(S,r,t)\frac{\partial V}{\partial r} = 0. \qquad (1)$$

It will be quite helpful if we think of solving the two-factor convertible bond problem.

For the general two-factor problem (1) to be parabolic we need

classifying a parabolic PDE.

$$\boxed{e(S,r,t)^2 < 4a(S,r,t)\,d(S,r,t).}$$

method of characteristics

As in the one-factor world, the variables must be discretized.

That is, we solve on a three-dimensional grid with

$V(S, r, t)$

new

$$S = i\,\delta S, \quad \boxed{r = j\,\delta r,} \quad \text{and} \quad t = T - k\,\delta t.$$

Expiry is $t = T$ or $k = 0$.

new index.

The indices range from zero to I and J for i and j respectively.

$$V(S, r, t) \longrightarrow V_{i\,j}^{k}$$

increase the subscripts.

We will assume that the interest rate model is only specified on $r \geq 0$.

This may not be the case, some simple interest rate models such as Vasicek, are defined over negative r as well.

The contract value is written as

$$V(S, r, t) = V_{ij}^{k}.$$

Whatever the problem to be solved, we must impose certain conditions on the solution.

First of all, we must specify the final condition.

This is the payoff function, telling us the value of the contract at the expiration of the contract.

Suppose that we are pricing a long-dated warrant with a call payoff.

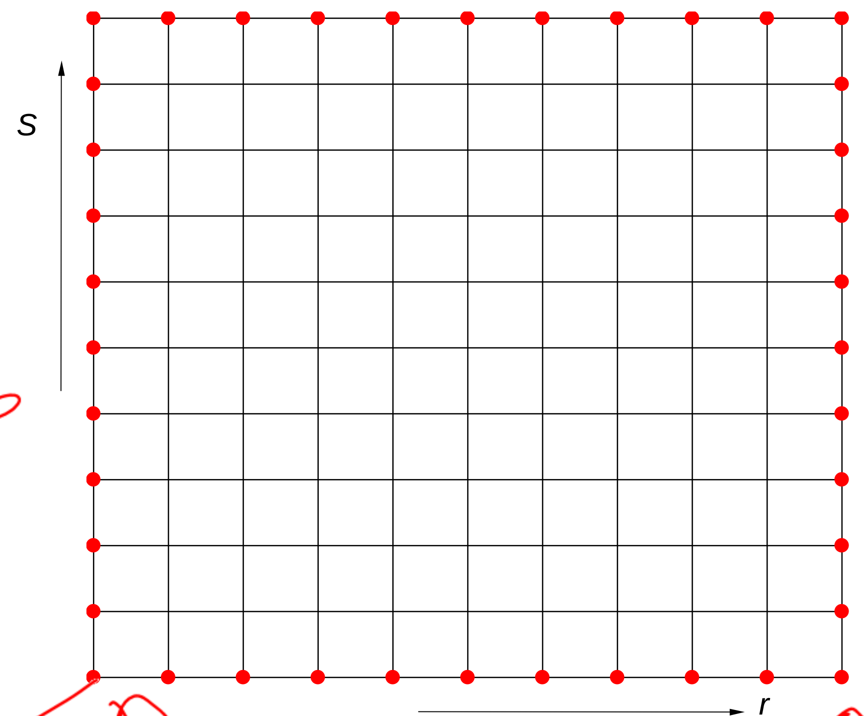
The final condition for this problem is then

$$V(S, r, T) = \underbrace{V_{ij}^{0} = \max(S - E, 0)}_{\text{Payoff at } t = T}.$$

$k = 0$

Boundary conditions must be imposed at all the grid points marked with a dot. The boundary conditions will depend on the contract.

Remember that there is also a time axis coming out of the page, and not drawn in this figure.



$S = 0$

$S = \infty$

$0 < r < \infty$

t

The explicit method

The one-factor explicit method can be extended to two-factors with very little effort.

In fact, the ease of programming make it a very good method for those new to the subject.

We will use symmetric central differences for all derivatives in (1).

This is the best way to approximate the second derivatives but may not be the best for the first derivatives.

We have seen how to use central differences for all of the terms with the exception of the second derivative with respect to both S and r,

How to discretise?

$$\boxed{\frac{\partial^2 V}{\partial S \partial r}.} = \frac{\partial^2 V}{\partial r \partial S}$$

We can approximate this by

$$\frac{\partial}{\partial S} V = V^k_{i+1} - V^k_{i-1} / 2\,\delta S$$

$$\frac{\partial V}{\partial r}$$

$$\frac{\partial \left(\frac{\partial V}{\partial r} \right)}{\partial S} \approx \frac{\frac{\partial V}{\partial r}(S + \delta S, r, t) - \frac{\partial V}{\partial r}(S - \delta S, r, t)}{2\,\delta S}.$$

But

$$= \frac{V^k_{j+1} - V^k_{j-1}}{2\,\delta r}$$

$$\frac{\partial V}{\partial r}(S + \delta S, r, t) \approx \frac{V^k_{i+1,j+1} - V^k_{i+1,j-1}}{2\,\delta r}.$$

This suggests that a suitable discretization might be

V_{i+1} V_{i-1}

$$\frac{\frac{V^k_{i+1,j+1}-V^k_{i+1,j-1}}{2\,\delta r}-\frac{V^k_{i-1,j+1}-V^k_{i-1,j-1}}{2\,\delta r}}{2\,\delta S}$$

$$\frac{\partial^2 V}{\partial S \partial r} = \frac{V^k_{i+1,j+1}-V^k_{i+1,j-1}-V^k_{i-1,j+1}+V^k_{i-1,j-1}}{4\,\delta S\,\delta r}.$$

This is particularly good since, not only is the error of the same error as in the other derivative approximations but also it preserves the property that

$$\frac{\partial^2 V}{\partial S \partial r} = \frac{\partial^2 V}{\partial r \partial S}.$$

The resulting explicit difference scheme is

$$\frac{V_{ij}^{k}-\boxed{V_{ij}^{k+1}}}{\delta t}+a_{ij}^{k}\left(\frac{V_{i+1,j}^{k}-2V_{ij}^{k}+V_{i-1,j}^{k}}{\delta S^{2}}\right)$$

$\partial V / \partial t$ $\quad$ $\partial^2 V/\partial S^2$

$$+b_{ij}^{k}\left(\frac{V_{i+1,j}^{k}-V_{i-1,j}^{k}}{2\,\delta S}\right)+c_{ij}^{k}V_{ij}^{k}$$

$-rV$ $\quad$ $\partial V/\partial S$

$$+d_{ij}^{k}\left(\frac{V_{i,j+1}^{k}-2V_{ij}^{k}+V_{i,j-1}^{k}}{\delta r^{2}}\right)$$

$\dfrac{\partial^2 V}{\partial r^2}$

rearrange as

$V_{ij}^{k+1} = \text{-------}$

$$e_{ij}^{k}\left(\frac{V_{i+1,j+1}^{k}-V_{i+1,j-1}^{k}-V_{i-1,j+1}^{k}+V_{i-1,j-1}^{k}}{4\,\delta S\,\delta r}\right)$$

$\dfrac{\partial^2 V}{\partial S \partial r}$

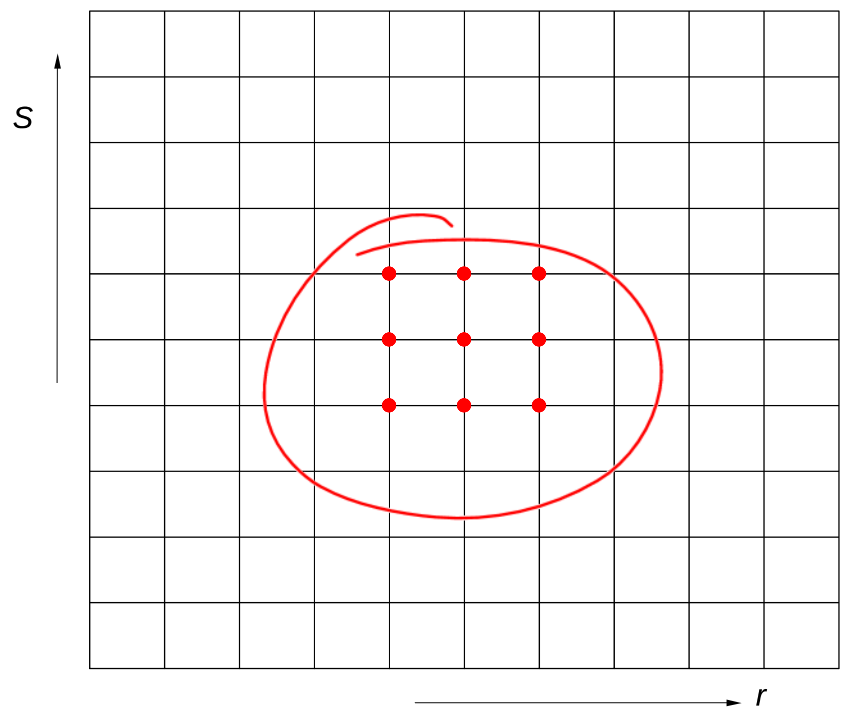
$$+f_{ij}^{k}\left(\frac{V_{i,j+1}^{k}-V_{i,j-1}^{k}}{2\,\delta r}\right)=O(\delta t,\delta S^{2},\delta r^{2}).$$

$\dfrac{\partial V}{\partial r}$

We could rewrite this in the form

$$V_{ij}^{k+1} = \ldots ,$$

where the right-hand side is a linear function of the nine option values shown schematically below.



The coefficients of these nine values at time step k are related to a, b etc.

It would not be very helpful to write the difference equation in this form, since the actual implementation is usually more transparent than this.

Note that in general all nine points (i, j), $(i \pm 1, j)$, $(i, j \pm 1)$, $(i \pm 1, j \pm 1)$ are used in the scheme.

If there is no cross derivative term then only the five points (i, j), $(i \pm 1, j \pm 1)$ are used.

This simplifies some of the methods.

Derivatives Market Practice in Historical Perspective

Prof. Espen Gaarder Haug
CQF, Fitch Learning, London

25 March 2020

Hello

Option traders now and then

Myth 1:

People Did not Properly price options before Black-Scholes-Merton.

Myth 2:

Option traders today use the Black-Scholes-Merton formula.

Put-Call Parity

Extremely Robust Arbitrage Principle

F-X=call-put

According to some modern text books and famous Professors put-call parity invented 1969!

Joseph de la Vega 1688

"We say of those who buy means of a forward call contract and sell at fixed term or of those who sell by means of a put contract and buy at a fixed term that they shift the course of their speculation."

Rubienstein

Reprint in the book "Extraordinary Popular Delusions and the Madness of Crowds & Confusion de Confusiones" 1996 John Wiley & Sons, edited by Martin Fridson

Put-Call Parity early 1900

A) As pure arbitrage constrain.

B) To convert calls to puts, and puts and calls to straddles for hedging purpose. Options with options.

Option prices affected by supply and demand of options.

Higgins 1902/Nelson 1904

"It may be worthy of remark that 'calls' are more often dealt than 'puts' the reason probably being that the majority of 'punters' in stocks and shares are more inclined to look at the bright side of things, and therefore more often 'see' a rise than a fall in prices.

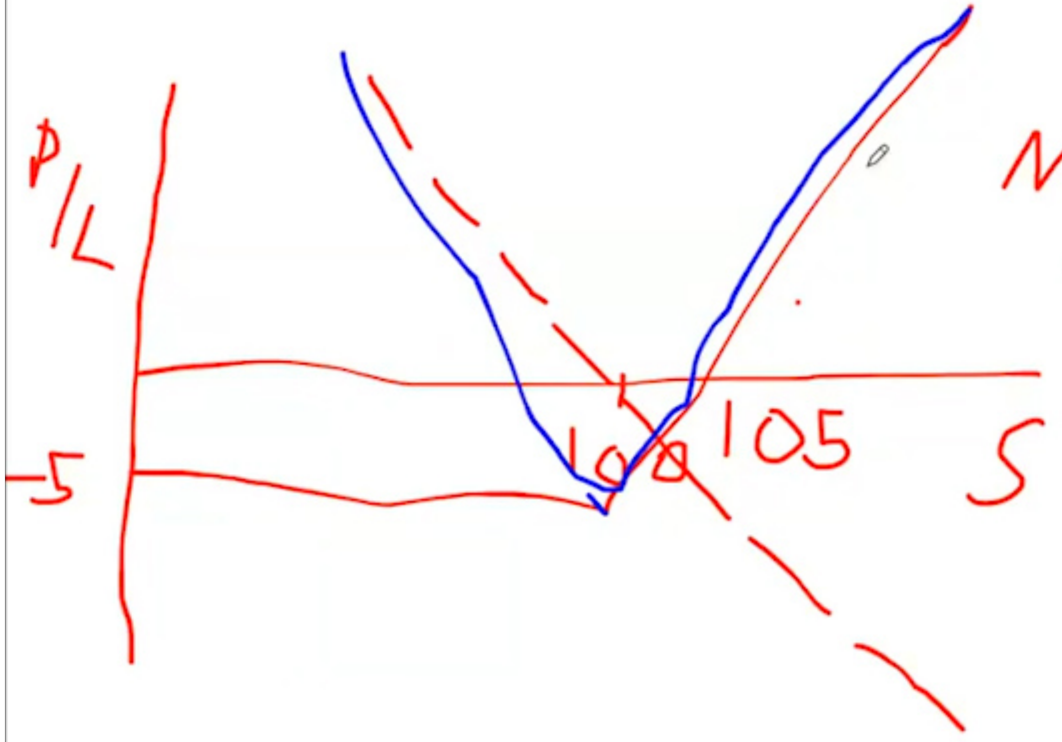
This special inclination to buy 'calls' and to leave the 'puts' severely alone dose not, however, tend to make 'calls' dear and 'puts' cheap, for it can be shown that the adroit dealer in options can convert a 'put' into a 'call,' a 'call' into a 'put', a 'call o' more' into a 'put- and-call,' in fact any option into another, by dealing against it in the stock. We may therefore assume, with tolerable accuracy, that the 'call' of a stock at any moment costs the same as the 'put' of that stock, and half as much as the Put-and-Call. "

Higgins 1902:

- 1. That a Call of a certain amount of stock can be converted into a Put-and- Call of half as much by selling one-half of the original amount.

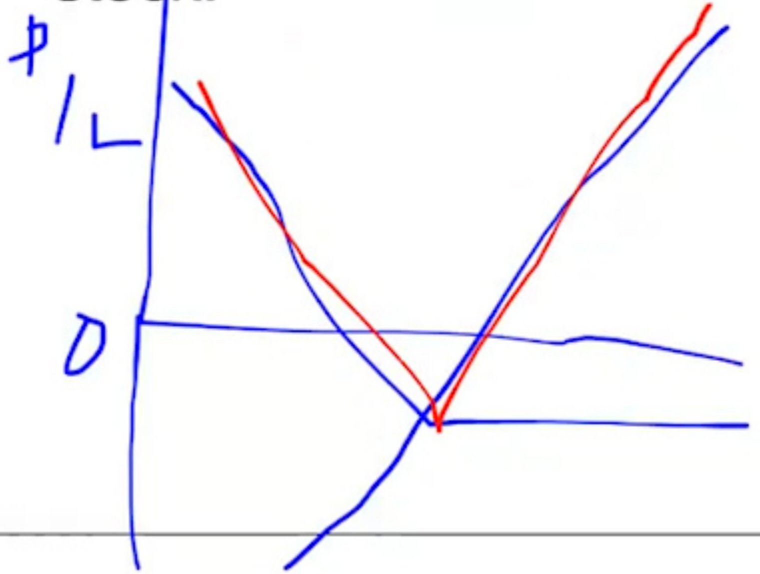
+50 c
+80 p

N=100
−50S



Higgins 1902:

- 1. That a Call of a certain amount of stock can be converted into a Put-and- Call of half as much by selling one-half of the original amount.
- 2. That a Put of a certain amount of stock can be converted into a Put-and- Call of half as much by buying one-half of the original amount.
- 3. That a Call can be turned into a Put by selling all the stock.



Higgins 1902/Nelson 1904

"It may be worthy of remark that 'calls' are more often dealt than 'puts' the reason probably being that the majority of 'punters' in stocks and shares are more inclined to look at the bright side of things, and therefore more often 'see' a rise than a fall in prices.

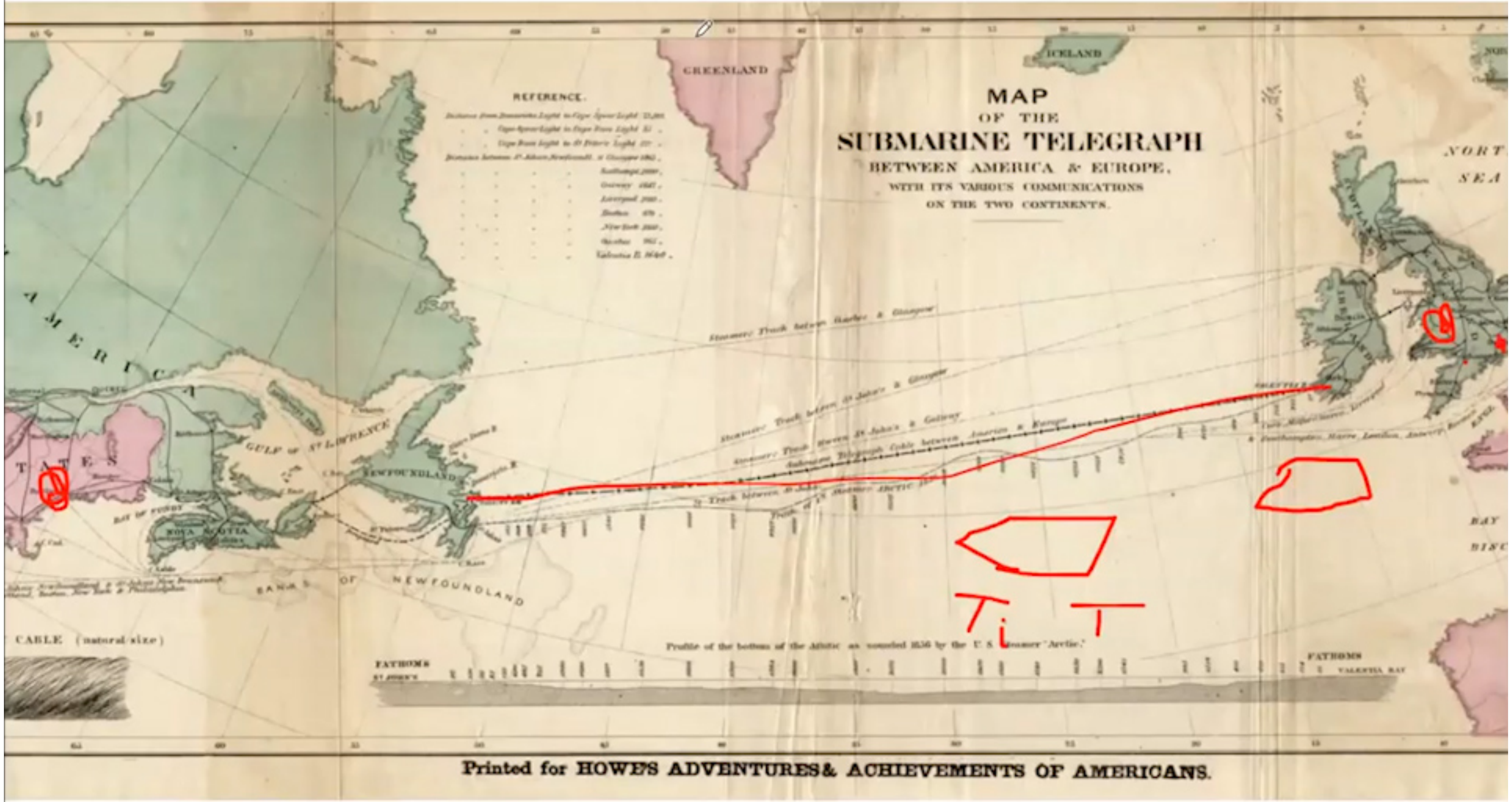
This special inclination to buy 'calls' and to leave the 'puts' severely alone dose not, however, tend to make 'calls' dear and 'puts' cheap, for it can be shown that the adroit dealer in options can convert a 'put' into a 'call,' a 'call' into a 'put', a 'call o' more' into a 'put- and-call,' in fact any option into another, by dealing against it in the stock. We may therefore assume, with tolerable accuracy, that the 'call' of a stock at any moment costs the same as the 'put' of that stock, and half as much as the Put-and-Call. "

Reinach 1961

"Although I have no figures to substantiate my claim, I estimate that over 60 per cent of all Calls are made possible by the existence of Converters."

Reinach understands put-call parity do not hold for American options, mention how converters try to arbitrage on it.

Five attempts 1857, two in 1858, 1865, 1866 lasting connection



International Arbitrage Trading

Telegraph

- 1858 0.1 Words per minute
- 1866 8 words per minute
- 1900 >120 words per minute

Horrible with such high frequency trading!
Is it not?

Options Arbitrage Between London and New York (Nelson 1904)

Up to 500 messages per hour and typically 2,000 to 3,000 messages per day where sent between the London and the New York market through the cable companies. Each message flashed over the wire system in less than a minute.

Empirical Research pre-1973 markets

Options priced much as today

There was a volatility skew

Empirical regularities in implied vol etc similar.

Mixon, S. (2008): "Option Markets and Implied Volatility: Past Versus Present

Kairys, J. P. and N. Valerio (1997) The Market for Equity Options in the 1870s, Journal of Finance, Vol LII, NO. 4., Pp 1707–1723.

Delta Hedging

Delta Hedging invented in 1960's to 1970's ?

Dynamic Delta Hedging Invented ?

Continuous Time Dynamic Delta Hedging invented in 1970's Black-Scholes-Merton

Nelson 1904

$N = 100$
$+50$

"Sellers of options in London as a result of long experience, if they sell a Call, straightway buy half the stock against which the Call is sold; or if a Put is sold; they sell half the stock immediately."

Standard options at that time in London always issued at-the-money forward and European style

Nelson 1904

"The regular London option is always either a Put or a Call, or both, at the market price of the stock at the time the bargain is made, to which is immediately added the cost of carrying or borrowing the stock until the maturity of the option."

Put-Call Parity 1902/1904

Static delta hedging 1902/1904 for at-the-money options. Nelson/Higgins

Static delta hedging any option 1967, Thorp

Idea of dynamic delta hedging 1904? 1969 Thorp

Put-Call Parity 1902/1904

Static delta hedging 1902/1904 for at-the-money options. Nelson/Higgins

Static delta hedging any option 1967, Thorp

Beat + dealer

market

Put-Call Parity 1902/1904

Static delta hedging 1902/1904 for at-the-money options. Nelson/Higgins

Static delta hedging any option 1967, Thorp

Idea of dynamic delta hedging 1904? 1969 Thorp

Continuous-time dynamic delta hedging 1973 Merton

Arbitrage early 1900

WWII

1923

1. Postage
2. Interest expenses
3. Insurance
4. Transportation
5. Transportation arbitrage
6. Credit risk

Option Pricing Formulas

Before Black-Scholes-Merton

- Bachelier (1900) Normal distribution. Says little about how to hedge out risk in options.

1000x

Brown

Option Pricing Formulas

Before Black-Scholes-Merton

- Bachelier (1900) Normal distribution. Says little about how to hedge out risk in options.

1000x

1905

Option Pricing Formulas Before Black-Scholes-Merton

Bachelier (1900) Normal distribution. Says little about how to hedge out risk in options.

Bronzin (1908) based on put-call parity, several types of distributions.

Sprenkle (1960), Book of Cootner

Ayres (1963), Book of Cootner

Boness (1964),

Samuelson (1965)

Thorp (1969)

Professor Bronzin 1908, Option Pricing:

$$F(x) = \int_x^{\infty} f(x)\, dx \text{ resp. } F_1(x) = \int_x^{\omega_1} f_1(x)\, dx \tag{5}$$

die Gesamtprobabilitäten dar, daß die Schwankungen über resp. unter B am Liquidationstermin die Größe x übersteigen; wir werden bald erfahren, welche bedeutende Rolle gerade diese Funktionen in den späteren Betrachtungen spielen werden.

Tragen wir auf einer horizontalen Geraden rechts von einem Punkte 0 die Marktschwankungen über B, links davon hingegen die Schwankungen unter B auf und errichten wir in den jeweiligen Endpunkten Senkrechte, welche die entsprechenden Funktionswerte $f(x)$ bezw. $f_1(x)$ darstellen sollen, so entstehen zwei kontinuierliche Kurven C und C_1, die wir füglich Schwankungswahrscheinlichkeitskurven nennen werden (siehe Fig. 23); die zwischen irgend zwei Ordinaten

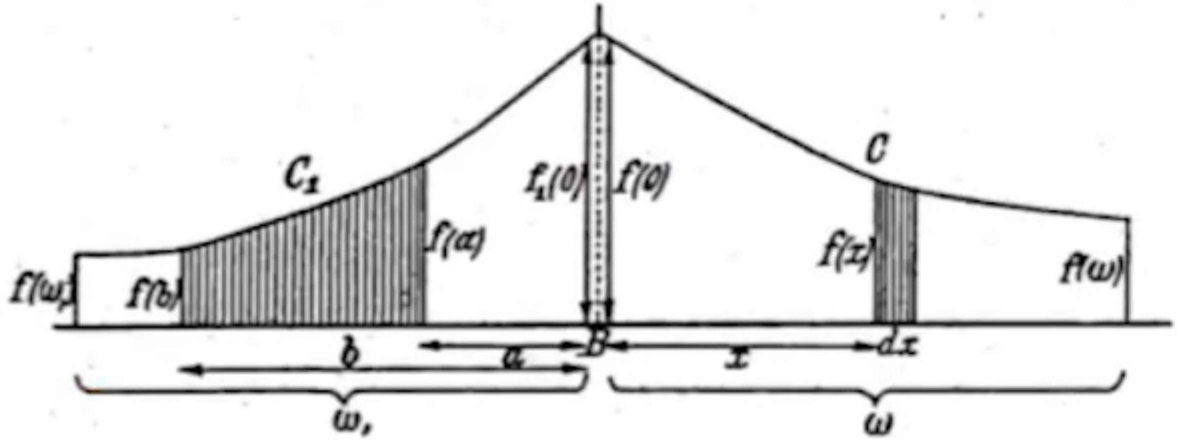
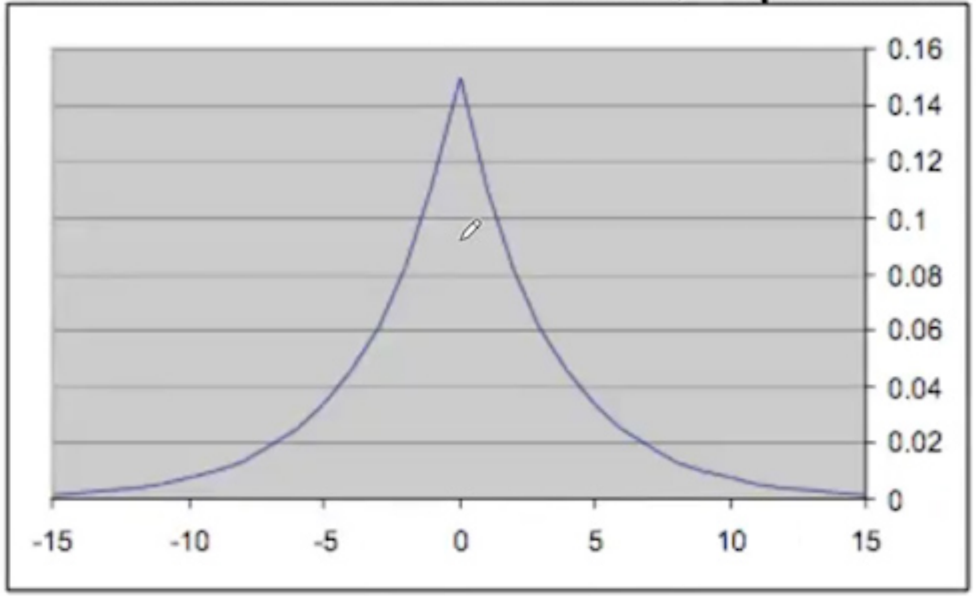
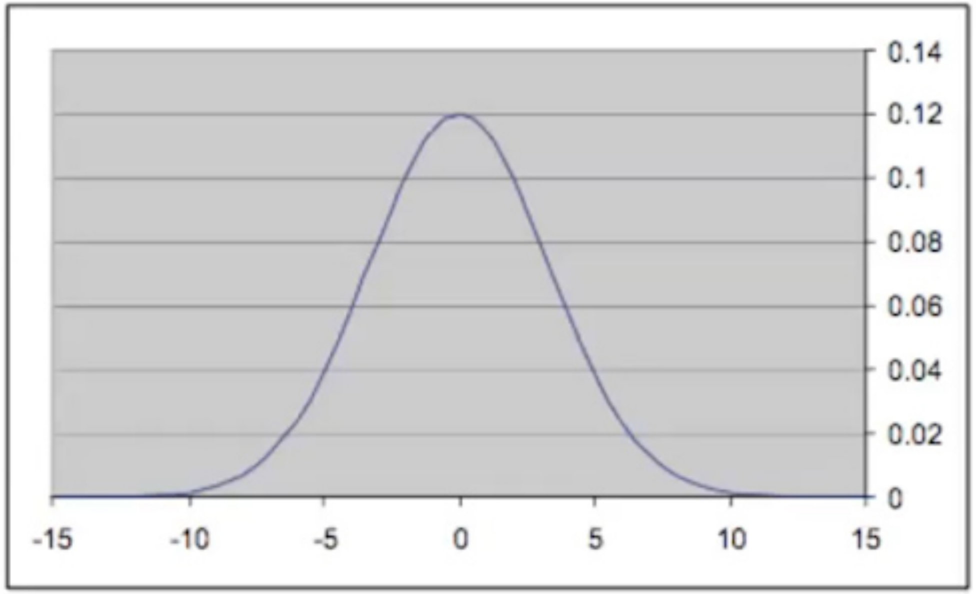


Fig. 23.

Professor Bronzin 1908, Option Pricing:



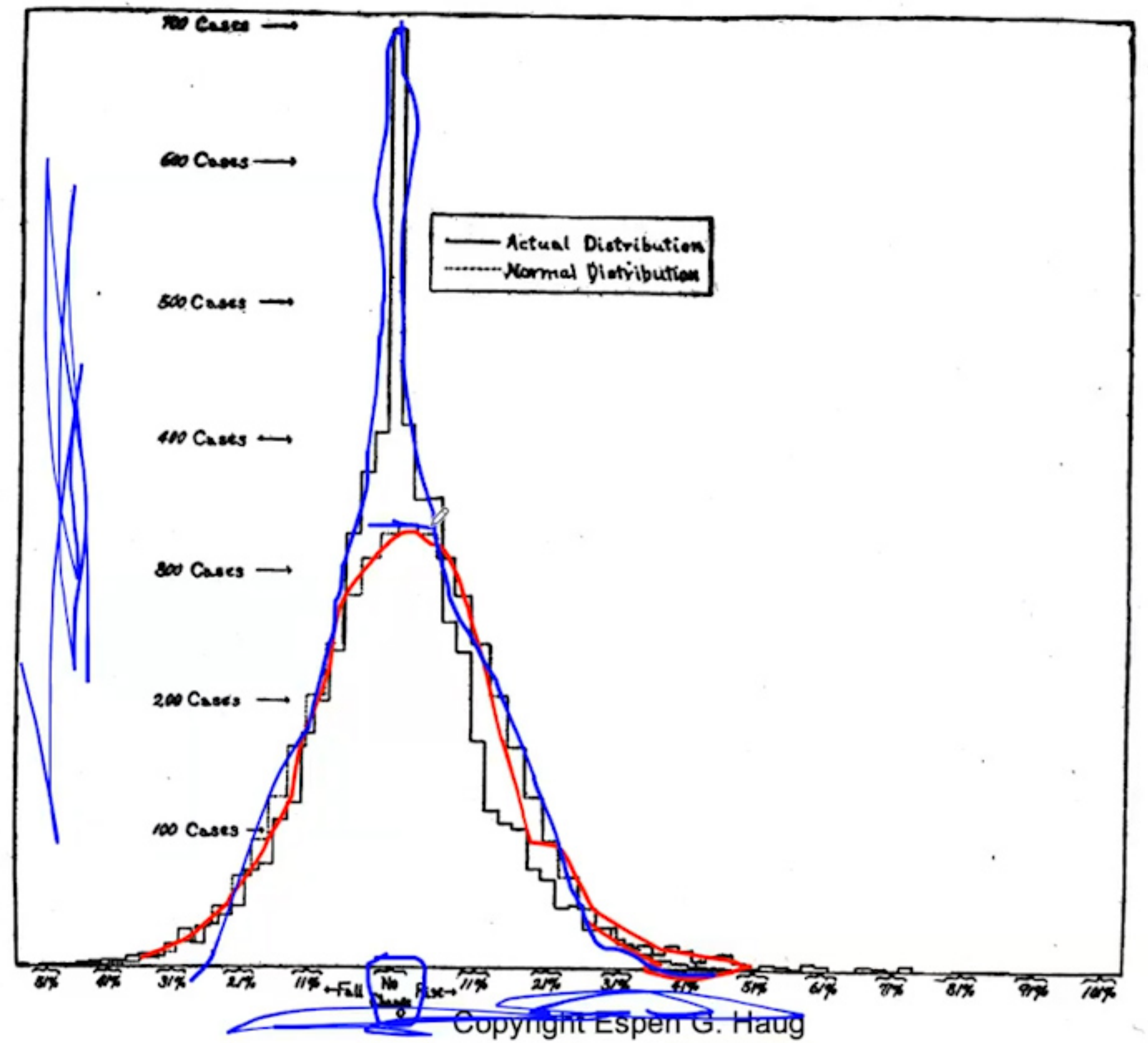


Fat-Tails in Price Data

Wesley Clair Mitchell
1874-1948

"The Making and Using of Index Numbers" published in 1915

CHART 2.—DISTRIBUTION OF 5,578 PRICE VARIATIONS (PERCENTAGES OF RISE OR FALL FROM PRICES OF PRECEDING YEAR).



Copyright Espen G. Haug

Mills rejects the Gaussian hypothesis.

“A distribution may depart widely from the Gaussian type because the influence of one or two extreme price changes.”

Mills, F. C. (1927): The Behaviour of Prices. New York: National Bureau of Economic Research, Albany: The Messenger Press.

Osborne 1958

“Brownian Motion in Stock Market”

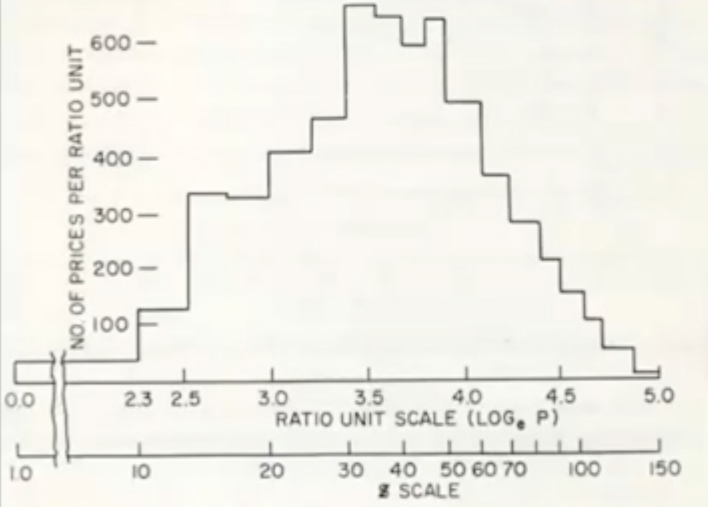


Fig. 4. Distribution function of $\log_e P$ for common stocks (NYSE, July 31, 1956).

This nearly normal distribution in the changes of logarithm of prices suggests that it may be a consequence of many independent random variables contributing to the changes in values (as defined by the Weber-Fechner law). The normal distribution arises in many stochastic proc-

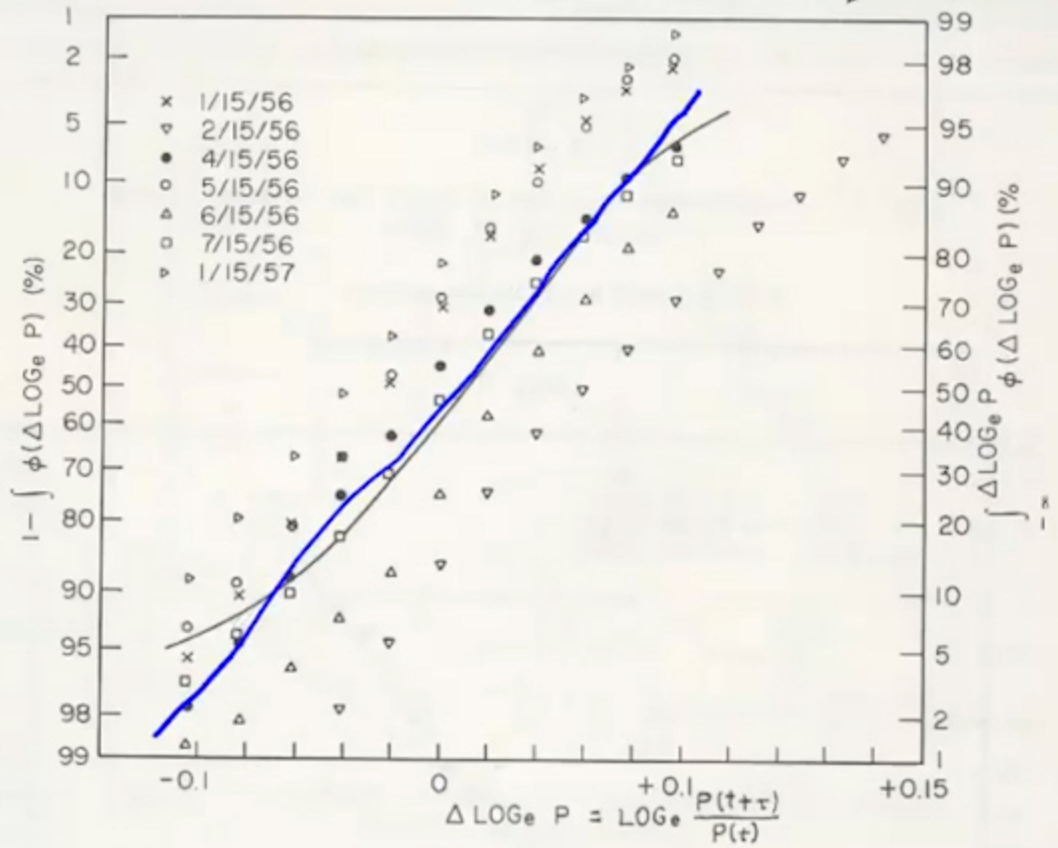


Fig. 7. Cumulated distributions of $\Delta \log_e P = \log_e[P(t+\tau)/P(t)]$ for $\tau = 1$ month (NYSE common stocks). These, and also Fig. 8, may be regarded as distributions of $S(\tau)$ for fixed $M^*(\tau)$. The solid line is the distribution of $Z(\tau) \approx M(\tau)$, transcribed from Fig. 12 for comparison.

esses involving large numbers of independent variables, and certainly the market place should fulfill this condition, at least.

Osborne (1959) detects fat-tails in price data, but basically ignores them and seems to be a strong believer in normal distributed returns.

Mandelbrot 1962

“The Variation of Certain Speculative Prices”

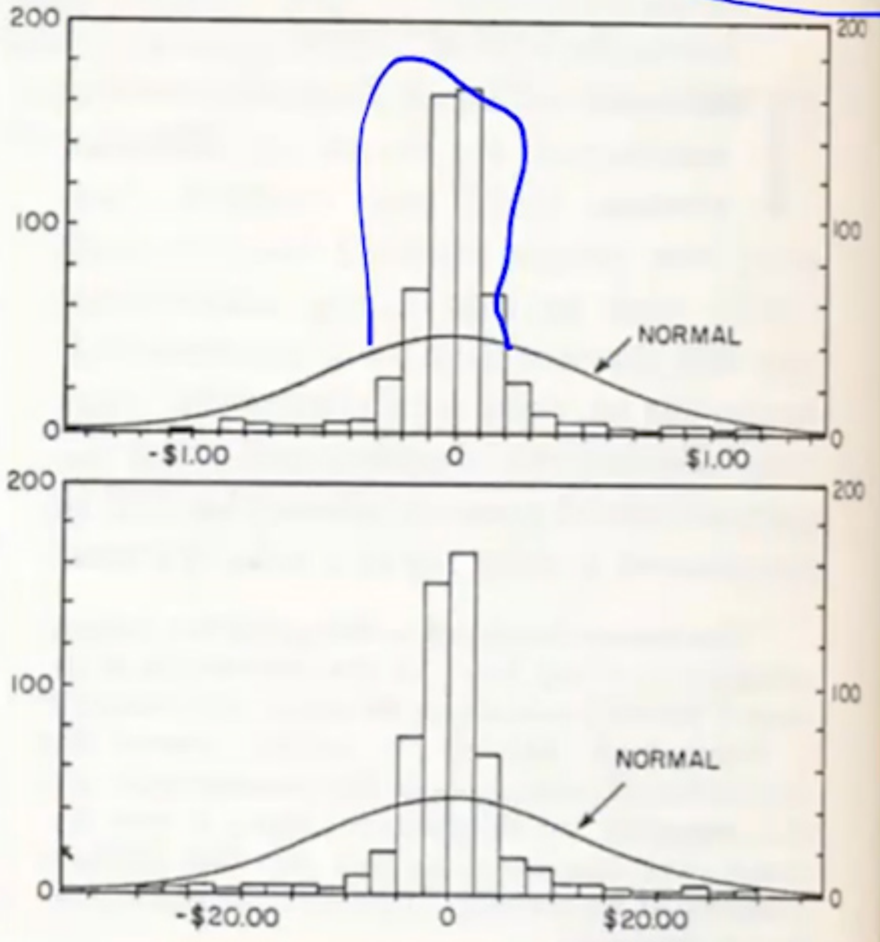


FIG. 1.—Two histograms illustrating departure from normality of the fifth and tenth difference of monthly wool prices, 1890–1937. In each case, the continuous bell-shaped curve represents the Gaussian “interpolate” based upon the sample variance. Source: Gerhard Tintner, *The Variate-Difference Method* (Bloomington, Ind., 1940).

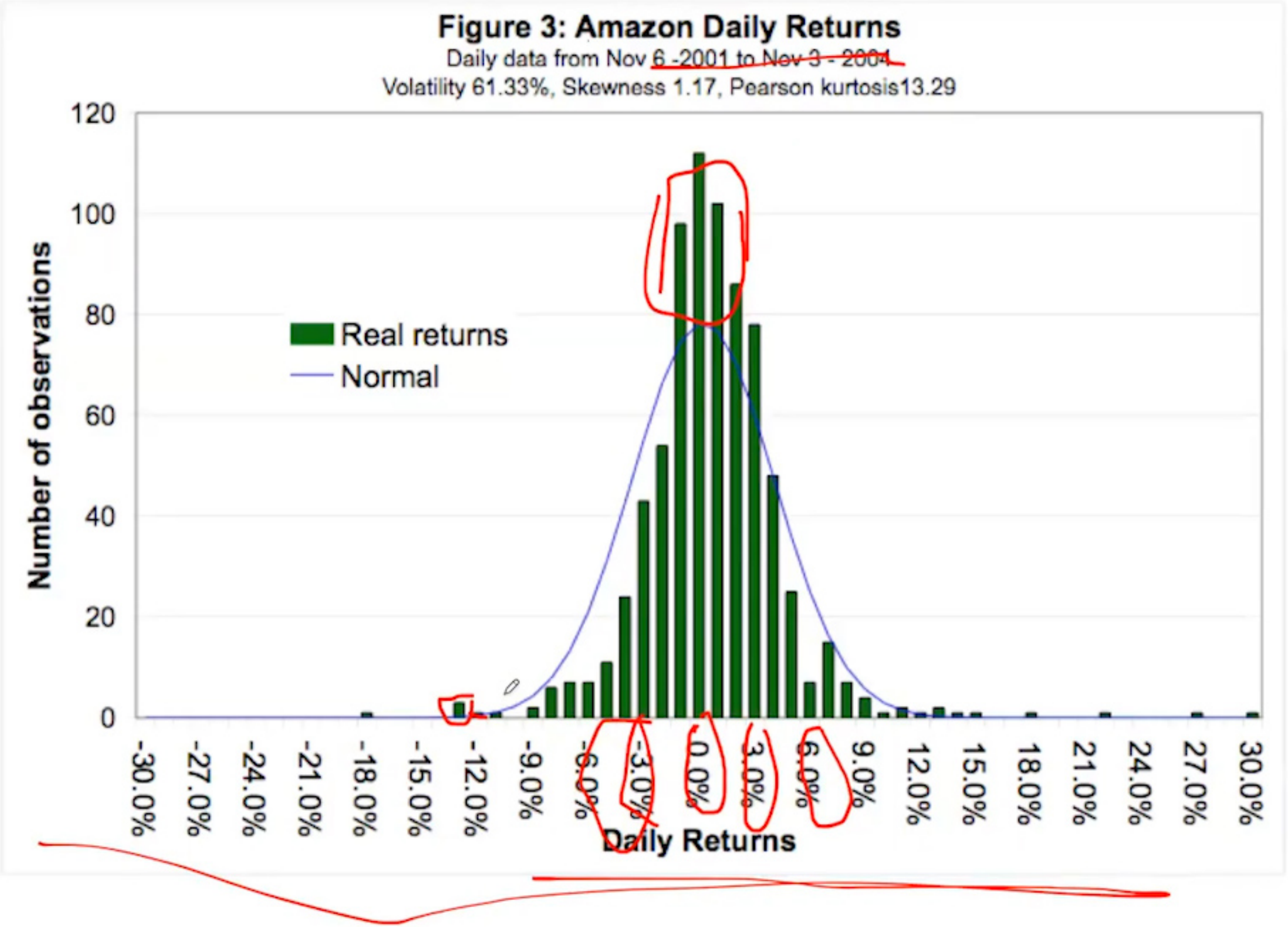


Figure 3: Amazon Daily Returns
Daily data from Nov ~~6 -2001 to Nov 3 - 2004~~
Volatility 61.33%, Skewness 1.17, Pearson kurtosis13.29
Number of observations
120
100
80
60
40
20
0
Real returns
Normal
-30.0%
-27.0%
-24.0%
-21.0%
-18.0%
-15.0%
-12.0%
-9.0%
-6.0%
-3.0%
0.0%
3.0%
6.0%
9.0%
12.0%
15.0%
18.0%
21.0%
24.0%
27.0%
30.0%
Daily Returns

If Gaussian

We can measure all risk by variance/standard deviation σ.

Easy to make models.

Consistent models

Cost: we are loosing out on important information if non-Gaussian

Some of the BIG Ideas in Finance

CAPM: Based on Gaussian!

Also

Sharpe Ratio: Based on Gaussian!

Black-Scholes-Merton: Based on Gaussian!

Before Black-Scholes-Merton

Option valuation by discounting expected value

$$dS = \mu S_t dt + \sigma S_t dz$$

Bachelier/Sprenkle/Boness/Thorp (1964/1969)

formula: Exact Boness formula

$$c = SN(d_1) - Xe^{-rT}N(d_2)$$

$$d_1 = \frac{\ln(S/X) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \quad d_2 = d_1 - \sigma\sqrt{T}$$

Before Black-Scholes-Merton

Option valuation by discounting expected value

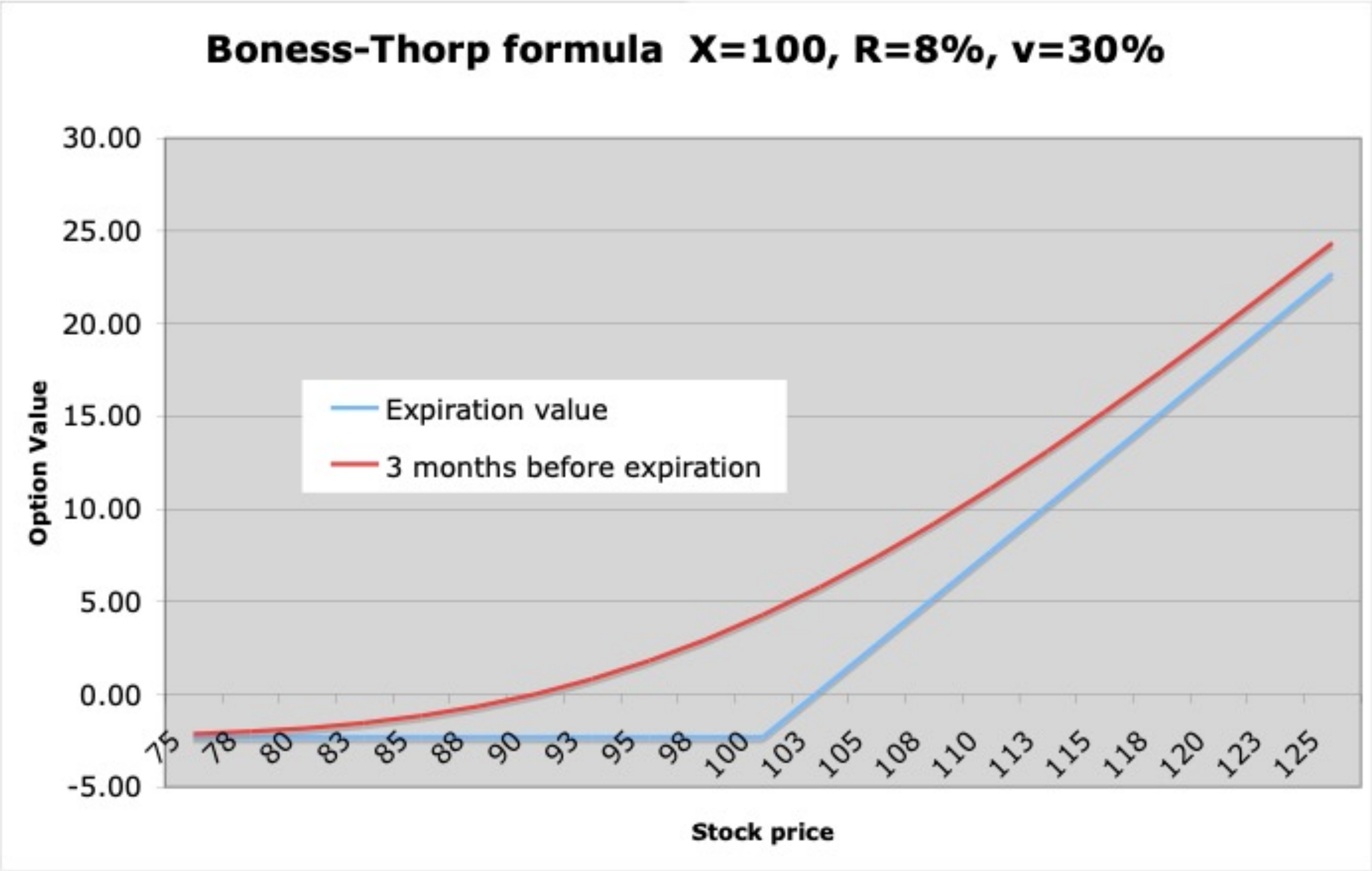
$$dS = \mu S_t dt + \sigma S_t dz$$

Bachelier/Sprenkle/Boness/Thorp (1964/1969)

formula: Exact Boness formula

$$c = SN(d_1) - Xe^{-\mu T}N(d_2)$$

$$d_1 = \frac{\ln(S/X) + (\mu + \sigma^2/2)T}{\sigma\sqrt{T}} \quad d_2 = d_1 - \sigma\sqrt{T}$$


Boness-Thorp formula X=100, R=8%, v=30%
Option Value
30.00
25.00
20.00
15.00
10.00
5.00
0.00
-5.00
Expiration value
3 months before expiration
75
78
80
83
85
88
90
93
95
98
100
103
105
108
110
113
115
118
120
123
125
Stock price

MODERN HISTORY DELTA HEDGING

Ed Thorp (2002) 1966: "We understood static hedging, dynamic hedging, and delta hedging,in particular market-neutral delta hedging."

Thorp and Kassouf (1967) "Beat the Market"

Thorp (1969) "Optimal Gambling Systems for Favorable Games"

Rubinstein (2006) "Thorp came close since he early on understood the idea of dynamically delta hedging an option with a position its underlying asset…."

Thorp 1967

“If, when the common changed price, the warrant moved along this line, then a 1 point increase in the common would result in a 1/3 point increase in the warrant. If we are short 3 warrants to one common long, then the gain on the common is completely offset by the loss on the warrant.”

Thorp 1969

“We have assumed so far that a hedge position is held unchanged until expiration, then closed out. This static or ‘desert island’ strategy is not optimal. In practice intermediate decisions in the spirit of dynamic programming lead to considerably superior dynamic strategies. The methods, technical details, and probabilistic summary are more complex so we defer the details for possibly subsequent publication.”

Arnold Bernhard & Co., 1970

EVALUATION OF WARRANT

	...bol	% Overvalued (+) Undervalued (−) Footnote (See Next Page)		Wt Leverage Projections (for these changes in the price of the underlying security) +50%	+25%	−25%	−50%	Hedge Rank	Hedge Ratio	Effective Per Share Exercise Price	Per Share Cost of Option	Total $ Premium	% Premiu... Tangible Value			
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	3...	
1	Y.WS		+40%	+45%	+23%	−40%	−70%	F	90	3.75	5.13	2.00	3.13	29%	Y	
2			+21	+65	+35	−40	−70	D	40	37.52	4.00	4.00	.00	28	ADP	
3		1	−1	+75	+35	−30	−60	C	50	12.00	2.25	2.25	.00	35	AMO	
4	T.WS	2	+25	+140	+60	−55	−85	C+	40	52.00	8.50	8.50	.00	19	T	
5			+55	+45	+25	−45	−75	F	55	28.00	6.50	6.50	.00	43	APL	
6			−16	+125	+70	−30	−65	A	85	47.50	42.00	−5.63	47.63	−6	ARA	NYS
7	ARCW	3	−10	+115	+50	−35	−65	B	35	110.00	11.25	11.25	.00	18	ARC	NYS
8	AZ>W	4	−4	+60	+30	−25	−55	C−	65	6.25	1.38	1.38	.00	55	AZ	NYS
9	ATOW	5	+30	+60	+30	−40	−70	D	35	30.00	2.13	2.13	.00	27	ATO	NYS
10	AV.W		+7	+80	+40	−35	−65	C−	35	33.56	2.88	2.88	.00	24	AV	NYS
	NKW	6	+75	+45	+25	−45	−75	F	30	55.00	1.88	1.88	.00	24	BNK	NYS
		7	+17	+70	+35	−40	−70	D	60	12.50	5.50	5.25	.25	41	BIW	NYS
		8	+995	−100	−100	−100	−100	F	25	11.98	.50	.50	.00	8	BKY	NYS
			−16	+100	+45	−30	−60	B	135	11.43	2.50	7.50	.00	29	BNF	NYS
		9	−18	+105	+50	−30	−60	B	45	7.26	1.38	1.38	.00	28	BWN	NYS
	Y	10	+18	+70	+35	−40	−70	D	45	14.00	2.38	2.38	.00	32	BGT	NYS
	CNY	11	+50	+60	+30	−45	−75	F	30	23.73	1.19	1.25	.00	21	CCN	NY
	CTCW		−13	+105	+50	−35	−60	B	[illegible]	22.38	6.25	6.25	.00	31	CTC	N
19	DLNW	12	−4	+90	+45	−35	−65	C+	60	12.94	5.38	4.19	1.19	30	DLN	
20		13	−7	+90	+45	−35	−65	C+	35	45.00	4.50	4.50	.00	23	DRY	

Black-Scholes reference to Thorp and Kassouf

One of the concepts that we use in developing our model is expressed by Thorp and Kassouf (1967). They obtain an empirical valuation formula for warrants by fitting a curve to actual warrant prices. Then they use this formula to calculate the ratio of shares of stock to options needed to create a hedged position by going long in one security and short in the other. What they fail to pursue is the fact that in equilibrium, the expected return on such a hedged position must be equal to the return on a riskless asset. What we show below is that this equilibrium condition can be used to derive a theoretical valuation formula.

Black-Scholes 1973

"If the hedge is maintained continuously, then the approximations mentioned above become exact, and the return on the hedged position is completely independent of the change in the value of the stock. In fact, the return on the hedged position becomes certain. This was pointed out to us by Robert Merton."

HISTORY

Black-Scholes (1973) and Merton (1973): "First" to publish the brilliant idea of removing all risk by holding the right combination of options and stocks, where the number of stocks are continuously rebalanced, known as dynamic: delta hedging, dynamic replication, dynamic spanning…

$$dS = \mu S_t dt + \sigma S_t dz$$

Ito's Lemma

$$dc = \left[\frac{\partial c}{\partial t} + \frac{\partial c}{\partial S}\mu S + \frac{1}{2}\frac{\partial^2 c}{\partial S^2}\sigma^2 S^2\right] dt + \frac{\partial c}{\partial S}\sigma S dz$$

Idea: risk free portfolio $$\boxed{-c} + \boxed{\frac{\partial c}{\partial S}} S$$

Change in value : $$dV = \boxed{-dc} + \frac{\partial c}{\partial S} \boxed{dS}$$

$$dV = -\frac{\partial c}{\partial t} dt - \frac{1}{2} \frac{\partial^2 c}{\partial S^2} \sigma^2 S^2 dt$$

$$dV = rV\,dt$$

$$\left[\frac{\partial c}{\partial t} + \frac{1}{2} \frac{\partial^2 c}{\partial S^2} \sigma^2 S^2 + r \frac{\partial c}{\partial S} S \right] dt = rc.$$

DYNAMIC DELTA HEDGING

-Works perfectly under the theoretical assumptions of continuous time continuous price. Brilliant mathematical idea!!

- Finance is not pure mathematics!

-Models are only models: we can not claim a model is bad simply because there are breaks on it’s assumptions!

- ROBUST?? : to what we really can do and how real markets behave?

Merton 1998

"A broader, and still open, research issue is the robustness of the pricing formula in the absence of a dynamic portfolio strategy that exactly replicates the payoffs to the option security."

DELTA HEDGING IN PRACTICE

-Discrete time and price steps.

- Transaction costs, bid-offer spreads, etc.

- Empirically we have fat tailed distributions.

- Jumps in prices as well as stochastic volatility, and even jumps in volatility.

-Boyle and Emanuel (1980)

-Gilster (1990) Systematic risk

-Mello Neuhaus (1998)

-Derman and Kamal (1999)

-Wilmott (2000)

-Derman-Taleb (2005)

Substantial Transaction Costs

Basele, Shows, Thorpe (1983)

Gilster and Lee (1984)

Lealand (1985)

Hoggard, Whalley, and Wilmott (1994)

Kabanov and Safarian (1997)

Grandits and Schachinger (2001)

Not path dependent, only S at maturity counts

Need reasonable speed (100 000 simulations)

Monte Carlo simulation put option no hedge and static delta hedge, 100 000 simulations.

$(S = 100, T = 30/365, r = b = 0)$

Vol	Strike	Delta	Value	Stdev %	Max error	Stdev %	Max error
				No Hedging		Static Delta Hedge	
10%	100.0000	-49.4%	1.14369	144.2%	856.2%	75.7%	457.9%
10%	98.1252	-25.0%	0.43383	231.9%	2204.3%	166.3%	1520.6%
10%	96.4322	-10.0%	0.13736	395.9%	6740.3%	339.5%	4897.5%
30%	100.0000	-48.3%	3.43014	139.1%	855.5%	75.6%	548.0%
30%	94.7136	-25.0%	1.34026	221.0%	1817.0%	159.9%	1239.6%
30%	89.8953	-10.0%	0.42222	374.2%	5390.1%	321.4%	4600.8%
60%	100.0000	-46.6%	6.85394	132.8%	680.4%	75.5%	619.1%
60%	90.3727	-25.0%	2.80468	202.2%	1380.2%	148.6%	981.0%
60%	81.4117	-10.0%	0.87655	335.6%	3620.1%	291.3%	3036.7%

VERY RISKY BUSINESS!!

Risk can be good or Bad!

Delta Hedging Replication:

Now path dependent

Need good random number generator

Need a lot of speed C++ (100 000 simulations)

3 to 30 million random number per number in table

Table 1: Monte Carlo simulation of dynamic delta replication put option
($S = 100$, $T = 30/365$, $r = b = 0$)

Vol	Strike	Delta	Value	Stdev %	Max	Stdev %	Max	Stdev %	Max
				$n = 30$		$n = 60$		$n = 300$	
10%	100.0000	-49.4%	1.14369	15.6%	100.7%	11.2%	70.2%	5.1%	35.0%
10%	98.1252	-25.0%	0.43383	34.9%	304.6%	25.3%	246.7%	11.4%	88.9%
10%	96.4322	-10.0%	0.13736	76.9%	854.8%	55.1%	663.0%	24.8%	275.6%
30%	100.0000	-48.3%	3.43014	15.6%	123.1%	11.2%	80.4%	5.1%	37.9%
30%	94.7136	-25.0%	1.34026	34.2%	304.2%	24.4%	216.4%	11.0%	77.6%
30%	89.8953	-10.0%	0.42222	73.4%	763.6%	52.0%	518.6%	23.7%	269.4%
60%	100.0000	-46.6%	6.85394	15.7%	127.6%	11.1%	93.8%	5.1%	35.4%
60%	90.3727	-25.0%	2.80468	31.9%	436.7%	22.7%	352.2%	10.3%	83.7%
60%	81.4117	-10.0%	0.87655	67.0%	661.5%	48.0%	534.9%	21.9%	262.9%

5% | 10%

Derman-Kamal analytic approximation 1998

F 0-r

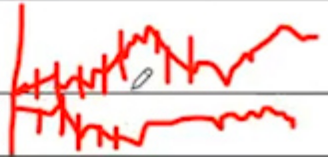
S r-r

$$\sigma_{P\&L} \approx \sqrt{\frac{\pi}{4}} S e^{(b-r)T} n(d_1)\sqrt{T}\frac{\sigma}{\sqrt{N}} = \sqrt{\frac{\pi}{4}} \text{Vega}\frac{\sigma}{\sqrt{N}}$$

Table 2: Derman-Kamal theoretical dynamic delta hedging replication error put option

$(S = 100,\ T = 30/365,\ r = 0)$

Vol	Strike	Delta	Value	Stdev % $n = 30$	Stdev % $n = 60$	Stdev % $n = 300$
10%	100.0000	-49.4%	1.14369	16.2%	11.4%	5.1%
10%	98.1252	-25.0%	0.43383	34.0%	24.0%	10.7%
10%	96.4322	-10.0%	0.13736	59.3%	41.9%	18.7%
30%	100.0000	-48.3%	3.43014	16.2%	11.4%	5.1%
30%	94.7136	-25.0%	1.34026	33.0%	23.3%	10.4%
30%	89.8953	-10.0%	0.42222	57.8%	40.9%	18.3%
60%	100.0000	-46.6%	6.85394	16.1%	11.4%	5.1%
60%	90.3727	-25.0%	2.80468	31.5%	22.3%	10.0%
60%	81.4117	-10.0%	0.87655	55.7%	39.4%	17.6%



A) VOLATILITY IS NOT CONSTANT IN DESCRETE TIME!!!

"sampling error" ? (yes and no)

Not really: Real effect: we can not hedge continuous in time!!

$$P\left[s\sqrt{\frac{(n-1)}{\chi^2_{(n-1;\alpha/2)}}} \leq \sigma \leq s\sqrt{\frac{(n-1)}{\chi^2_{(n-1;1-\alpha/2)}}}\right] = 1-\alpha$$

Example: N=50, 95% confidence interval, 30% volatility: 25.06% to 37.38%

N=300 : 95% confidence: 27.78% to 32.61%

B) Delta for OTM options very sensitive to volatility. High DDeltaDVol

$\frac{\partial^2 c}{\partial S \partial \sigma}$

$\Delta = 50\%$

$\sigma = 25\%$

$\Delta = 15\%$

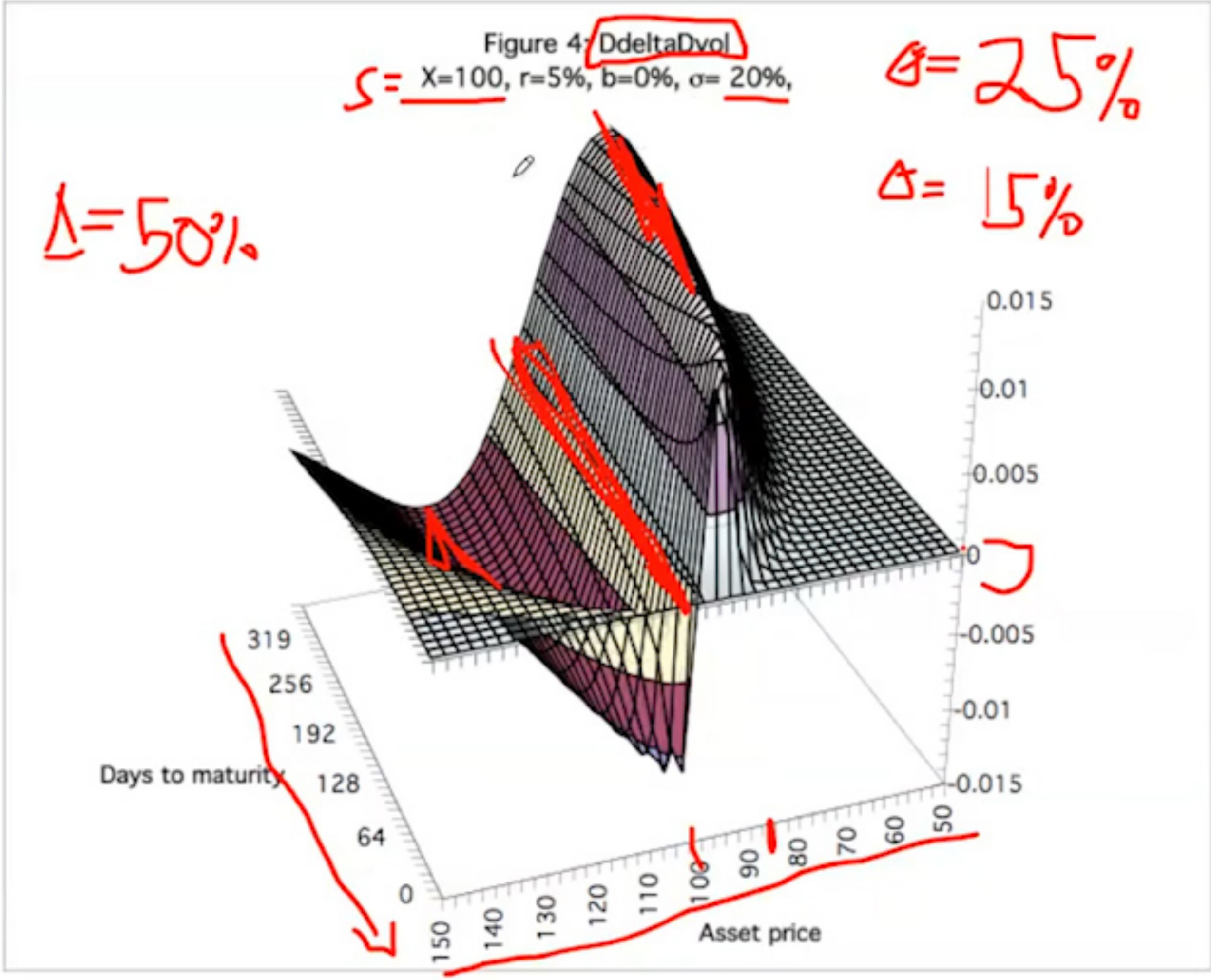


Figure 4: DdeltaDvol
X=100, r=5%, b=0%, σ= 20%,

SO FAR ONLY LOOKED AT GBM

What about stochastic volatility, jumps etc.?

We already got weak-form stochastic volatility or small price jumps from discrete time hedging.

Can only make things worse when we add jumps or stochastic volatility using BSM delta.

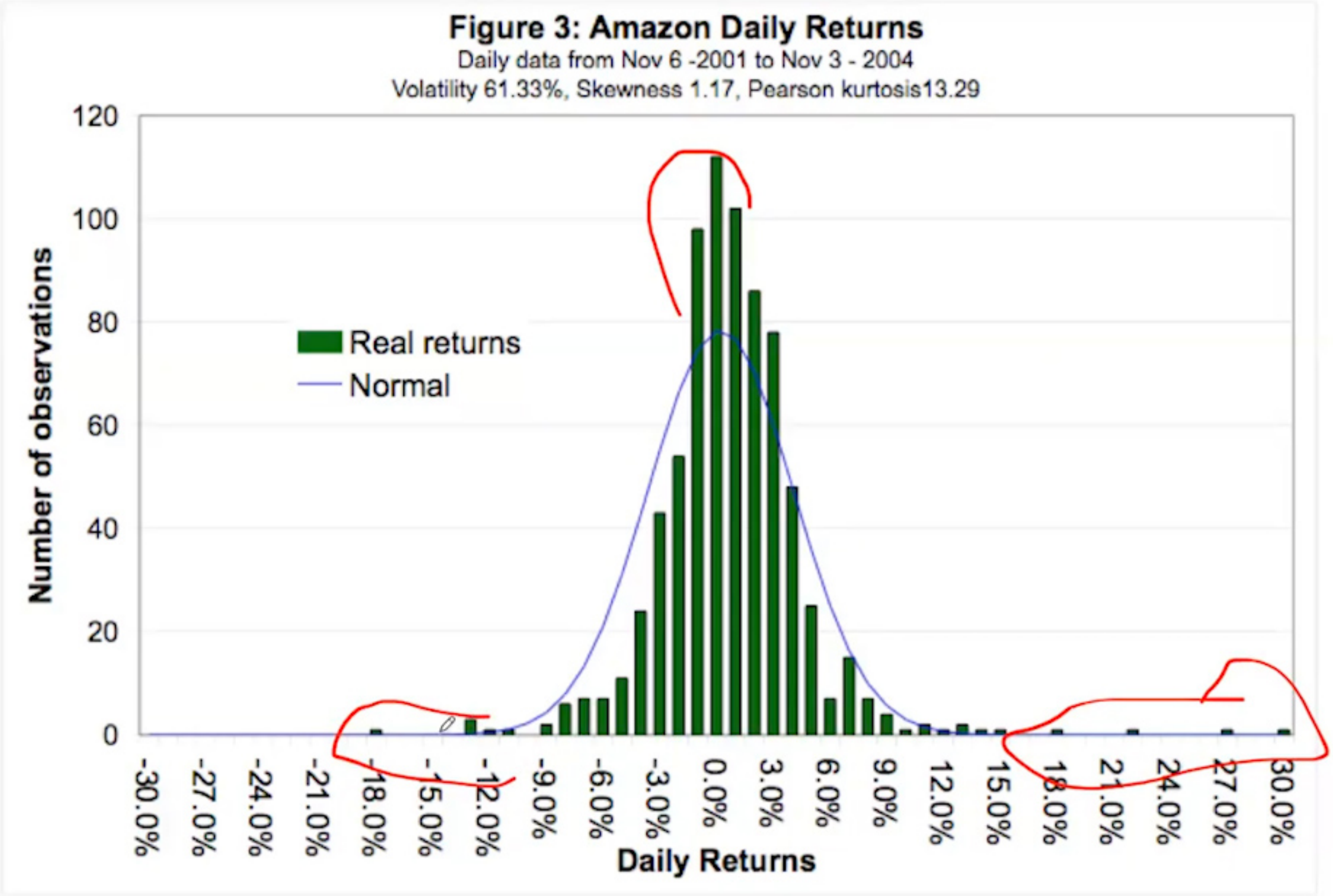


Figure 3: Amazon Daily Returns
Daily data from Nov 6 -2001 to Nov 3 - 2004
Volatility 61.33%, Skewness 1.17, Pearson kurtosis13.29
Number of observations
120
100
80
60
40
20
0
Real returns
Normal
-30.0%
-27.0%
-24.0%
-21.0%
-18.0%
-15.0%
-12.0%
-9.0%
-6.0%
-3.0%
0.0%
3.0%
6.0%
9.0%
12.0%
15.0%
18.0%
21.0%
24.0%
27.0%
30.0%
Daily Returns

What about jumps

Jumps at random time during option lifetime

100 000 simulations per run

Table 2: Monte Carlo simulation of dynamic delta replication put option with jumps

($S = 100$, $T = 30/365$, $r = b = 0$, $\sigma = 0.30$)

Jump Size	Strike	Delta	Value	Stdev %	Max	Stdev %	Max	Stdev %	Max
				$n = 30$		$n = 60$		$n = 300$	
3%	100.0000	-48.3%	3.43014	17.7%	141.6%	13.7%	104.2%	9.1%	73.7%
3%	94.7136	-25.0%	1.34026	41.6%	349.7%	32.0%	252.0%	20.9%	191.0%
3%	89.8953	-10.0%	0.42222	97.4%	1014.9%	74.7%	731.8%	47.4%	619.0%
5%	100.0000	-48.3%	3.43014	25.6%	173.1%	23.1%	160.0%	20.8%	127.1%
5%	94.7136	-25.0%	1.34026	62.3%	394.6%	55.1%	403.1%	48.9%	312.8%
5%	89.8953	-10.0%	0.42222	149.1%	1504.2%	131.3%	1040.0%	114.1%	950.4%
10%	100.0000	-48.3%	3.43014	73.7%	287.9%	73.6%	275.2%	73.6%	268.8%
10%	94.7136	-25.0%	1.34026	189.1%	773.7%	188.5%	658.1%	188.1%	680.9%
10%	89.8953	-10.0%	0.42222	478.8%	2314.6%	473.3%	2026.8%	468.9%	2089.0%

100

90

NON-ROBUST

Delta Hedging Replication:

Now path dependent

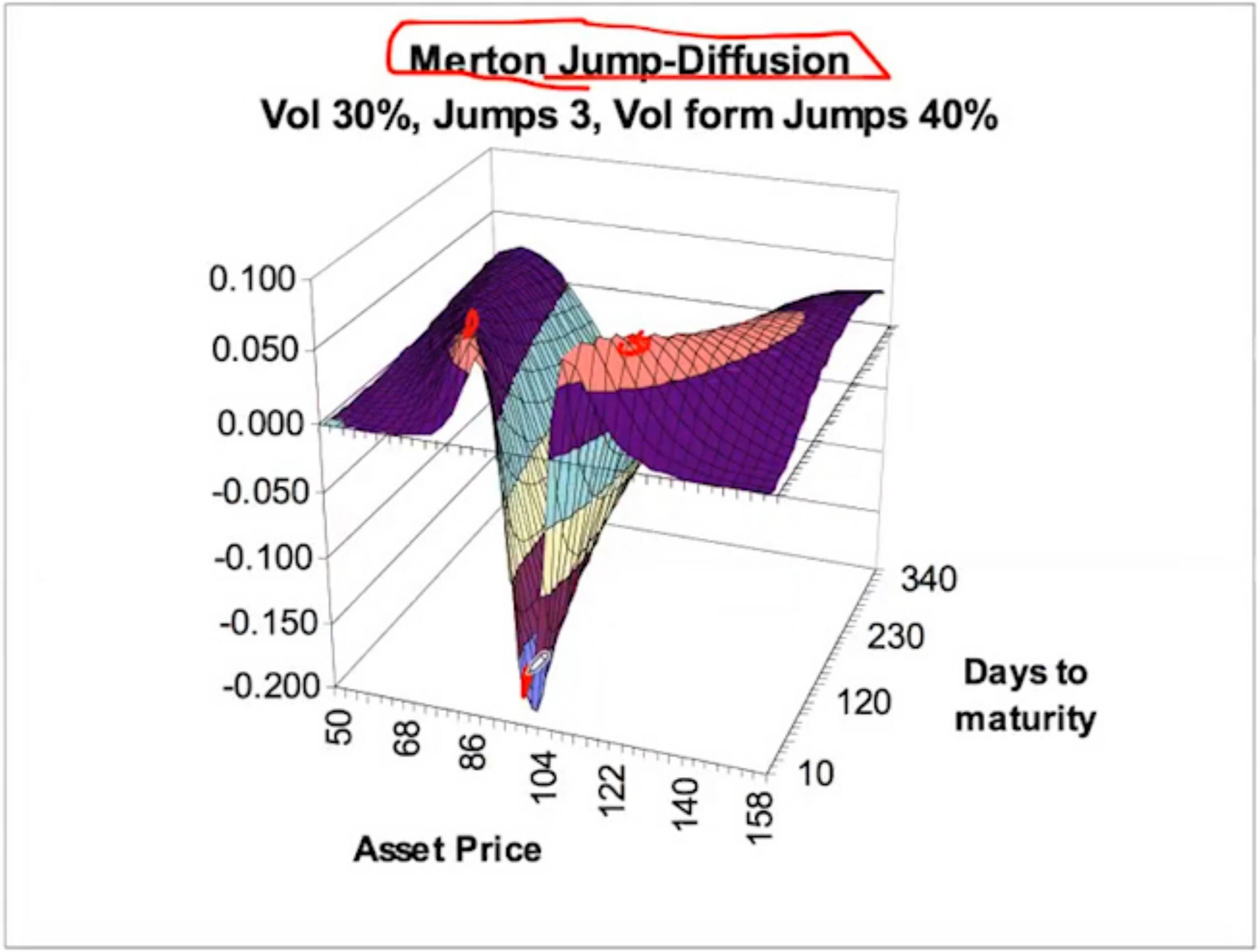
Need good random number generator

Need a lot of speed C++ (100 000 simulations)

3 to 30 million random number per number in table

Table 1: Monte Carlo simulation of dynamic delta replication put option ($S = 100$, $T = 30/365$, $r = b = 0$)

Vol	Strike	Delta	Value	Stdev %	Max	Stdev %	Max	Stdev %	Max
				$n = 30$		$n = 60$		$n = 300$	
10%	100.0000	-49.4%	1.14369	15.6%	100.7%	11.2%	70.2%	5.1%	35.0%
10%	98.1252	-25.0%	0.43383	34.9%	304.6%	25.3%	246.7%	11.4%	88.9%
10%	96.4322	-10.0%	0.13736	76.9%	854.8%	55.1%	663.0%	24.8%	275.6%
30%	100.0000	-48.3%	3.43014	15.6%	123.1%	11.2%	80.4%	5.1%	37.9%
30%	94.7136	-25.0%	1.34026	34.2%	304.2%	24.4%	216.4%	11.0%	77.6%
30%	89.8953	-10.0%	0.42222	73.4%	763.6%	52.0%	518.6%	23.7%	269.4%
60%	100.0000	-46.6%	6.85394	15.7%	127.6%	11.1%	93.8%	5.1%	35.4%
60%	90.3727	-25.0%	2.80468	31.9%	436.7%	22.7%	352.2%	10.3%	83.7%
60%	81.4117	-10.0%	0.87655	67.0%	661.5%	48.0%	534.9%	21.9%	262.9%



Call	
Asset price (S)	80.00
Strike price (X)	100.00
Time to maturity (T)	0.25
Risk-free rate (r)	5.00%
Volatility (σ)	30.00%
Jumps per year (λ)	3.00
Percent of total volatility (γ)	40.00%
Value	0.5255

Merton 1976

Bates Jump-Diffusion

- Jumps are allowed to be asymmetric, i.e. with non-zero mean.
- Since we often have to do with options on stock index futures (e.g. S&P index options), it is hardly plausible to maintain Merton's simplifying assumption that jump risk is idiosyncratic and thus fully diversifiable.

The Bates (1991) jump-diffusion model is consistent with an asymmetric volatility smile (generated form a BSM type model). This is often what we observe in practice.

The “many world interpretation of Jumps”

Diversification, do the professional diversify?
Much less than many would like to think!!

ARCH, GARCH, STOCHASTIC VOL, JUMP MODELS OF ANY USE??

Yes: helps us understand what stochastic process causes fat tails, helps us in risk management..help us value out-of-the-money options. Help us improve delta-hedge.

Do not improved dynamic delta hedging enough to save risk-neutrality.

They are all good fudge models!! (rooted in Gaussian??)

Static/Semi-Static Hedging

Higgins 1902/ Nelson 1904 Put-Call Parity Fully Understood!!

Mello Neuhaus (1998) Hedging options with options

Carr and Wu (2002)

"Put Call Reversal" by Peter Carr and Jesper Andreasen April 25, 2002

Derman-Taleb 2005: only for European options "Complete Static Hedging Argument"

Exotic options: Carr 1994, Derman 1995, Haug 1998 and many others, also variance swaps and volatility swaps.

MODELS ARE ONLY MODELS In practice: semi-static hedging **not** same strike on call and put **different** maturities

BUT THIS IS A ROBUST MODEL!!!!!

Most brilliant ideas often diffuse start:
Black-Scholes-Merton not first model for pricing derivatives based on arbitrage and risk-neutrality:

Forward price must be risk-neutral* , due to covered arbitrage

$$F = Se^{bT}$$

John Maynard Keynes., A Tract on Monetary Reform, 1923 (2000) (Prometheus Books: Amherst).

Blau, G. “*Some Aspects of the Theory of Futures Trading.*” The Review of Economic Studies XII (1944-45), 1-30.

With no delta hedging: Derman-Taleb argument:

$$c = e^{-RT}(E(S-X)_+) = e^{-RT}[S^{\mu T}N(d_1) - XN(d_2)]$$

Put-call parity long call + short put = forward, so options must also be priced with discount rate = r and drift in stock =r

$$c = e^{-rT}[S^{rT}N(d_1) - XN(d_2)] = SN(d_1) - Xe^{-rT}N(d_2)$$

Derman-Taleb Criticized October 2006

Not consistent with specific well known equilibrium model! ?

Not consistent with different discount rate for call and put ?

Extended Argument: If hedging options with options and synthetic delta-replication is far from fully available then demand and supply of options will have to affect option values!

$$c = SN(d_1) - Xe^{-rT}N(d_2)$$

$$d_1 = \frac{\ln(S/X) + (r + \sigma_i^2/2)T}{\sigma_i\sqrt{T}}$$

$$d_2 = d_1 - \sigma_i\sqrt{(T)}$$

But can we just do this, is this not a crank model? Normally yes, but not in this case!!! All we say vol must not break with PCP

Black-Scholes/Merton: Dynamic Hedging model

Brilliant mathematical idea

NOT ROBUST IN PRACTICE!

Static/semi-static hedging:

VERY ROBUST IN PRACTICE!

Robust for discrete time and price moves

Robust for stochastic volatility

Robust for jumps

What traders actually try to use!

"In the end, a theory is accepted not because it is confirmed by conventional empirical tests, but because researchers persuade one another that the theory is correct and relevant."

Fischer Black

Summary

- Dynamic delta hedging brilliant mathematical idea!

- Dynamic hedging reduces risk considerably compared to no hedging or static delta hedging!

- Dynamic delta hedging do not do what it promises!

Fischer Black was himself skeptical to dynamic hedging as solution to risk-neutrality argument (?)

- Dynamic delta hedging in practice is not in any way good enough to use risk-neutral valuation!

-Dynamic delta hedging works very poorly for OTM options.

What to do

- Hedging options with options, supply demand driven.

-Truncating your tails.

-Delta hedging to remove risk, but not to relay on risk-neutral valuation.

- Only stay long options? No, can take in option premium.

- Remove risk in robust way and/or construct portfolio in such a way that can live with risk.

References

ARROW, K. (1953): *The Role of Securities in the Optimal Allocation of Risk-Bearing*. Econometrie.

BATES, D. S. (2005): "Hedging the Smirk," *Finance Research Letters*, 2(4), 195–200.

BLACK, F., AND M. SCHOLES (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

BLAU, G. (1944-1945): "Some Aspects of The Theory of Futures Trading," *The Review of Economic Studies*, 12(1).

BONESS, A. (1964): "Elements of a Theory of Stock-Option Value," *Journal of Political Economy*, 72, 163–175.

BOYLE, P., AND D. EMANUEL (1980): "Discretely Adjusted Option Hedges," *Journal of Financial Economics*, 8, 259–282.

CARR, P. (2002): "Static Hedging of Standard Options," *Working paper, Courant Institute, New York University.*

COX, J. C., S. A. ROSS, AND M. RUBINSTEIN (1979): "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7, 229–263.

DERMAN, E. (2004): *My Life as a Quant*. New York: John Wiley & Sons.

DERMAN, E., AND M. KAMAL (1999): "When You Cannot Hedge Continuously, The Corrections of Black-Scholes," *Risk Magazine*, 12, 82–85.

DERMAN, E., AND N. TALEB (2005): "The Illusion of Dynamic Delta Replication," *Quantitative Finance*, 5(4), 323–326.

GALAI, D. (1983): "The Components of Return from Hedging Options Against Stocks," *Journal of Business*, 56(1), 45–54.

GILSTER, J. E. (1990): "The Systematic Risk of Discretely Rebalanced Options Hedges," *Journal of Financial and quantitative analysis*, 25(4), 507–516.

HAUG, E. G. (1997): *The Complete Guide To Option Pricing Formulas*. McGraw-Hill, New York.

HOGGARD, T., A. E. WHALLEY, AND P. WILMOTT (1994): "Hedging Option Portfolios in the Presence of Transaction Costs," *Advances in Futures and Options Research*, 7, 21–35.

KEYNES, J. M. (1924): *A Tract on Monetary Reform*. Re-printed 2000. Amherst New York: Prometheus Books.

LELAND, H. (1985): ""Option Pricing and Replication with Transactions Costs"," *Journal of Finance*, XL(5), 1283–1301.

MANDELBROT, B. (1963): "The Variation of Certain Speculative Prices," *Journal of Business*, 36, 394–419.

MELLO, A. S., AND H. J. NEUHAUS (1992): "Market Making in the Options Markets and the Costs of Discrete Hedge Rebalancing," *Journal of Finance*, 47(2), 765–779.

——— (1998): "A Portfolio Approach to Risk Reduction in Discretely Rebalanced Option Hedges," *Management Science*, 44(7), 921–934.

MERCURIO, F., AND T. VORST (1996): "Option Pricing with Hedging at Fixed Trading Dates," *Applied Mathematical Finance*, 3, 135–158.

MERTON, R. C. (1973): "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4, 141–183.

——— (1976): "Option Pricing When Underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, 3, 125–144.

MITCHELL, WESLEY, C. (1915): "The Making and Using of Index Numbers," *Introduction to Index Numbers and Wholesale Prices in the United States and Foreign Countries (published in 1915 as Bulletin No. 173 of the U.S. Bureau of Labor Statistics, reprinted in 1921 as Bulletin No. 284).*

RENDLEMAN, R. J., AND B. J. BARTTER (1979): "Two-State Option Pricing," *Journal of Finance*, 34, 1093–1110.

RUBINSTEIN, M. (2006): *A History of The Theory of Investments*. New York: John Wiley & Sons.

SAMUELSON, P. (1965): "Rational Theory of Warrant Pricing," *Industrial Management Review*, 6, 13–31.

SCHACHTER, B., AND R. ROBINS (1994): "An Analysis of the Risk in Discretely Rebalanced Option Hedges and Delta-Based Techniques," *Management Science*, 40(6), 798–808.

SPRENKLE, C. (1960): *Warrant Prices as Indicators of Expectations and Preferences*. Dr. Philosophy thesis Yale University.

——— (1964): *Warrant Prices as Indicators of Expectations and Preferences: P. Cootner, ed., 1964, The Random Character of Stock Market Prices*,MIT Press, Cambridge, Mass.

THORP, E. O., AND S. T. KASSOUF (1967): *Beat the Market*. New York: Random House.

WILMOTT, P. (2006): *Paul Wilmott on Quantiative Finance, Second Edition*. New York: John Wiley & Sons.

ARCH, GARCH, STOCHASTIC VOL, JUMP MODELS OF ANY USE??

Yes: helps us understand what stochastic process causes fat tails, helps us in risk management..help us value out-of-the-money options. Help us improve delta-hedge.

Do not improved dynamic delta hedging enough to save risk-neutrality.

They are all good fudge models!! (rooted in Gaussian??)

Know Your Weapon

Dr. Espen Gaarder Haug

CQF, Fitch, March 26 – 2020

Hello

Market Formula (Bachelier-Thorp)

$$c = Se^{(b-r)T}N(d_1) - Xe^{-rT}N(d_2)$$

$$p = Xe^{-rT}N(-d_2) - Se^{(b-r)T}N(-d_1)$$

Where:

$$d_1 = \frac{\ln(S/X) + (b + \sigma_{X,T}^{2}/2)T}{\sigma_{X,T}\sqrt{T}}$$

$$d_2 = d_1 - \sigma_{X,T}\sqrt{T}$$

S = Asset price
X = Strike
T = Years to maturity
r = risk-free-rate
b = cost-of-carry
$\sigma_{X,T}$ = volatility that can be different for each strike and maturity
See Haug 2007 "Derivatives Models on Models" chapter 2

Black-Scholes-Merton

$$c = Se^{(b-r)T}N(d_1) - Xe^{-rT}N(d_2)$$

$$p = Xe^{-rT}N(-d_2) - Se^{(b-r)T}N(-d_1)$$

Where:

$$d_1 = \frac{\ln(S/X) + (b + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

S = Asset price

X = Strike

T = Years to maturity

r = risk-free rate

b = cost-of-carry

σ = volatility

U

F

$b = 0$ F Black 76

$b = r$ S B-S-M 73

$b = r_d - r_f$ S_{FX} G-K 84

$b = r - q$ S_i M-73

Delta Greeks

- Delta
- Delta mirror strikes
- Strike from delta
- Elasticity

Delta higher than one

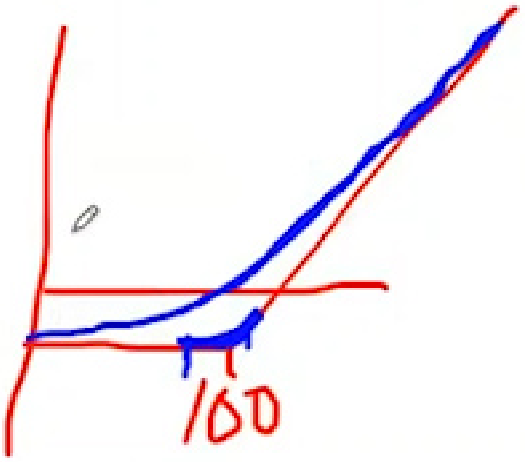
$$\Delta_{call} = \boxed{\frac{\partial c}{\partial S}} = \boxed{e^{(b-r)T}}\,\boxed{N(d_1)}$$

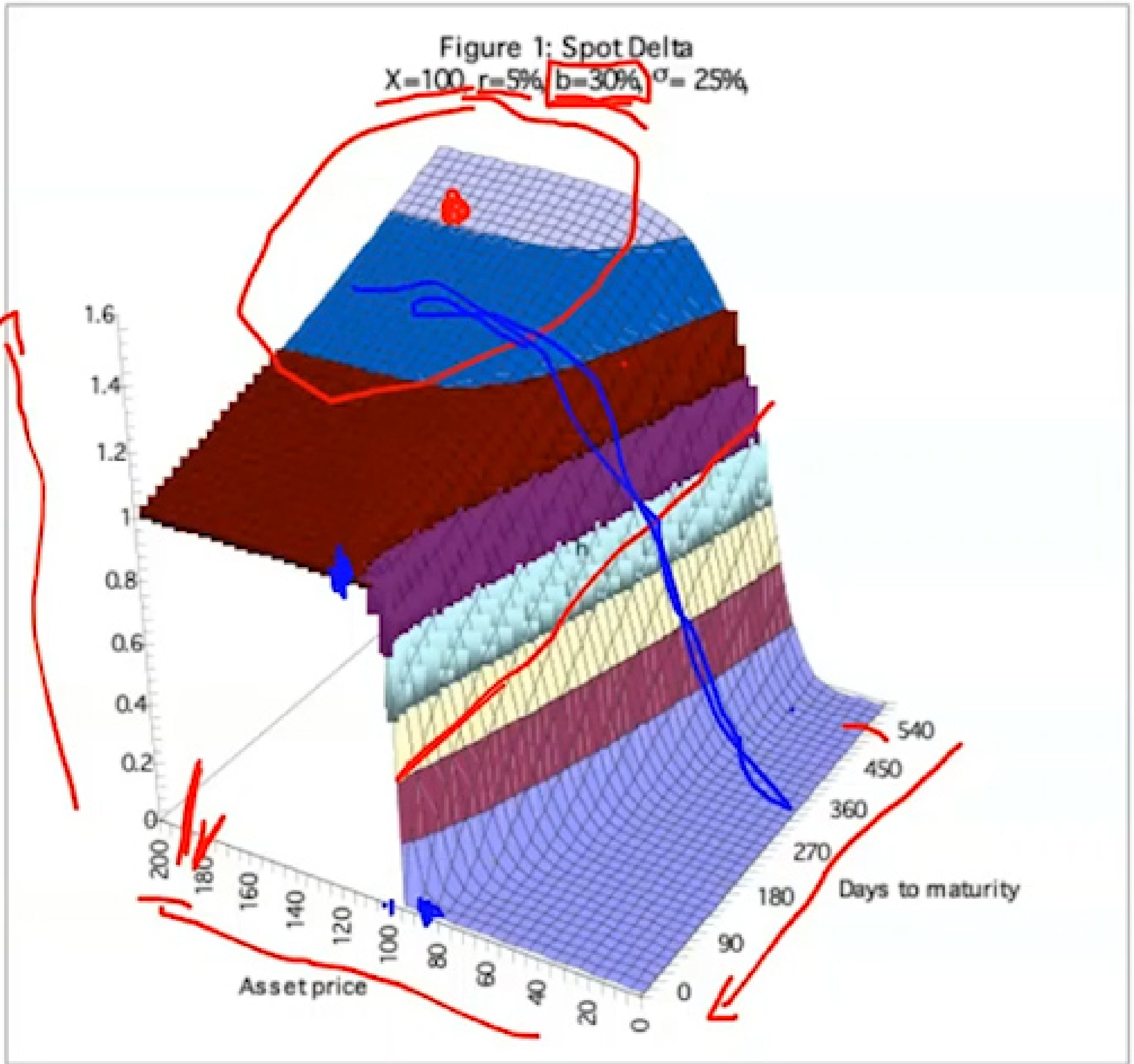
$$\Delta_{put} = \frac{\partial p}{\partial S} = \boxed{-e^{(b-r)T} N(-d_1)}$$

$N(d_1) \le 1$

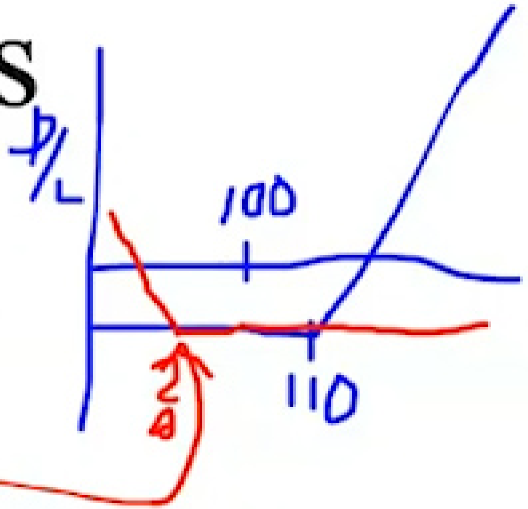
$b > r$

$e^{0 \cdot T} = 1$





Delta Mirror Strikes



100^2

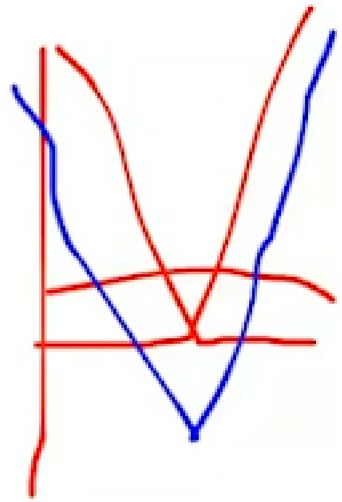
$$X_P = \frac{S^2}{\boxed{X_C}} e^{(2b+\sigma^2)T}, \quad X_C = \frac{S^2}{X_P} e^{(2b+\sigma^2)T}$$

$= 110$

Special case delta symmetric straddle (Wystrup(1999)):

$$\boxed{X_C = X_P} = \boxed{Se^{(b+\sigma^2/2)T}}$$

$\ln(S/X)$



Delta symmetric asset: $S = Xe^{(-b-\sigma^2/2)T}$

At this strike the delta is $\Delta_C = \frac{e^{(b-r)T}}{2}, \quad \Delta_P = -\frac{e^{(b-r)T}}{2}$

$$c = \frac{Se^{(b-r)T}}{2} - X^{-rT}N(-\sigma\sqrt{T}), \quad p = X^{-rT}N(\sigma\sqrt{T}) - \frac{Se^{(b-r)T}}{2}$$

Strikes from delta

FX 95%
· Interbank

Wystrup(1999):

$$\boxed{X_C} = S\exp[N^{-1}(\Delta_C e^{(r-b)T})\sigma\sqrt{T} + (b+\sigma^2/2)T]$$

USD/JPY
100

$$X_P = S\exp[N^{-1}(-\Delta_P e^{(r-b)T})\sigma\sqrt{T} + (b+\sigma^2/2)T]$$

C USD/JPY
T = 3M

Robust and accurate approximation of inverse cumulative normal distribution needed, Moro (1995).

$\boxed{\Delta = 25\%}$

110
105

σ = 20% - 21%

DdeltaDvol

$$\frac{\partial c}{\partial S \partial \sigma} = \frac{\partial p}{\partial S \partial \sigma} = \frac{-e^{(b-r)T} d_2}{\sigma} n(d_1)$$

Maximal value at

$$S_L = X e^{-bT - \sigma \sqrt{T} \sqrt{4 + T\sigma^2}/2}$$

Minimal value at

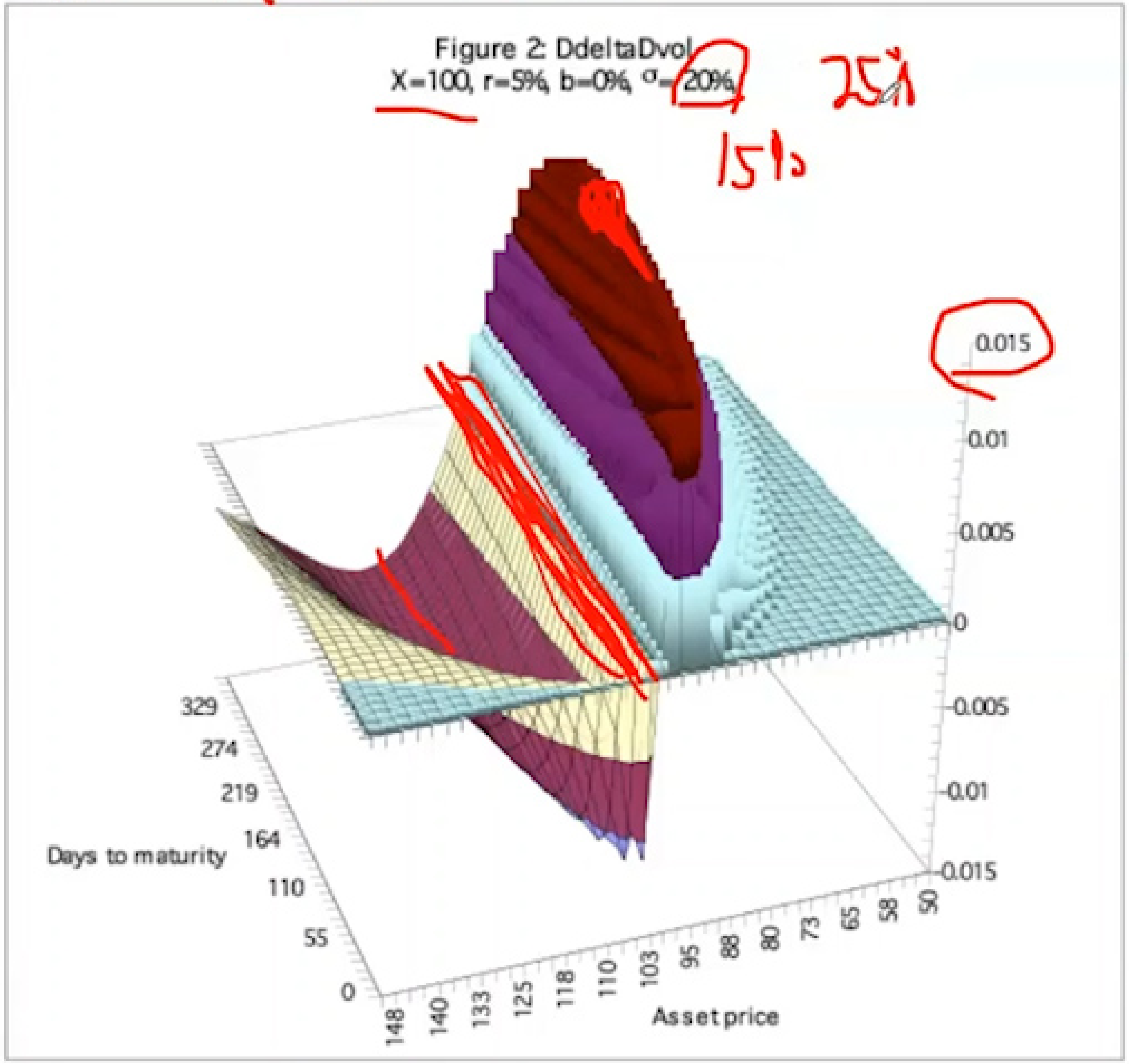
$$S_U = X e^{-bT + \sigma \sqrt{T} \sqrt{4 + T\sigma^2}/2}$$

Minimal value at

$$X_L = S e^{bT - \sigma \sqrt{T} \sqrt{4 + T\sigma^2}/2}$$

Maximal value at

$$X_U = S e^{bT + \sigma \sqrt{T} \sqrt{4 + T\sigma^2}/2}$$



Useful Tools

- A library
- Paper and pencil
- Mathematica
- Maple
- Matlab (?)
- Others ?

Implementation:

VBA, VB, C/C++, Java….

you name it

Elasticity

$$\Lambda_{call} = \Delta_{call}\frac{S}{call} > 1, \quad \Lambda_{put} = \Delta_{put}\frac{S}{put} < -1$$

Option volatility: $\sigma_O \approx \sigma\,|\Lambda|$ Compound options $S = C$ σ_O

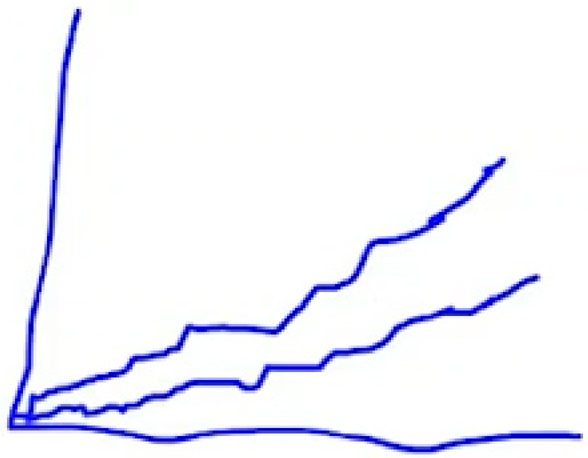
Option Beta, expected return satisfy the CAPM equation (Merton-71):

$$E[return] = r + E[r_m - r]\beta_i \quad > \beta$$

$$\beta_C = \frac{S}{call}\Delta_C\beta_S = \Lambda_C\beta_S, \quad \beta_P = \frac{S}{put}\Delta_P\beta_S = \Lambda_P\beta_S$$

Option Sharpe ratios

$$\frac{\mu_O - r}{\sigma_O} = \frac{\mu_S - r}{\sigma}$$



Smile?

Elasticity

$$\Lambda_{call} = \Delta_{call}\frac{S}{call}, \quad \Lambda_{put} = \Delta_{put}\frac{S}{put}$$

Option volatility: $\sigma_O \approx \sigma\,|\Lambda|$ Compound options

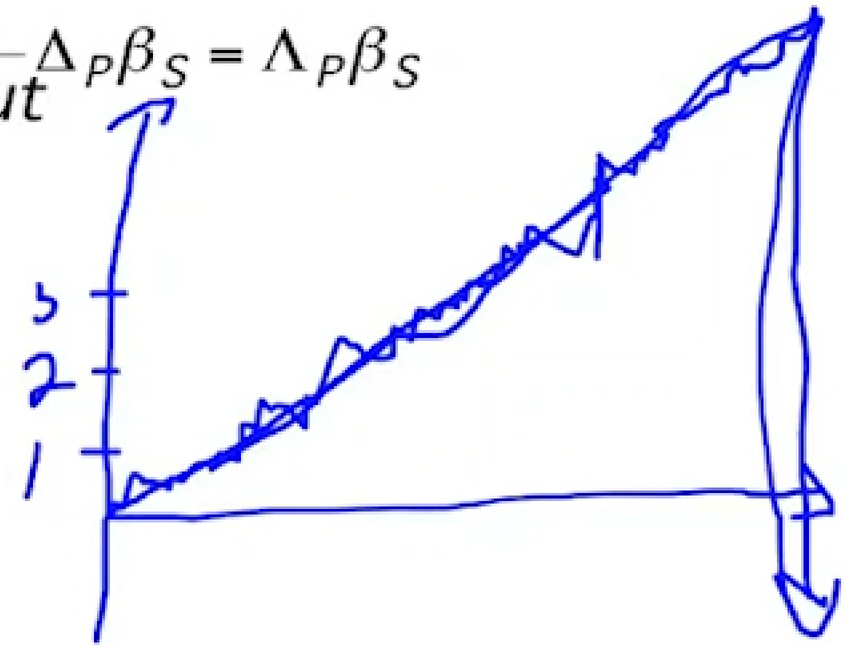
Option Beta, expected return satisfy the CAPM equation (Merton-71):

$$E[return] = r + E[r_m - r]\beta_i$$

$$\beta_C = \frac{S}{call}\Delta_C\beta_S = \Lambda_C\beta_S, \quad \beta_P = \frac{S}{put}\Delta_P\beta_S = \Lambda_P\beta_S$$

Option Sharpe ratios

$$\frac{\mu_O - r}{\sigma_O} = \frac{\mu_S - r}{\sigma}$$



Smile?

Charm

$$\frac{\partial \Delta_C}{\partial T} = -e^{(b-r)T}\left[n(d_1)\left(\frac{b}{\sigma\sqrt{T}} - \frac{d_2}{2T}\right) + (b-r)N(d_1)\right]$$

$$\frac{\partial \Delta_P}{\partial T} = -e^{(b-r)T}\left[n(d_1)\left(\frac{b}{\sigma\sqrt{T}} - \frac{d_2}{2T}\right) - (b-r)N(-d_1)\right]$$

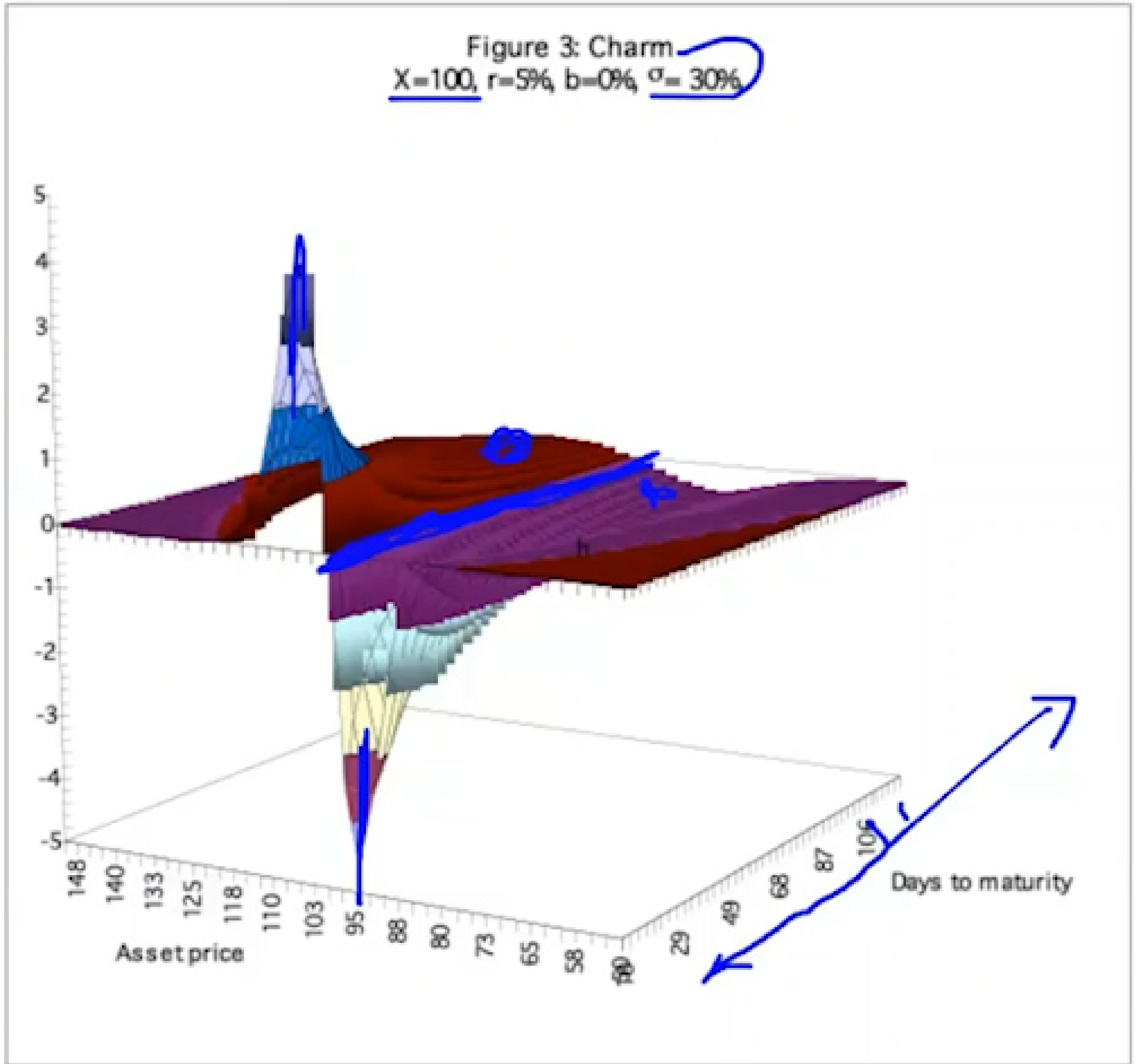
100

$S \uparrow\downarrow$

$\sigma \uparrow\downarrow$

$T \downarrow$ 2W

14d



Gamma Greeks

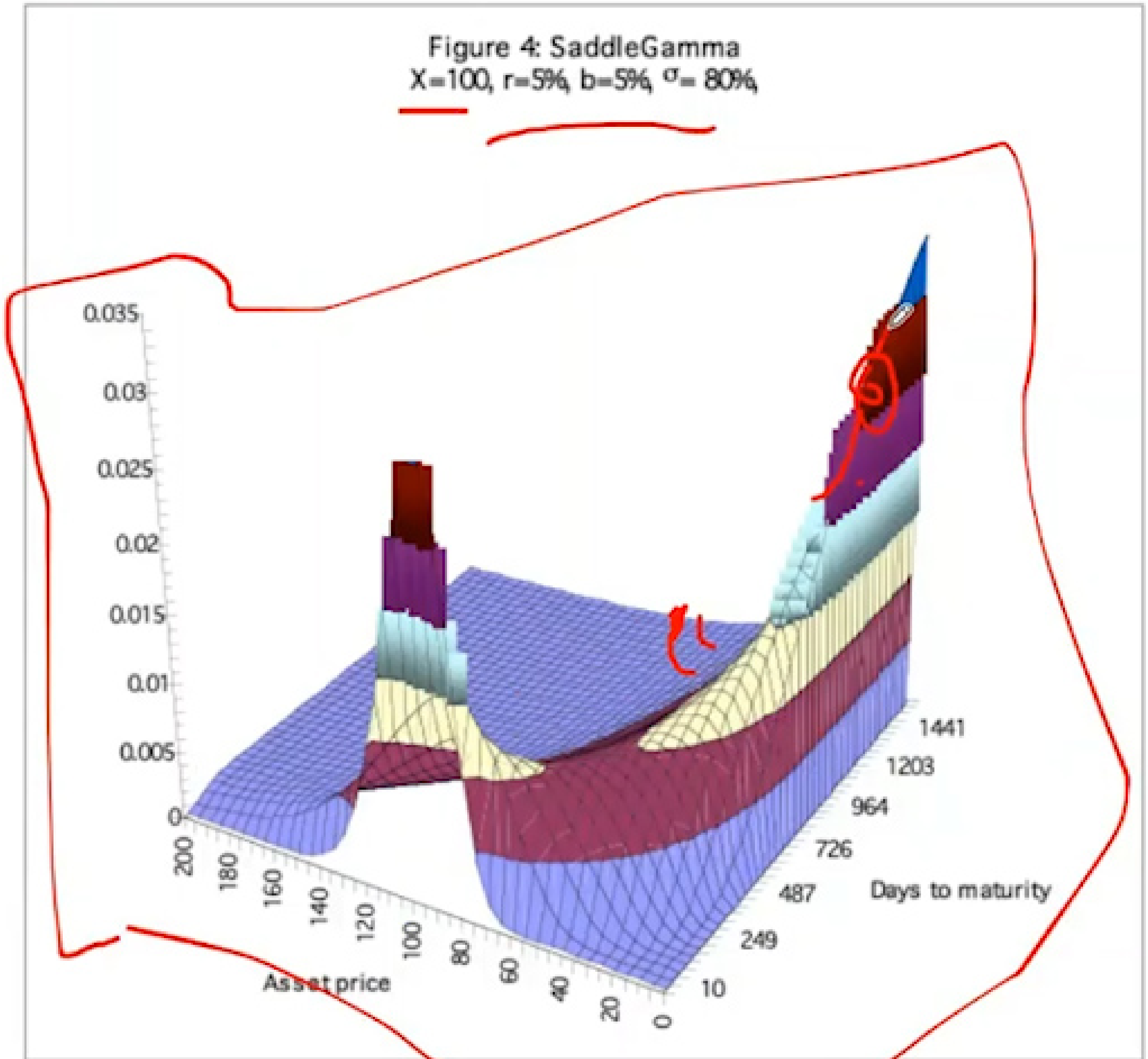
- Gamma
- Saddle gamma
- GammaP
- Gamma symmetry
- DGammaDVol
- DGammaDspot
- DGammaDTime

Saddle Gamma
Alexander Adamchuk www.wilmott.com

$$T_\Gamma = \frac{1}{2(\sigma^2 + b)}$$

$$S_\Gamma = X e^{(-b-3\sigma^2/2)T_S}$$

$$\Gamma_S = \frac{e^{(b-r)T}\sqrt{\frac{e}{\pi}}\sqrt{\frac{b}{\sigma^2}+1}}{X}$$



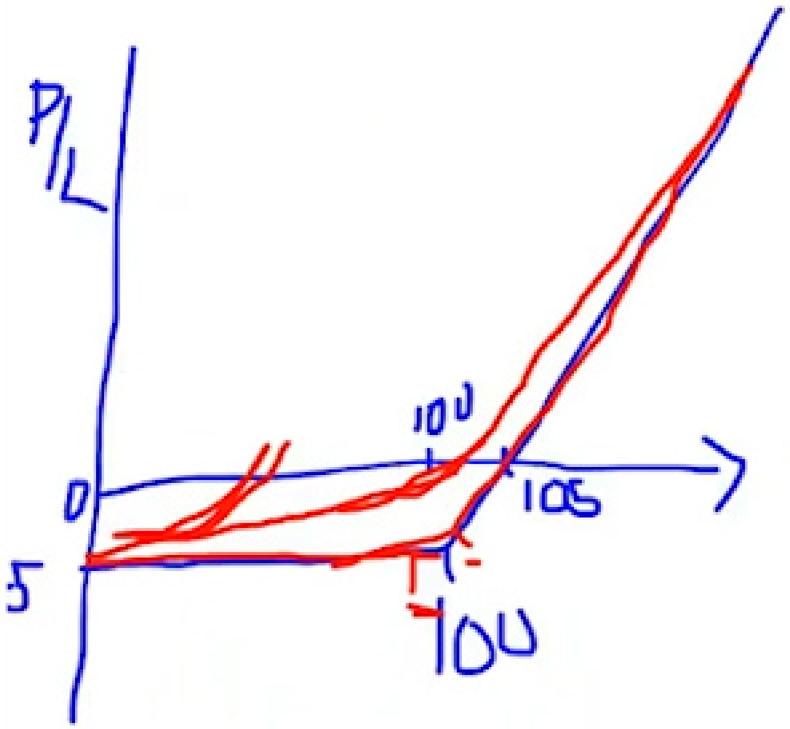
GammaP

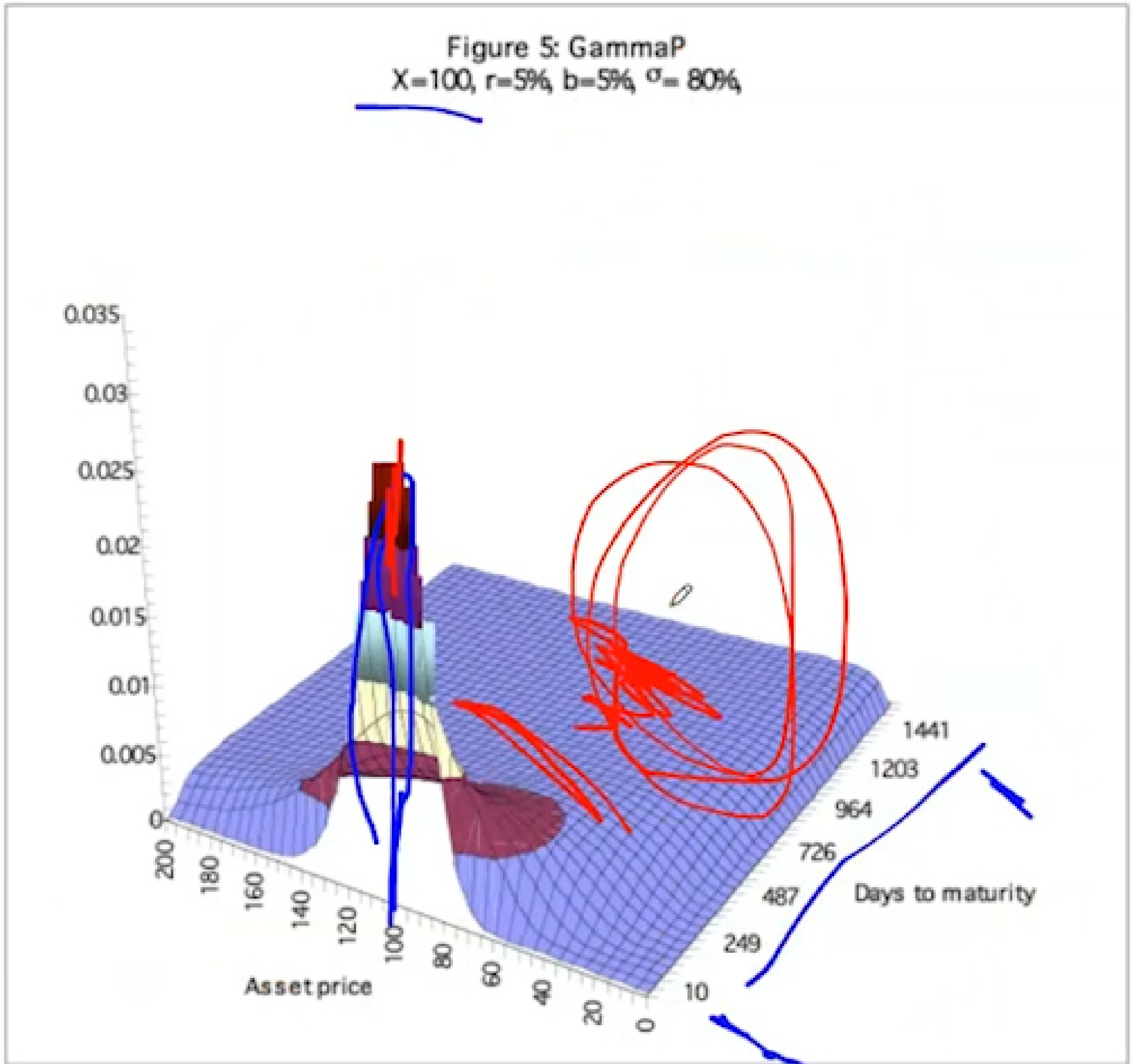
$$\Gamma_P = \Gamma \frac{S}{100}$$

Max GammaP at

$$S = Xe^{(-b-\sigma^2/2)T}$$

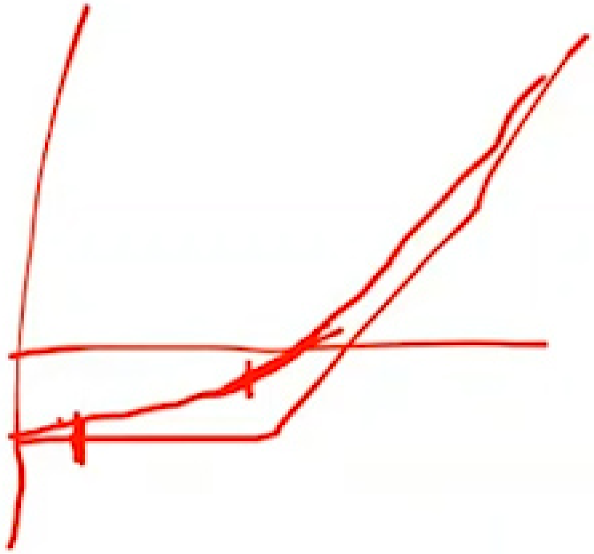
$$X = Se^{(b+\sigma^2/2)T}$$

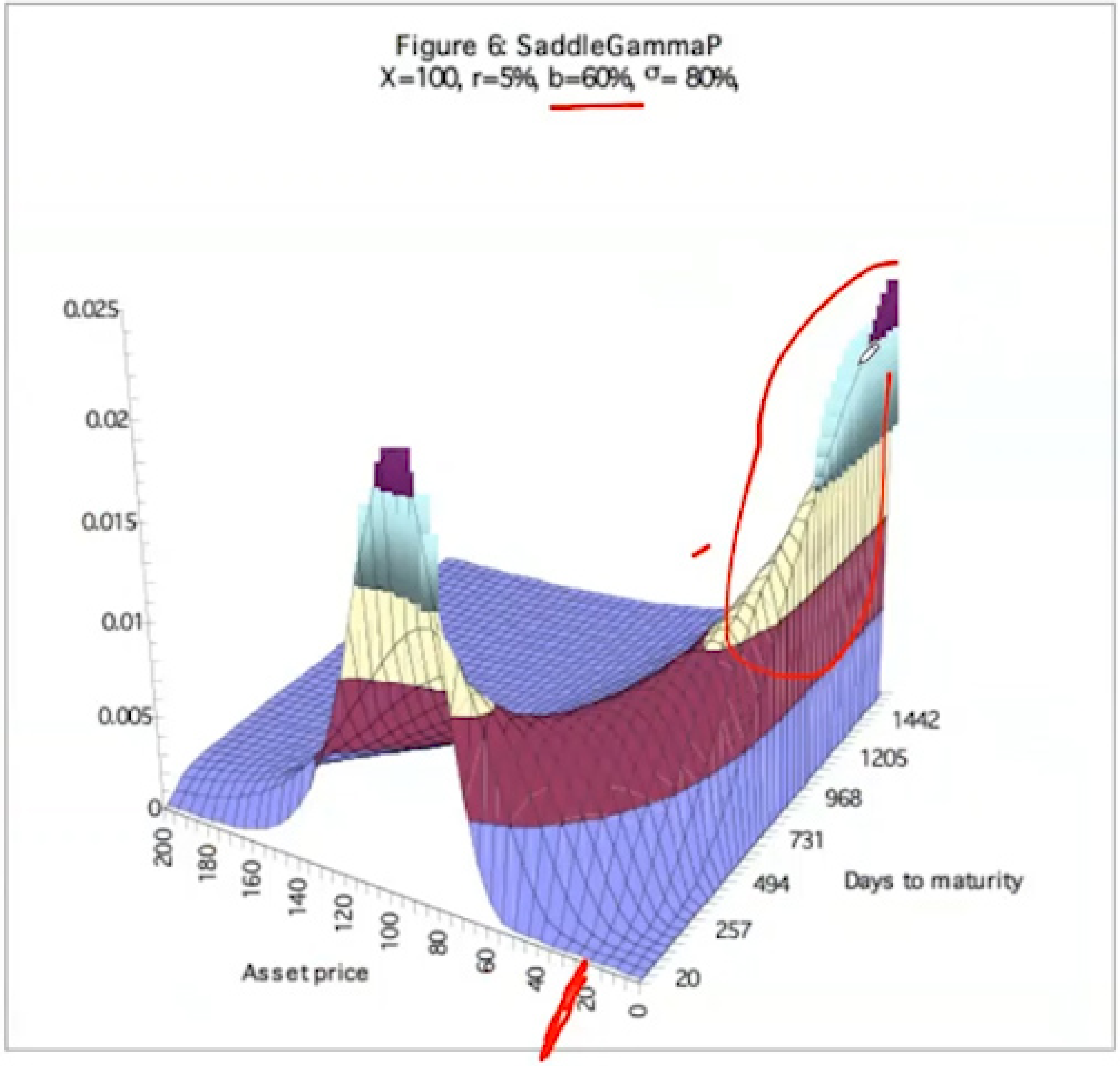




Saddle GammaP

- •Spot gamma
- •Forward gamma





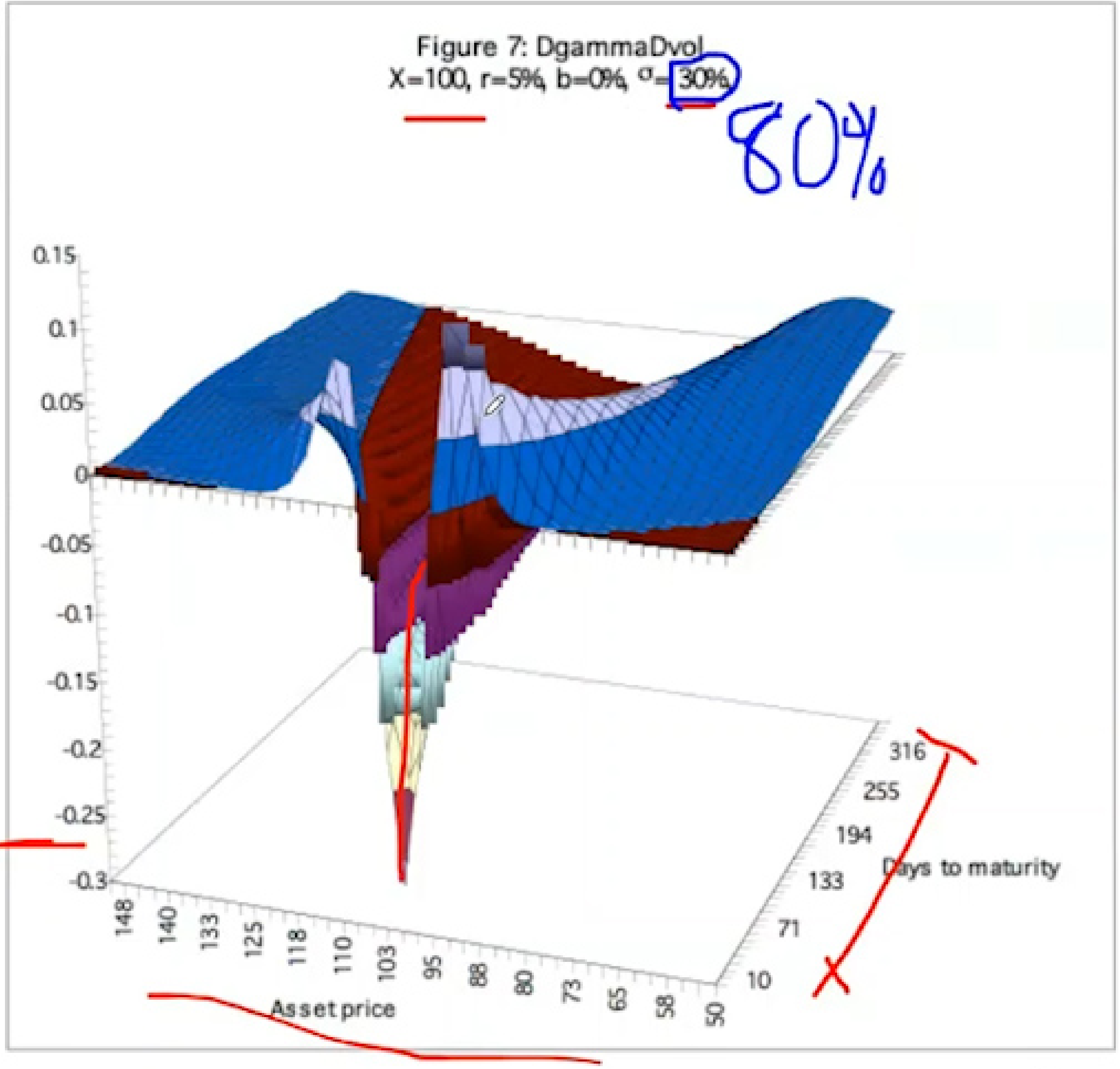
DgammaDvol

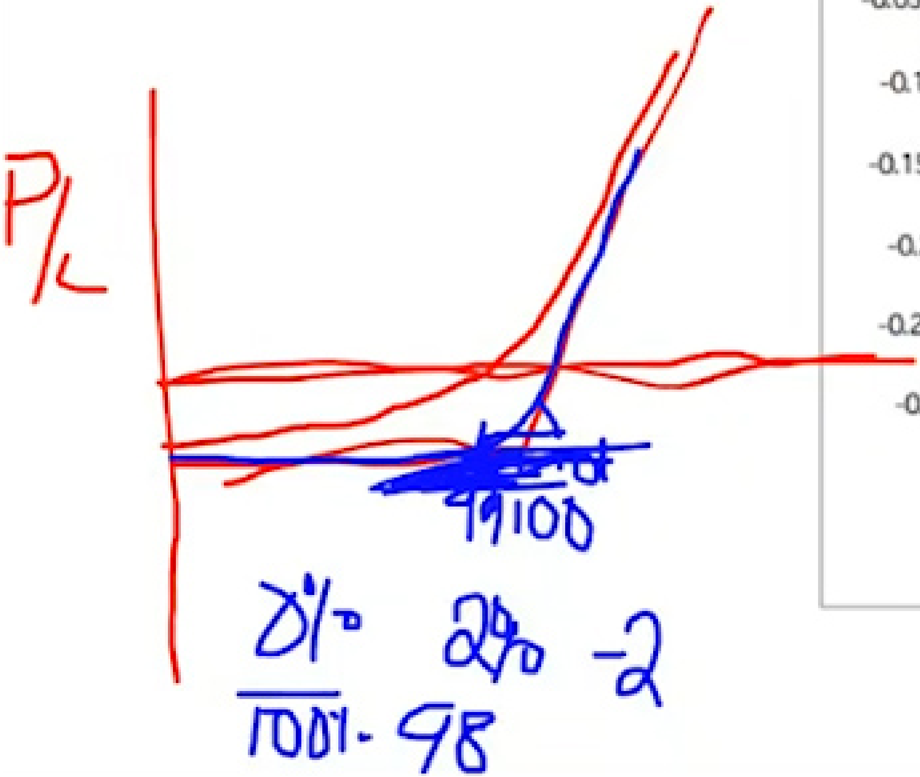
$$\frac{\partial \Gamma}{\partial \sigma} = \Gamma\left(\frac{d_1 d_2 - 1}{\sigma}\right) \qquad \frac{\partial \Gamma_P}{\partial \sigma} = \Gamma_P\left(\frac{d_1 d_2 - 1}{\sigma}\right)$$

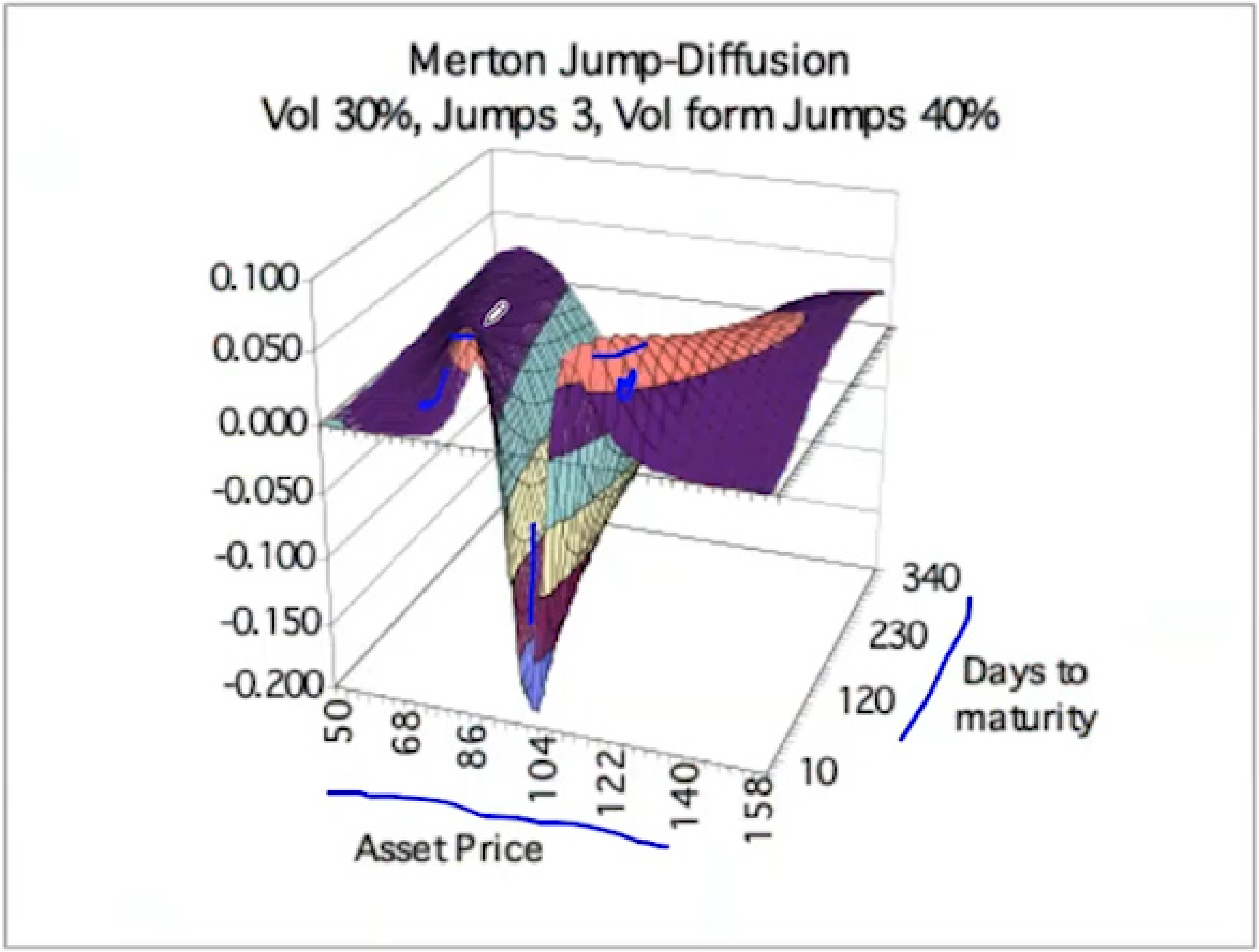
Positive outside interval

$$S_L = Xe^{-bT - \sigma\sqrt{T}\sqrt{4 + T\sigma^2}/2}$$

$$S_U = Xe^{-bT + \sigma\sqrt{T}\sqrt{4 + T\sigma^2}/2}$$





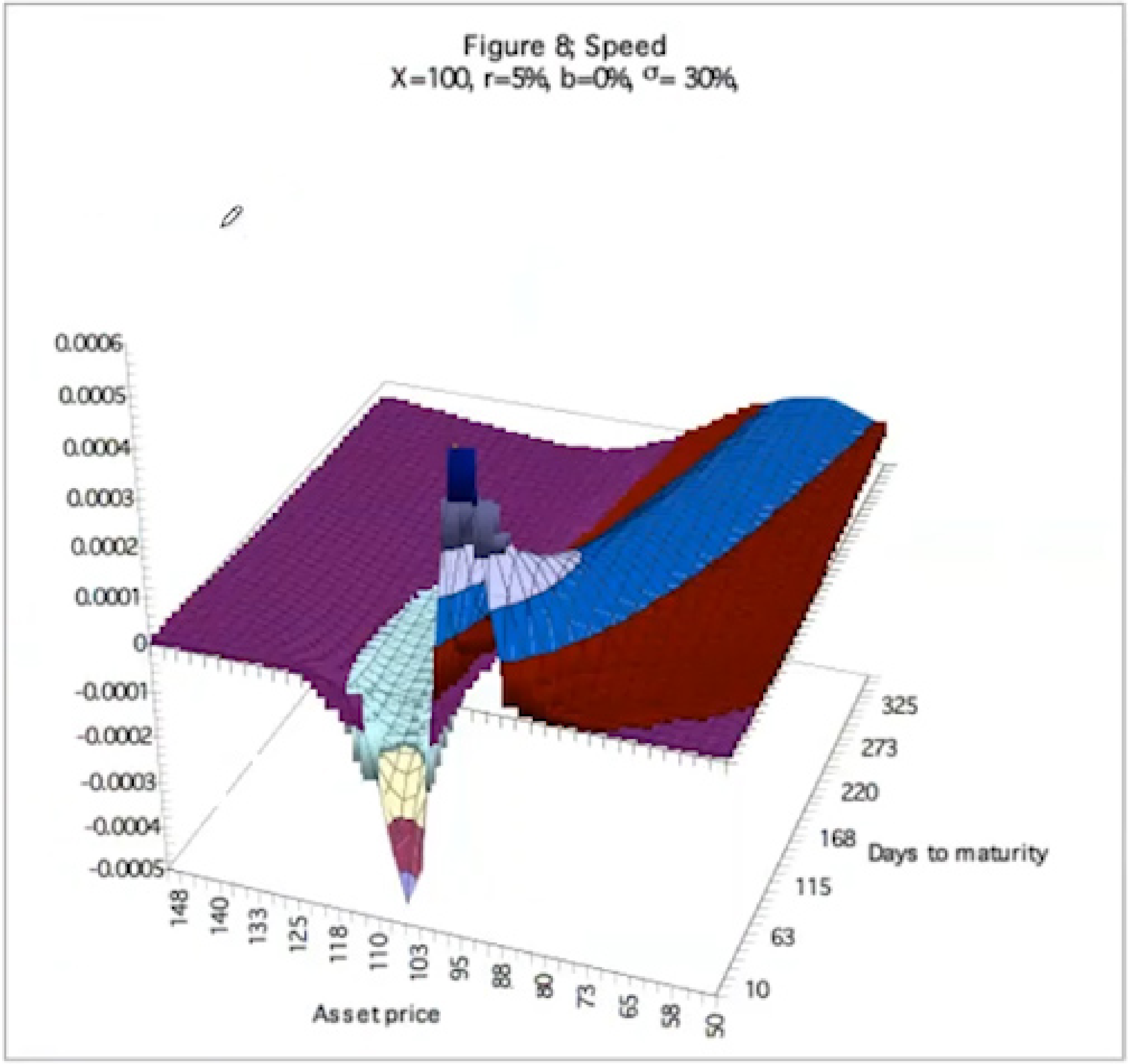


Call	
Asset price (S)	80.00
Strike price (X)	100.00
Time to maturity (T)	0.25
Risk-free rate (r)	5.00%
Volatility (σ)	30.00%
Jumps per year (λ)	3.00
Percent of total volatility (γ)	40.00%
Value	0.5255

Speed (DgammaDspot)

$$\frac{\partial^3 c}{\partial S^3} = -\frac{\Gamma\left(1 + \frac{d_1}{\sigma\sqrt{T}}\right)}{S}$$

$$SpeedP \approx -\Gamma \frac{d_1}{S}$$

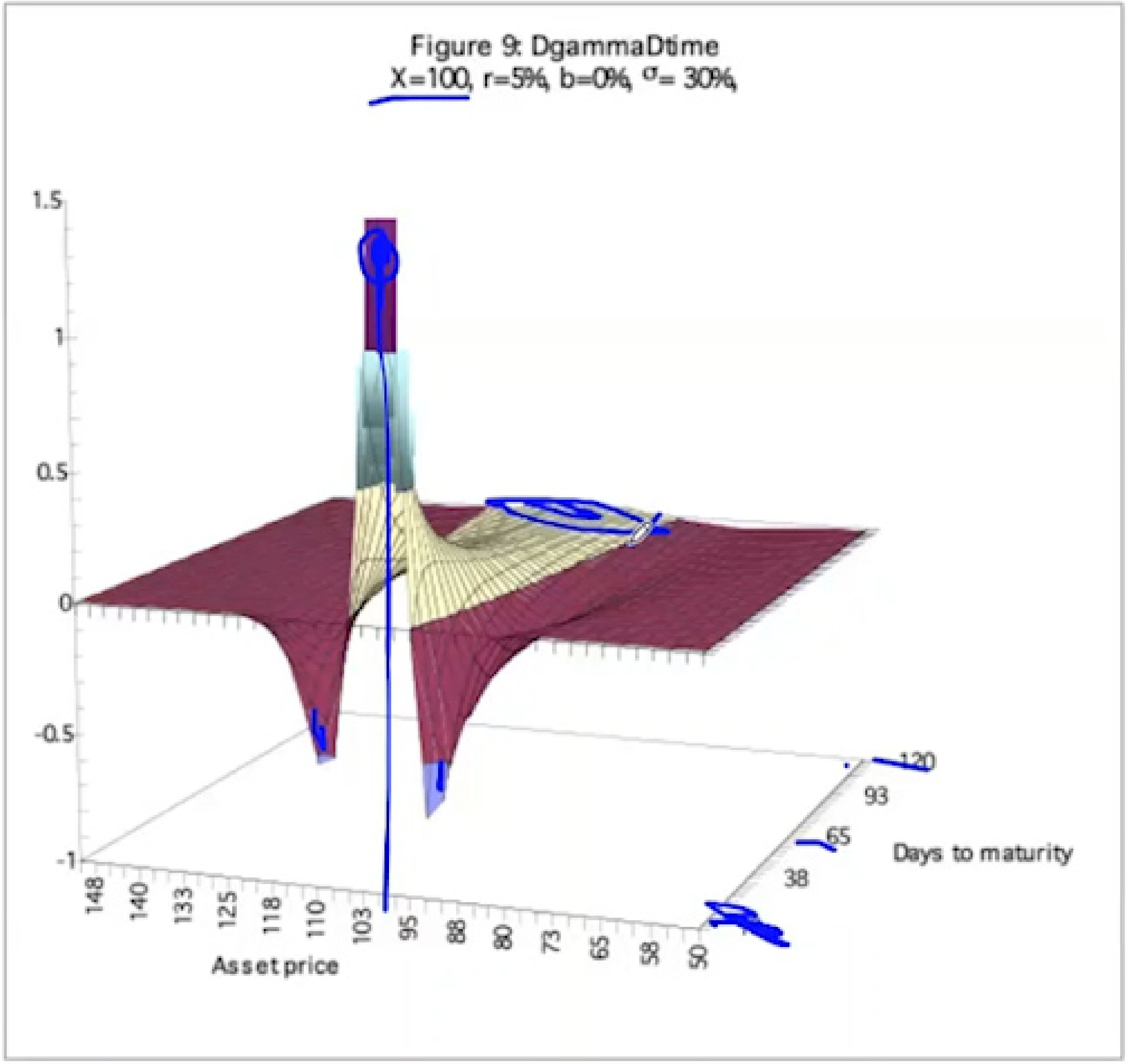


Speed is used by Fouque, Papanicolaou, and Sircar (2000) as part of stochastic vol model

DgammaDtime

$$\frac{\partial \Gamma}{\partial T} = \Gamma\left(r - b + \frac{bd_1}{\sigma\sqrt{T}} + \frac{1 - d_1 d_2}{2T} \right)$$

$$\frac{\partial \Gamma_P}{\partial T} = \Gamma_P\left(r - b + \frac{bd_1}{\sigma\sqrt{T}} + \frac{1 - d_1 d_2}{2T} \right)$$



Numerical Greeks

- More robust (?)
- Model independent
- Faster to implement (?)

$-\Delta S$

$S = 100 + \Delta S$

$\Delta S = \cancel{4}\ 0.01 \cdot S$

Two-sided finite difference

$$\Delta_C \approx \frac{c(S+\Delta S, X, T, r, b, \sigma) - c(S-\Delta S, X, T, r, b, \sigma)}{\boxed{2\Delta S}}$$

Backward derivative

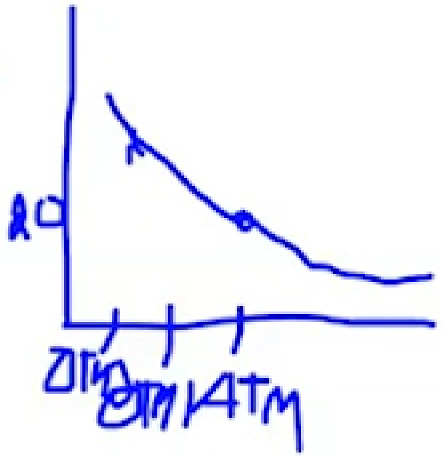
$\Delta = \frac{1}{365.25}$

$$\Theta \approx \frac{c(S, X, T, r, b, \sigma) - c(S, X, T-\Delta T, r, b, \sigma)}{\Delta T}$$

Numerical Greeks

14% 22%

$$\Delta_C \approx \frac{c(S+\Delta S, X, T, r, b, \sigma_1) - c(S-\Delta S, X, T, r, b, \sigma_2)}{2\Delta S}$$



$$\Theta \approx \frac{c(S, X, T, r, b, \sigma_1) - c(S, X, T-\Delta T, r, b, \sigma_2)}{\Delta T}$$

Gamma and other second derivatives, central finite difference

$$\Gamma \approx \frac{c(S+\Delta S, \ldots) - 2c(S, \ldots) + c(S-\Delta S, \ldots)}{\Delta S^2}$$

Speed and other third order derivatives, central finite difference

$$Speed \approx \frac{1}{\Delta S^3}\left[c(S+2\Delta S, \ldots) - 3c(S+\Delta S, \ldots) + 3c(S, \ldots) - c(S-\Delta S, \ldots)\right)$$

Numerical Greeks

What about mixed derivatives? For example DdeltaDvol and Charm

$$DdeltaDvol \approx \frac{1}{4\Delta S \Delta\sigma}\left[c(S+\Delta S,\ldots,\sigma+\Delta\sigma) - c(S+\Delta S,\ldots,\sigma-\Delta\sigma)\right.$$
$$\left. - c(S-\Delta S,\ldots,\sigma+\Delta\sigma) + c(S-\Delta S,\ldots,\sigma-\Delta\sigma)\right]$$

Vega $\quad \dfrac{\partial c}{\partial \sigma} = Se^{(b-r)T} n(d_1)\sqrt{T}$

Vega local max $b = 0$

$$S = Xe^{(-b+\sigma^2/2)T}$$

$$X = Se^{(b+\sigma^2/2)T}$$

Global maximum

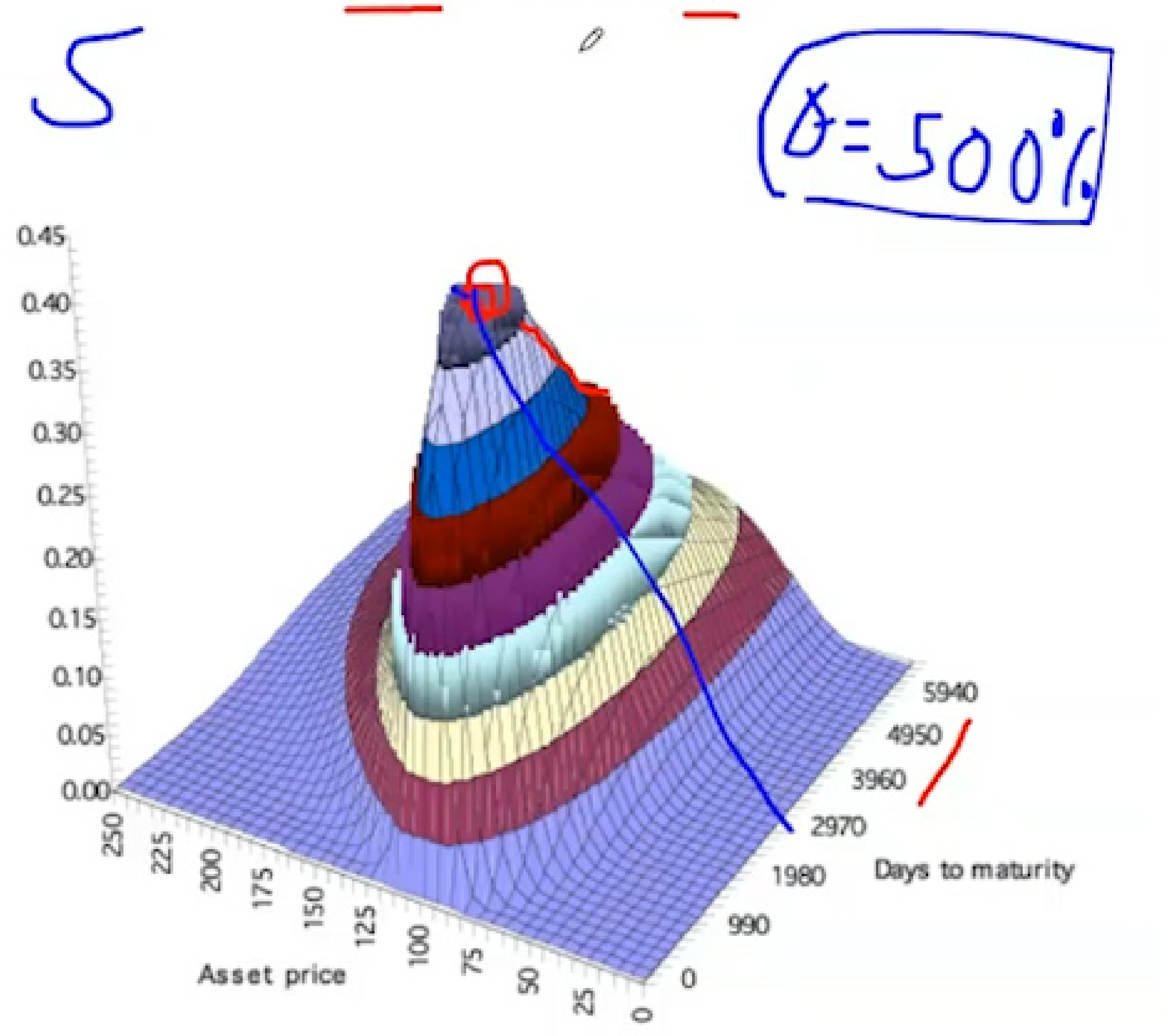
$$T_V = \frac{1}{2r}$$

$$S_V = Xe^{(-b+\sigma^2/2)T_V}$$

$$= Xe^{\frac{-b+\sigma^2/2}{2r}}$$

$$Vega(S_V, T_V) = \frac{X}{2\sqrt{re\pi}}$$

Figure 1. Vega
X=100, r=15%, b=0%, σ= 12%

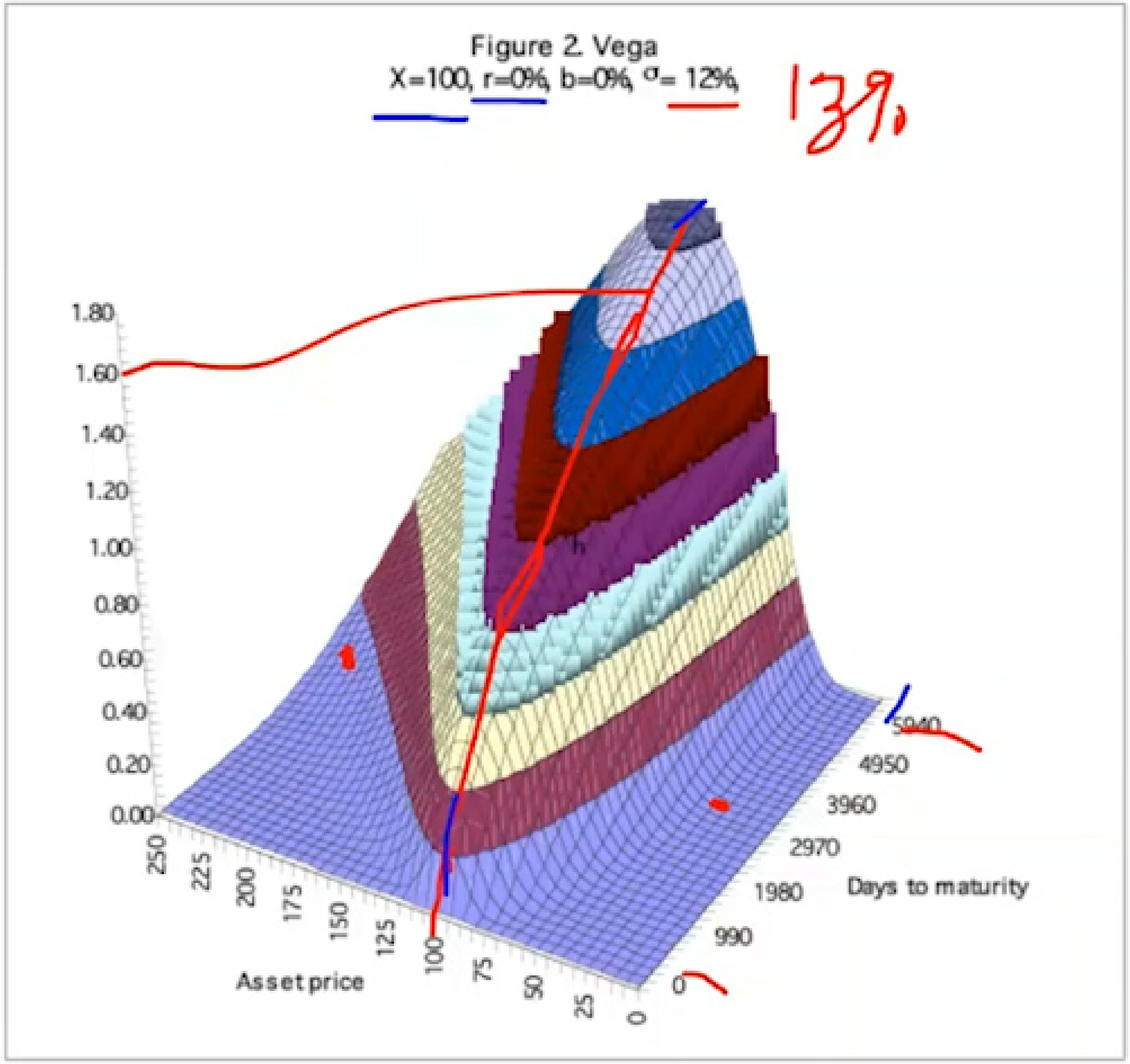


$F = X$

$C \approx p \approx F \cdot 0.4\, \sigma \cdot \sqrt{T}\, e^{-rT}$

Why the Vega top?

Discounting at some point will dominate over volatility effect (Vega).



Vega-symmetry

Put-call symmetry Bates(1991) and Carr and Bowie (1994):

$$c(S,X,T,r,b,\sigma)=\frac{X}{Se^{bT}}\,p\left(S,\frac{(Se^{bT})^2}{X},T,r,b,\sigma\right)$$

Vega-symmetry

$$Vega(S,X,T,r,b,\sigma)=\frac{X}{Se^{bT}}\,Vega\left(S,\frac{(Se^{bT})^2}{X},T,r,b,\sigma\right)$$

Also gives gamma and cost-of-carry symmetry

Probability "Greeks"

Risk neutral probability of ending up in-the-money

$$\zeta_C = N(d_2) > 0, \quad \zeta_P = N(-d_2) > 0$$

20% 30%

Strike-delta

$$\frac{\partial c}{\partial X} = -e^{-rT} N(d_2), \quad \frac{\partial p}{\partial X} = e^{-rT} N(-d_2)$$

Probability mirror strikes

$$X_P = \frac{S^2}{X_C} e^{(2b-\sigma^2)T} \quad X_C = \frac{S^2}{X_P} e^{(2b-\sigma^2)T}$$

Probability neutral straddle

$$X_C = X_P = S e^{(b-\sigma^2/2)T}$$

Probability "Greeks"

Strikes from probability

$$X_C = S\exp[-N^{-1}(p_i)\sigma\sqrt{T} + (b-\sigma^2/2)T]$$
$$X_P = S\exp[N^{-1}(-p_i)\sigma\sqrt{T} + (b-\sigma^2/2)T]$$

Risk neutral probability density

$$\underline{RND} = \boxed{\frac{\partial^2 c}{\partial X^2}} = \frac{\partial^2 p}{\partial X^2} = \frac{n(d_2)e^{-rT}}{X\sigma\sqrt{T}}$$

Probability neutral straddle

$$X_C = X_P = Se^{(b-\sigma^2/2)T}$$

Risk neutral probability density

$$RND = \boxed{\frac{\partial^2 c}{\partial X^2}} = \frac{\partial^2 p}{\partial X^2} = \frac{n(d_2)e^{-rT}}{X\sigma\sqrt{T}}$$

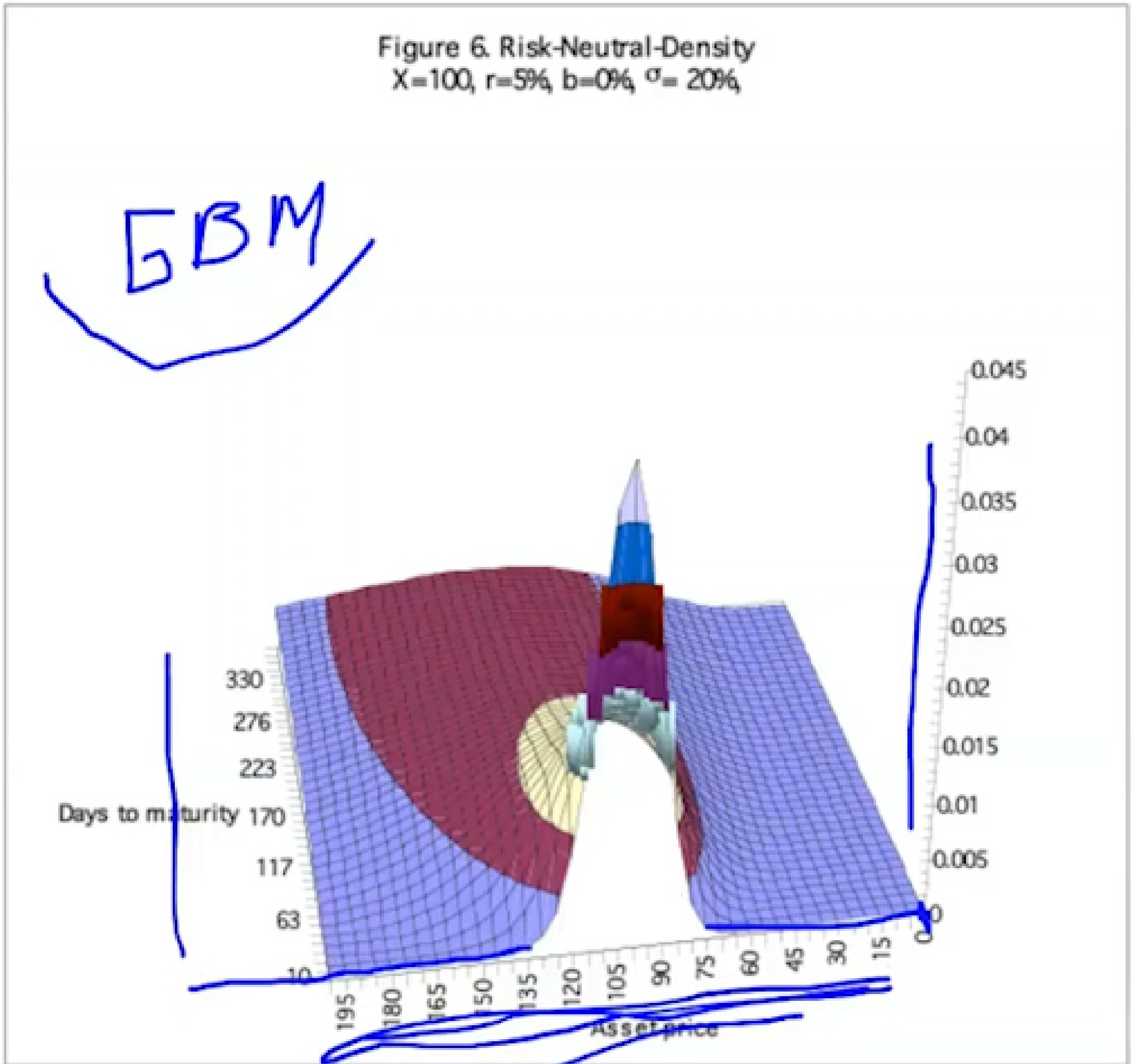
Breeden and Litzenberger (1978)

J.R 94

D.K 94

B.D. 94

GBM



Probability "Greeks"

Risk neutral probability of ever being in-the-money

$$p_C = (X/S)^{\mu+\lambda} N(-z) + (X/S)^{\mu-\lambda} N(-z + 2\lambda\sigma\sqrt{T})$$

$$p_p = (X/S)^{\mu+\lambda} N(z) + (X/S)^{\mu-\lambda} N(z - 2\lambda\sigma\sqrt{T})$$

where

$$z = \frac{\ln(X/S)}{\sigma\sqrt{T}} + \lambda\sigma\sqrt{T}, \quad \mu = \frac{b - \sigma^2/2}{\sigma^2}, \quad \lambda = \sqrt{\mu^2 + \frac{2r}{\sigma^2}}$$

‘Advanced’ Volatility Modeling in Complete Markets

In this lecture. . .

- the relationship between implied volatility and actual volatility in a deterministic world
- the local volatility surface
- the difference between ‘random’ and ‘uncertain’
- non-linear pricing equations
- optimal static hedging with traded options
- how non-linear equations make a mockery of calibration

By the end of this lecture you will

- understand the mathematics behind deterministic volatility
- be able to use uncertainty, instead of randomness, for modeling
- be able to optimally statically hedge a portfolio

Introduction

In this lecture we will see two extremes to modeling volatility.

- From deterministic to hardly modeled at all.
- From one value to an infinite number of values.
- From assuming that the market is correct to assuming that it knows nothing.

The first subject is a return to deterministic volatility and volatility surfaces.

The second subject is that of uncertain parameters.

Volatility surfaces, smiles and skews revisited

Recall that we earlier looked at how implied volatility varies with strike and expiration in practice.

Volatility matrices

Here is an example of price versus strike and expiry. . .

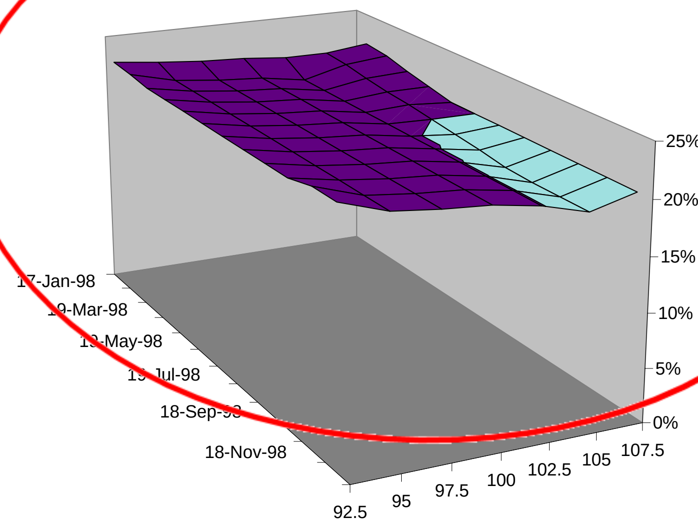
	95	**100**	**105**
1 month	11.80	7.24	3.50
3 months	13.33	9.23	5.76
7 months	16.02	12.13	7.97

This can be reinterpreted in terms of implied volatility versus strike and expiry. . .

	95	**100**	**105**
1 month	24.1%	22.9%	21.2%
3 months	22.7%	21.5%	20.5%
7 months	21.8%	20.5%	19.4%

Volatility surfaces

We can show implied volatility against both maturity and strike in a three-dimensional plot.



This implied volatility surface represents the constant value of volatility that gives each traded option a theoretical value equal to the market value.

There are various ways of interpreting this.

One of the key points is to ask whether we believe that the option market is right or wrong in its assessment of implied volatility.

In other words, is implied volatility a good forecast of actual volatility?

For this half of the lecture we assume that

- the market is correct, implied volatility is a perfect predictor of future actual volatility

- actual volatility is a deterministic function of stock price and time

The last assumption is the simplest that is potentially consistent with varying implied volatility.

Notation

$+\sigma S \, dX$

Actual volatility:

$$\sigma(S,t)$$

Implied volatility:

$$\sigma_\text{i}(E,T)$$

$$\sigma(S,t) \not\Rightarrow \sigma_i(E,T)$$

$$\Leftarrow \quad \checkmark$$

Our task is to find $\sigma(S,t)$ given market prices of options i.e. knowing $\sigma_{\text{i}}(E,T)$.

This is working backwards!

Normally we say that we know the parameters (here the 'actual' volatility) and then calculate the solution (here the 'implied' volatility).

$\sigma(S,t)$ is the **local volatility surface**.

Backing out the local volatility surface from European call option prices

Market prices of traded vanilla options are never consistent with the constant volatility assumed by Black & Scholes.

To match the theoretical prices of traded options to their market prices always requires a volatility structure that is a function of both the asset price, S, and time, t i.e. $\sigma(S,t)$.

To back out the local volatility surface from the prices of market traded instruments we are going to assume that we have a distribution of European call prices of all strikes and maturities. These prices will be denoted by $V(E,T)$.

This notation is vastly different from before.

Previously, we had the option value as a function of the underlying and time.

Now the asset and time are fixed at S^* and t^*, today's values.

We will use the dependence of the market prices on strike and expiry to calculate the volatility structure.

Assume that the risk-neutral random walk for S is

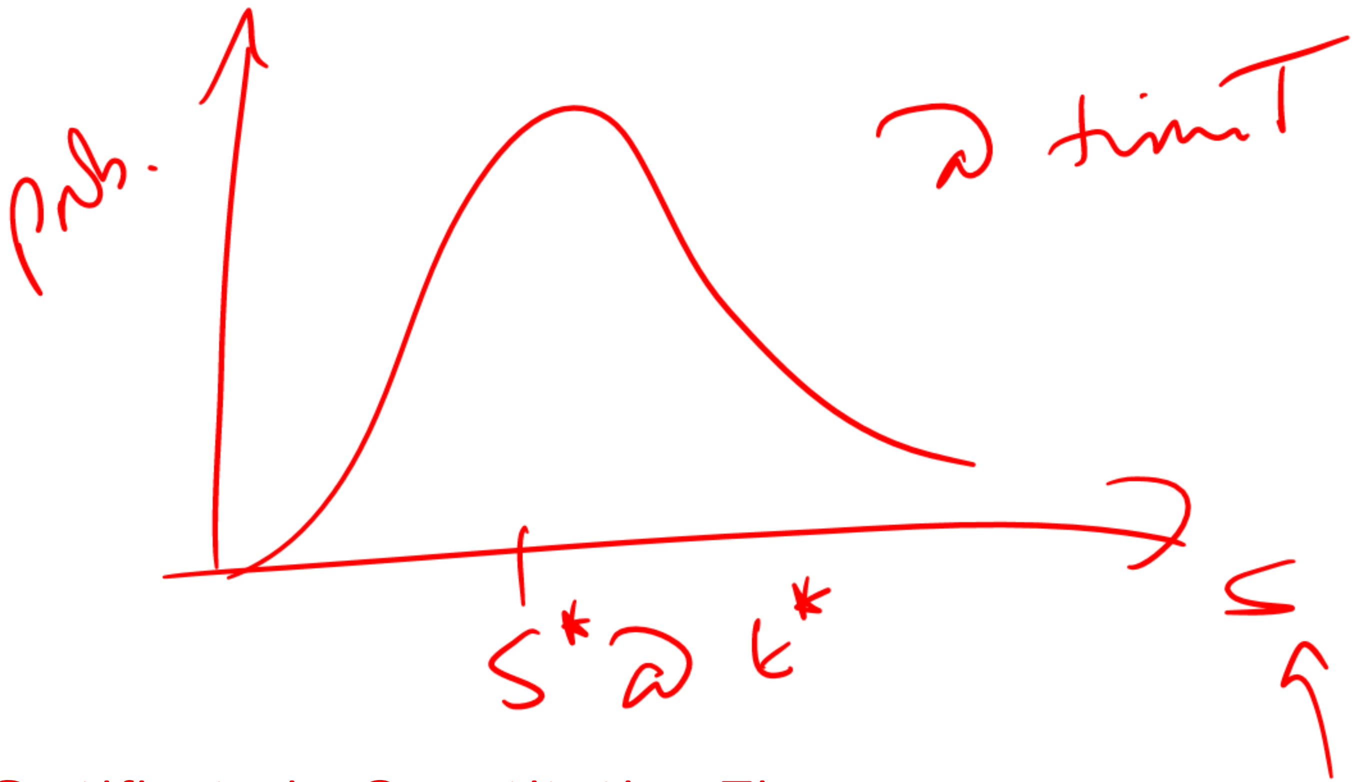
$$dS = rS\,dt + \sigma(S,t)S\,dX.$$

This is our usual one-factor model for which all the building blocks of delta hedging and arbitrage-free pricing hold.

The only novelty is that the volatility is dependent on the level of the asset and time.

In the following, we are going to rely heavily on the transition probability density function $p(S^*, t^*; S, T)$ for the risk-neutral random walk.

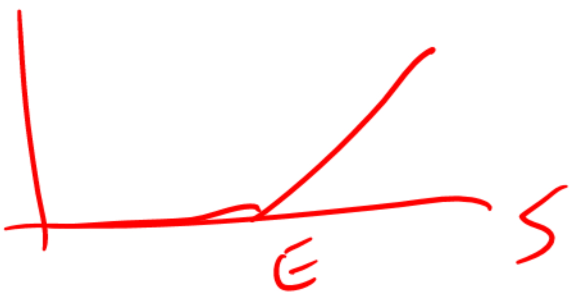
Note that the backward variables are fixed at today's values and the forward time variable is T.

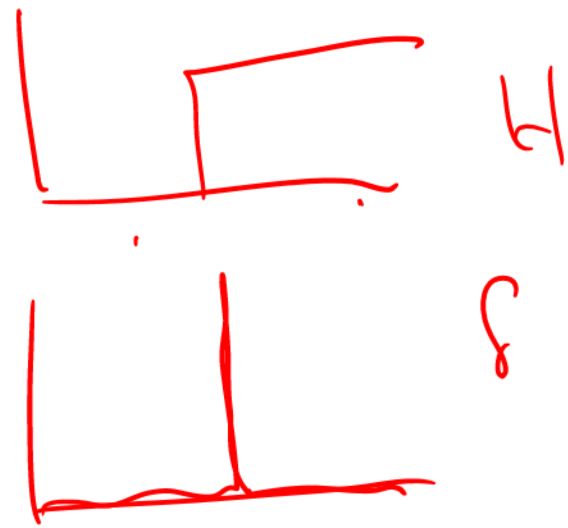


Remember that one of the interpretations of the Binomial/Black–Scholes model is that value of an option is the present value of the expected payoff.

So for a call option we can write

$$V(E,T)=e^{-r(T-t^*)}\int_0^\infty \max(S-E,0)\,p(S^*,t^*;S,T)\,dS.$$





$$V = e^{-r(T-t^*)} \int_E^\infty (S-E)\, p(S^*,t^*;S,T)\, dS. \quad (1)$$

int. egn

We are *very* lucky that the payoff is the maximum function so that after differentiating with respect to E we get

Leibniz

$$\frac{\partial V}{\partial E} = -e^{-r(T-t^*)} \int_E^\infty p(S^*,t^*;S,T)\, dS.$$

And after another differentiation, we arrive at

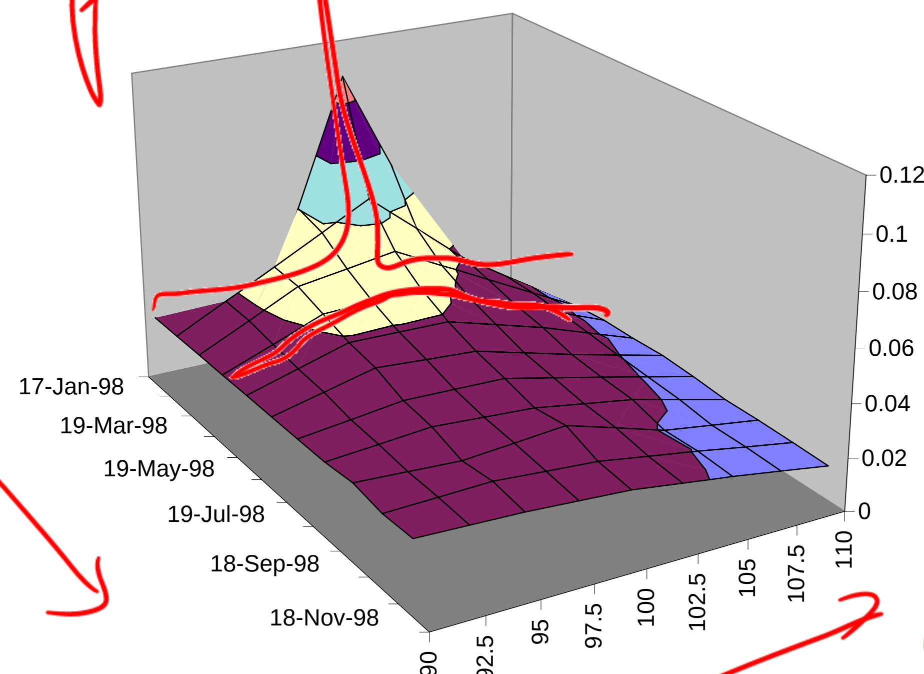
$$\frac{\partial^2 V}{\partial E^2} = e^{-r(T-t^*)} p(S^*,t^*;E,T).$$

There is therefore a surprisingly simple relationship between the second derivative of the option value with respect to strike and the risk-neutral probability density function:

$$p(S^*, t^*; E, T) = e^{r(T-t^*)} \frac{\partial^2 V}{\partial E^2}. \tag{2}$$

Before even calculating volatilities we can find the transition probability density function. In a sense, this is the market's view of the future distribution.

But it's the market view of the risk-neutral distribution and not the real one.



Our transition density function satisfies two equations, the forward and the backward.

We are going to exploit the forward equation for the transition probability density function, the Fokker–Planck equation,

$$\frac{\partial p}{\partial T} = \tfrac{1}{2}\frac{\partial^2}{\partial S^2}(\sigma^2 S^2 p) - \frac{\partial}{\partial S}(rSp). \qquad (3)$$

Here σ is our, still unknown, function of S and t. However, *in this equation* $\sigma(S,t)$ *is evaluated at* $t = T$.

$V = e^{\cdots} \int_E^{\infty} (S-E)\, p\, dS$

From (1) we have

$$\frac{\partial V}{\partial T} = -rV + e^{-r(T-t^*)} \int_E^{\infty} (S-E) \frac{\partial p}{\partial T} \, dS.$$

This can be written as

$$\frac{\partial V}{\partial T} = -rV + e^{-r(T-t^*)} \int_E^{\infty} \left(\frac{1}{2} \frac{\partial^2(\sigma^2 S^2 p)}{\partial S^2} - \frac{\partial (rSp)}{\partial S} \right) (S-E) \, dS$$

using the forward equation (3).

Integrating this by parts twice, assuming that p and its first S derivative tend to zero sufficiently fast as S goes to infinity we get

$$\frac{\partial V}{\partial T} = -rV + \frac{1}{2} e^{-r(T-t^*)} \sigma^2 E^2 p + r e^{-r(T-t^*)} \int_E^\infty S p \, dS.$$

In this expression $\sigma(S,t)$ has $S = E$ and $t = T$.

Writing

$$\int_E^\infty Sp\,dS = \int_E^\infty (S-E)p\,dS + E\int_E^\infty p\,dS$$

and collecting terms, we get

$$\frac{\partial V}{\partial T} = \tfrac{1}{2}\sigma^2 E^2 \frac{\partial^2 V}{\partial E^2} - rE\frac{\partial V}{\partial E}.$$

Bruno Dupire
Mark Rubinstein
Emanuel Derman

Rearranging this we find that

$$\sigma = \sqrt{\frac{\frac{\partial V}{\partial T} + rE\frac{\partial V}{\partial E}}{\frac{1}{2}E^2\frac{\partial^2 V}{\partial E^2}}}.$$

This is the answer we have been seeking.

There is one subtle point... In this σ is a function of E and T. We must relabel the variables to get $\sigma(S,t)$.

This calculation of the volatility surface from option prices worked because of the particular form of the payoff, the call payoff, which allowed us to derive the very simple relationship between derivatives of the option price and the transition probability density function.

When there is a constant and continuous dividend yield on the underlying the relationship between call prices and the local volatility is

$$\sigma = \sqrt{\frac{\frac{\partial V}{\partial T} + (r - D)E\frac{\partial V}{\partial E} + DV}{\frac{1}{2}E^2\frac{\partial^2 V}{\partial E^2}}}. \tag{4}$$

One of the problems with this expression concerns data far in or far out of the money.

Unless we are close to at the money both the numerator and denominator of (4) are small, leading to inaccuracies when we divide one small number by another.

One way of avoiding this is to relate the local volatility surface to the implied volatility surface.

In the same way that we found a relationship between the local volatility and the implied volatility in the purely time-dependent case we can find a relationship in the general case of asset- and time-dependent local volatility.

$V = S N(d_1) - \cdots$

The result is

$d_1 = \cdots \sigma_i \cdots$

$$\sigma(E,T)^2 =$$

$$\frac{\sigma_i^2 + 2(T-t^*)\sigma_i \frac{\partial \sigma_i}{\partial T} + 2(r-D)E(T-t^*)\sigma_i \frac{\partial \sigma_i}{\partial E}}{\left(1 + E d_1 \sqrt{T-t^*}\frac{\partial \sigma_i}{\partial E}\right)^2 + E^2 (T-t^*)\sigma_i \left(\frac{\partial^2 \sigma_i}{\partial E^2} - d_1 \left(\frac{\partial \sigma_i}{\partial E}\right)^2 \sqrt{T-t^*}\right)} \tag{5}$$

where

$$d_1 = \frac{\log(S^*/E) + (r - D + \frac{1}{2}\sigma_i^2)(T-t^*)}{\sigma_i \sqrt{T-t^*}}.$$

Again, we must relabel the variables to get $\sigma(S,t)$.

In terms of the implied volatility the implied risk-neutral probability density function is

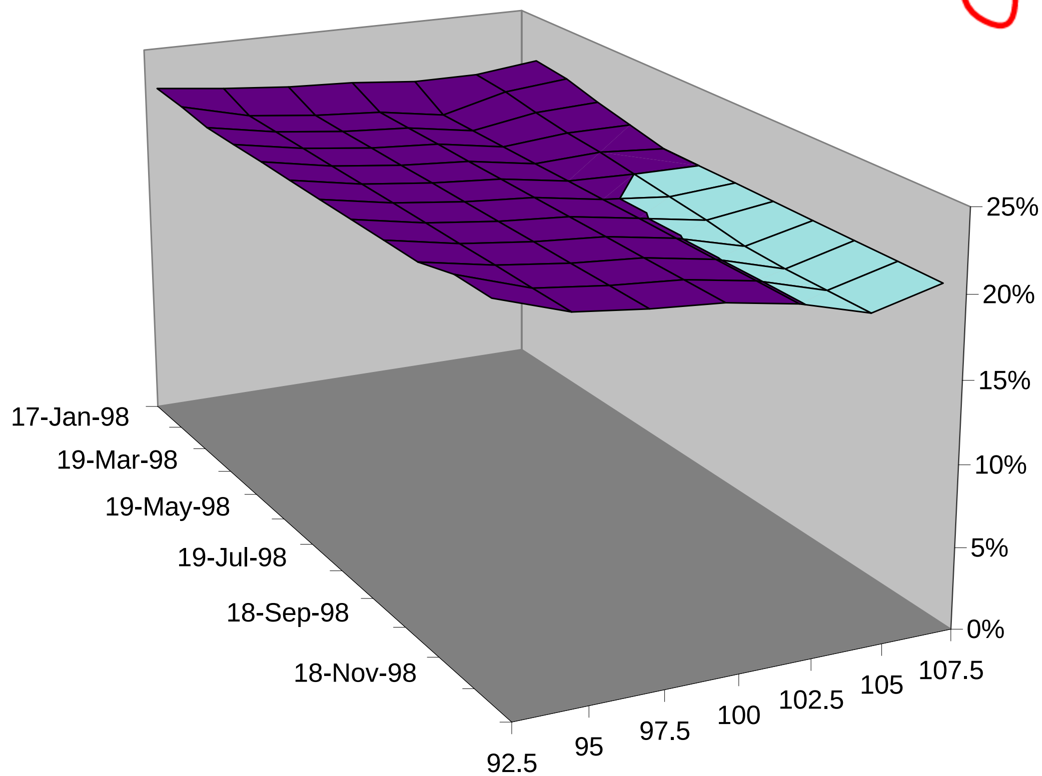
$$p(S^*, t^*; E, T) = \frac{1}{E\sigma_{\mathsf{i}}\sqrt{2\pi(T-t^*)}} e^{-\frac{1}{2}d_2^2}$$

$$\left(\left(1 + E d_1 \sqrt{T-t^*} \frac{\partial \sigma_{\mathsf{i}}}{\partial E} \right)^2 + E^2 (T-t^*) \sigma_{\mathsf{i}} \left(\frac{\partial^2 \sigma_{\mathsf{i}}}{\partial E^2} - d_1 \left(\frac{\partial \sigma_{\mathsf{i}}}{\partial E} \right)^2 \sqrt{T-t^*} \right) \right).$$

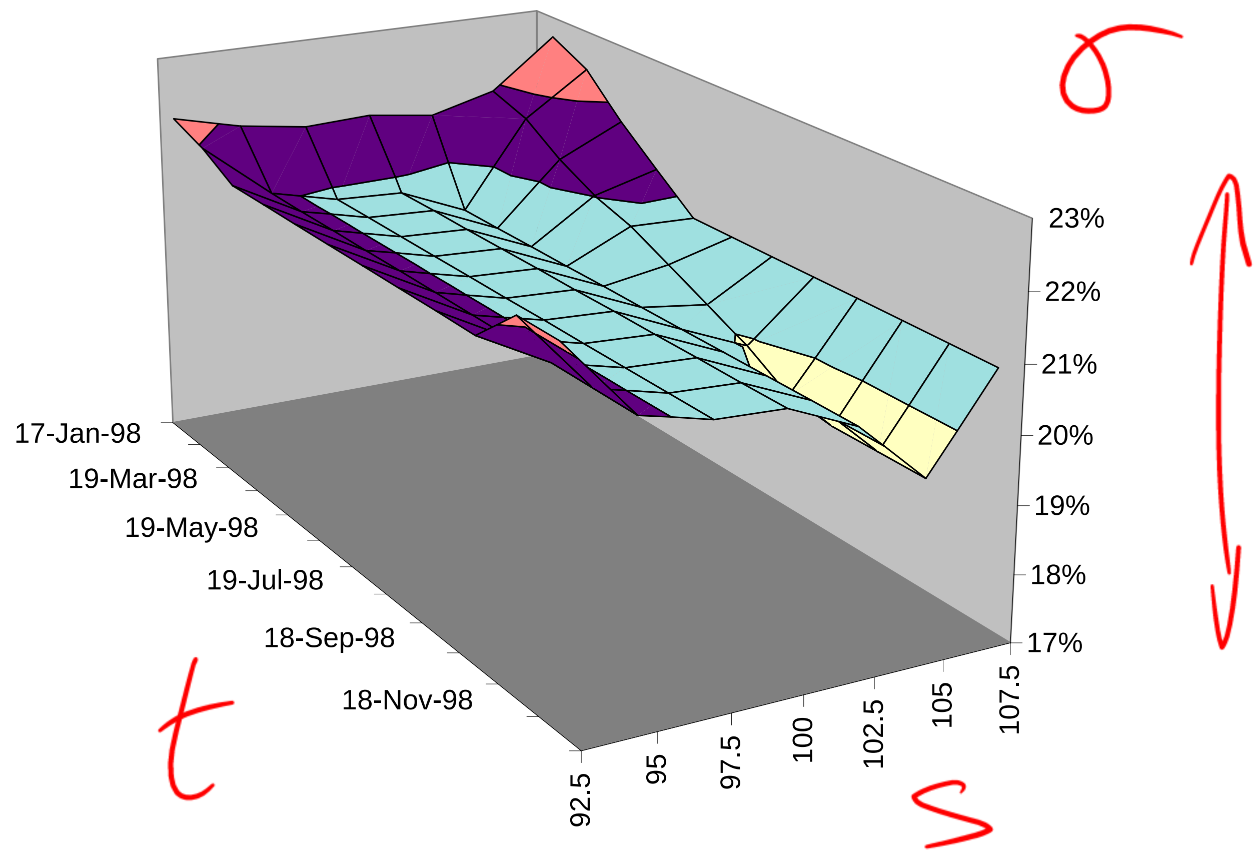
And again, we must relabel the variables to get $p(S^*, t^*; E, T)$ in terms of its natural variables S and t instead of E and T.

One of the advantages of writing the local volatility and probability density function in terms of the implied volatility surface is that if you put in a flat implied volatility surface you get out a flat local surface and a lognormal distribution.

Recall that we started with this as the *implied* volatility surface. . .



Now we have the *local* volatility surface. . .



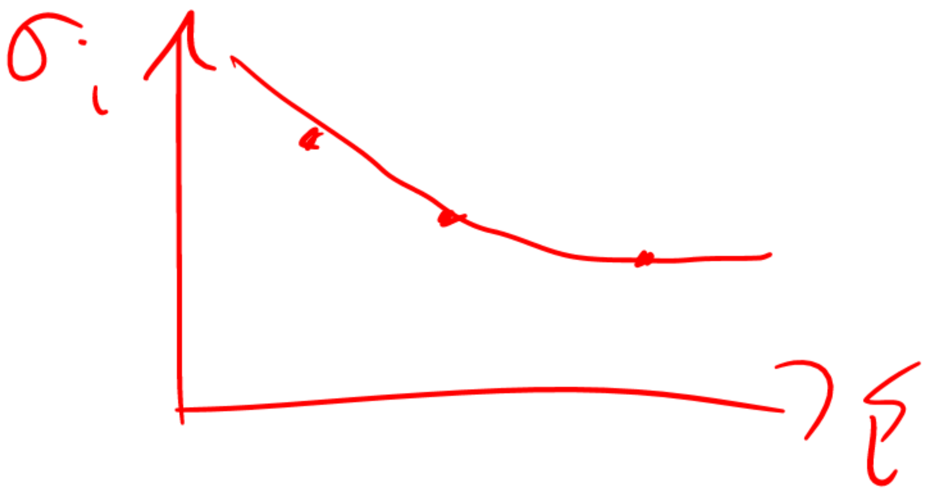
Local volatility surface calculated from European call prices.

Much less well behaved!

In practice there only exist a finite, discretely-spaced set of call prices.

To deduce a local volatility surface from this data requires some interpolation and extrapolation.

This can be done in a number of ways and there is no correct way.



One of the problems with these approaches is that the final result depends sensitively on the form of the interpolation.

The problem is actually 'ill-posed,' meaning that a small change in the input can lead to a large change in the output.

There are many ways to get around this ill-posedness, coming under the general heading of 'regularization.'

What do we do with the local volatility surface?

Once we've found $\sigma(S, t)$ we can now use it to find the values of other (non-vanilla) products!

This can be done, for example, by Monte Carlo simulation or by finite difference solution of the relevant PDE.

You can then say that the value of other derivatives "is consistent with the traded options."

Pros and cons of the deterministic volatility model

Pros:

- Famous people invented it
- It is easy to fool people that it is a good model

Cons:

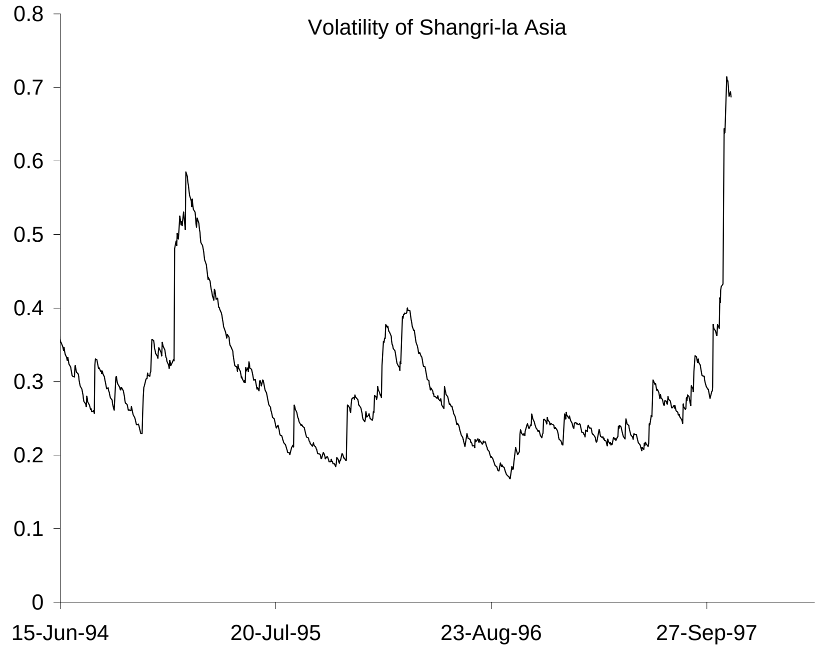
- The assumption $\sigma(S,t)$ is not sensible
- Even if it were sensible, the problem is an inverse problem with serious numerical instability difficulties (the lesson of CSI Miami)
- The results are unstable (the lesson of the fortune teller), proving that the model does not capture correct dynamics

Uncertain Parameters

In the Black–Scholes model which variables and parameters are easily measurable?

- Asset price
- Time to expiry
- Risk-free interest rate (to some extent)
- Dividends (to some extent)

- Volatility: Volatility is certainly not constant as assumed in the simple Black–Scholes formulæ. Here is an example of a volatility time series.



A typical time series for historical volatility; an implied volatility time series would look similar.

How far can we get by not even attempting to ‘model’ volatility?!

What is ‘uncertainty’?

‘Uncertainty’ is when you do not have either a deterministic or probabilistic description.

You have neither a formula nor a probability density function.

You say what is *not* allowed to happen, with no description of what *is*!

We will address the problem of how to value options when parameter values are *uncertain*.

- We assume that all we know about the parameters is that *they lie within specified ranges*.

- We do not find a *single* value for an option, instead we find that the option's value can also lie within a range: there is no such thing as *the* value.

There are many possible values any of which *might* turn out to be correct.

We will see that this problem is non linear, and thus an option valued in isolation has a different range of values from an option valued as part of a portfolio.

If we put other options into the portfolio this will change the value of the original portfolio.

This leads to the idea of incorporating traded options into an OTC portfolio in such a way as to maximize its value. This is called optimal static hedging.

Best and worst cases

The first step in valuing options with uncertain parameters is to acknowledge that we can do no better than give ranges for the future values of the parameters.

For volatility, for example, this range may be the range of past historical volatility, or implied volatilities, or encompass both of these. Then again, it may just be an educated guess.

The range we choose represents our estimate of the upper and lower bounds for the parameter value for the life of the option or portfolio in question. These ranges for parameters lead to ranges for the option's value.

Thus it is natural to think in terms of a lowest and highest possible option value; if you are long the option, then we can also call the lowest value the *worst* value and the highest the *best*.

Uncertain volatility: the model of Avellaneda, Levy & Parás and Lyons (1995)

Suppose that the volatility lies within the band

$$\sigma^{-} < \sigma < \sigma^{+}.$$

Follow the Black–Scholes hedging and no-arbitrage argument as far as we can and see where it leads us.

Construct a portfolio of one option, with value $V(S,t)$, and hedge it with $-\Delta$ of the underlying asset.

The value of this portfolio is thus

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

We still have

$$dS = \mu S\, dt + \sigma S\, dX,$$

even though σ is unknown.

The change in the value of this portfolio is

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \left(\frac{\partial V}{\partial S} - \Delta \right) dS.$$

Even with the volatility unknown, the choice of $\Delta = \partial V / \partial S$ eliminates the risk:

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt.$$

At this stage we would normally say that if we know V then we know $d\Pi$. This is no longer the case since we do not know σ.

The argument now deviates subtly from Black–Scholes.

- We will be pessimistic: the volatility over the next time step is such that the portfolio increases by the least amount.

If we have a long position in a call option, for example, we assume that the volatility is at the lower bound σ^-; for a short call we assume that the volatility is high.

But let's do the analysis in full generality. (It's not much of a model if it will only tell us how to price single calls or puts.)

The return on this worst-case portfolio is then set equal to the risk-free rate:

$$\min_{\sigma^- < \sigma < \sigma^+} (d\Pi) = r\Pi \, dt.$$

Thus we set

$$\min_{\sigma^- < \sigma < \sigma^+} \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt = r \left(V - S\frac{\partial V}{\partial S} \right) dt.$$

Observe that the volatility term is multiplied by the option's gamma.

Therefore the value of σ that will give this its minimum value depends on the sign of the gamma.

When the gamma is positive we choose σ to be the lowest value σ^- and when it is negative we choose σ to be its highest value σ^+.

The worst-case value V^- satisfies

$$\frac{\partial V^-}{\partial t} + \frac{1}{2}\sigma(\Gamma)^2 S^2 \frac{\partial^2 V^-}{\partial S^2} + rS\frac{\partial V^-}{\partial S} - rV^- = 0 \tag{6}$$

where $\Gamma = \frac{\partial^2 V^-}{\partial S^2}$ and

$$\sigma(\Gamma) = \begin{cases} \sigma^+ & \text{if } \Gamma < 0 \\ \sigma^- & \text{if } \Gamma > 0. \end{cases}$$

We can find the best option value V^+, and hence the range of possible values by solving

$$\frac{\partial V^+}{\partial t} + \frac{1}{2}\sigma(\Gamma)^2 S^2 \frac{\partial^2 V^+}{\partial S^2} + rS\frac{\partial V^+}{\partial S} - rV^+ = 0$$

where $\Gamma = \frac{\partial^2 V^+}{\partial S^2}$ but this time

$$\sigma(\Gamma) = \begin{cases} \sigma^+ & \text{if } \Gamma > 0 \\ \sigma^- & \text{if } \Gamma < 0. \end{cases}$$

Two observations about the best case:

1. The best case for a long position is the same as the worst case for a short position. (To see this just change a few signs in the PDE.)

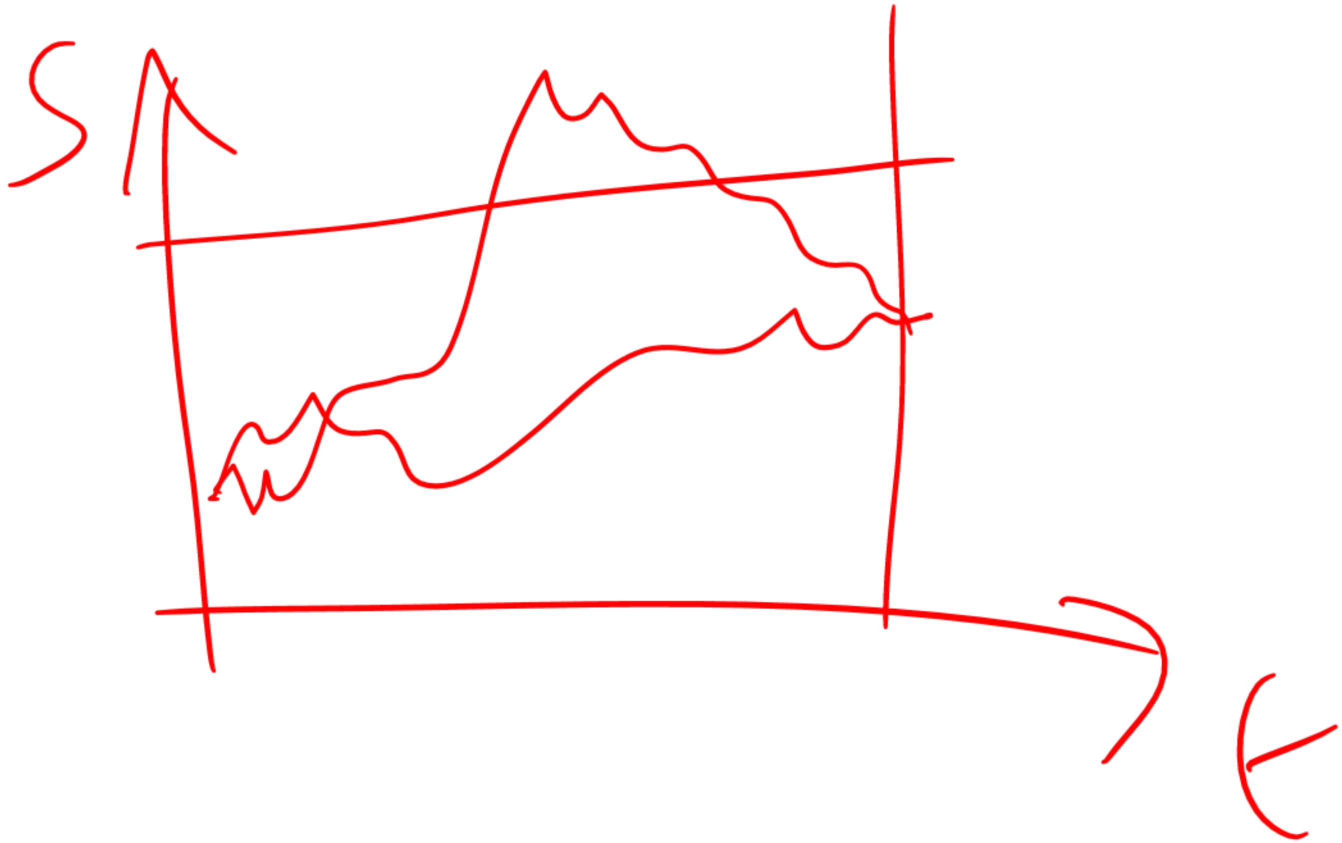
2. In practice it would be madness to go around assuming the best-case outcome for anything!

Equation (6) (Avellaneda, Levy & Parás and Lyons) is the same as the Hoggard–Whalley–Wilmott transaction cost model.

The equation must in general be solved numerically, because it is non linear.

Example: An up-and-out call

Value an up-and-out call. Volatility lies in the range 17% to 23%.

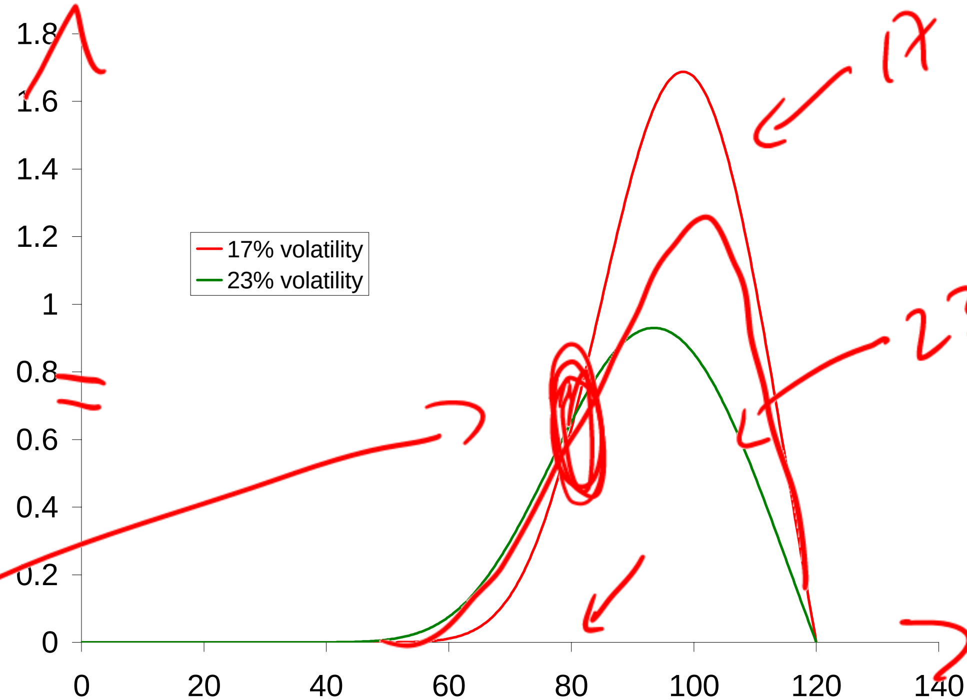


Naive approach: assume volatility is constant. . .

If we naively priced the option using first a 17% volatility and then a 23% volatility we would get two curves looking like those in the following figure.

Vega

V



17%

23%.

$\partial V / \partial \sigma \neq 0$

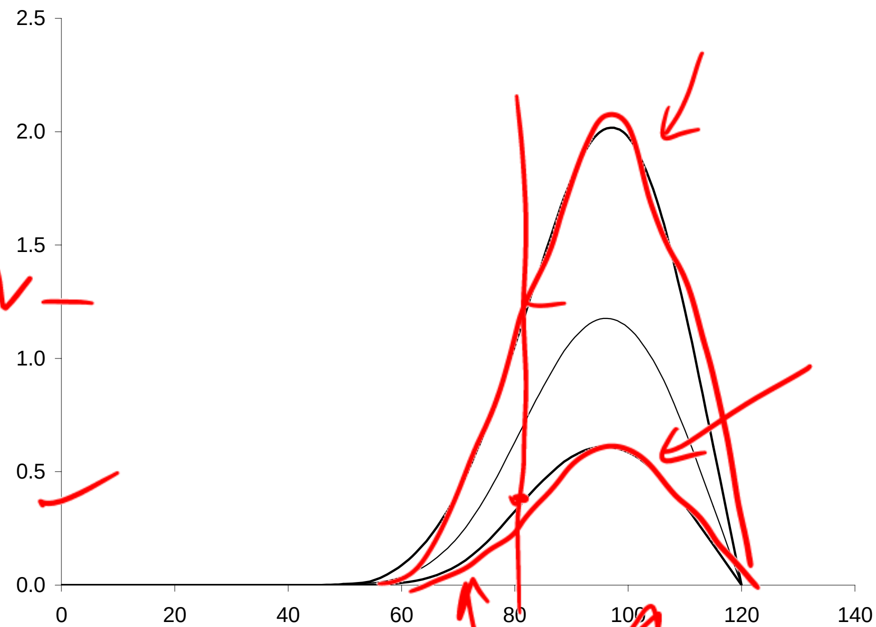
S

This figure suggests that there is a point at which the value is insensitive to the volatility. Vega is zero.

Actually, the value is very sensitive to volatility as we shall now see (and in the process also see why vega can be a poor measure of sensitivity).

In the next figure are shown the best and worst prices for an up-and-out call option *using the non-linear model*.

In the figure is a Black–Scholes value, the middle line, using a volatility of 20%. The other two bold lines give the worst-case long and short values assuming a volatility ranging over 17% to 23%.



1.2

0.3

FACTOR OF × 10

Note:

- We must solve the non-linear equation numerically because the gamma for this contract is not single signed.

- The problem is genuinely non linear, and we cannot just substitute each of 17% and 23% into a Black–Scholes formula.

- Observe that the best/worst of these two curves (the 'envelope') is not the same as the best/worst of the previous figure.

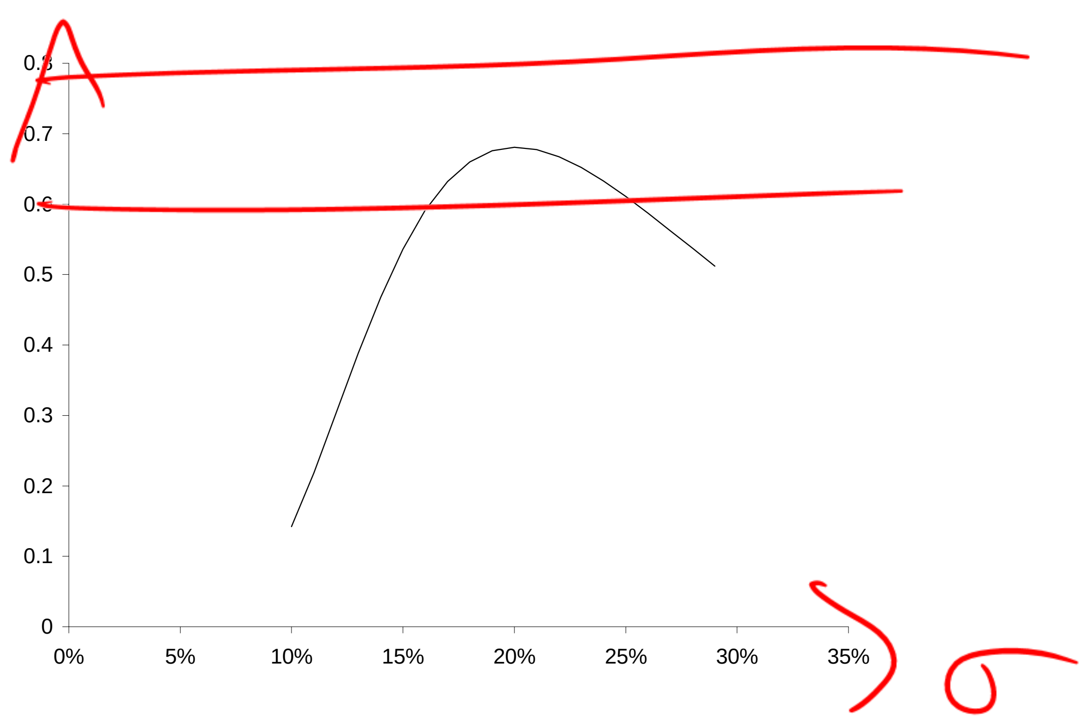
- Where is the price most/least sensitive to the volatility?

It can be extremely dangerous to calculate a contract's vega when the contract has a gamma that changes sign.

Continuing with this barrier option example, let us look at implied volatilities.

In the figure below is shown the Black–Scholes value of an up-and-out call option as a function of the volatility.

V



σ

This contract has a gamma that changes sign, and a price that is not monotonic in the volatility.

This figure shows that there is a maximum option value of 0.68 when the volatility is about 20%.

Suppose that the market is pricing this contract at 0.55. From the figure we can see that there are *two* volatilities that correspond to this market price. Which, if either, is correct?

The question is probably meaningless because of the non-single-signed gamma of this contract.

Take this example further. What if the market price is 0.72? This value cannot be reached by any single volatility.

Does this mean that there are arbitrage opportunities? Not necessarily.

This could be due to the market pricing with a non-constant volatility, either with a volatility surface, stochastic volatility or a volatility range.

As we have seen from the best/worst prices for this contract, the uncertainty in the option value may be large enough to cover the market price of the option, and there may be no guaranteed arbitrage at all.

Nonlinearity

The uncertain parameter partial differential equations that we have derived are non linear.

Because of this nonlinearity, we must distinguish between long and short positions.

For example for a long call we have

$$V^{-}(S,T) = \max(S-E,0)$$

and for a short call

$$V^{-}(S,T) = -\max(S-E,0).$$

Because Equation (6) is non linear the value of a portfolio of options is not necessarily the same as the sum of the values of the individual components.

Long and short positions have different values.

For example, a long call position has a lower value than a short call. In both cases we are being pessimistic: if we own the call we assume that it has a low value, if we are short the call and thus may have to pay out at expiry, we assume that the value of the option to its holder is higher.

Note that here we mean a long (or short) position valued *in isolation*.

Obviously, if we hold one of each simultaneously then they will cancel each other regardless of the behavior of any parameters.

This is a very important point to understand: *the value of a contract depends on what else is in the portfolio*.

A problem with the model . . . which will be sorted out

Unfortunately, the model as it stands predicts very wide spreads on options.

For example, suppose that we have a European call, strike price \$100, today's asset price is \$100, there are six months to expiry, no dividends but a spot interest rate that we expect to lie between 5% and 6% and a volatility between 20% and 30%.

We can calculate the values for long and short calls assuming these ranges for the parameters directly from the Black–Scholes formulæ. *This is because the gamma and the portfolio value are single-signed for a call.*

A long call position is worth **\$6.89** (the Black–Scholes value using a volatility of 20% and an interest rate of 5%) and a short call is worth **\$9.88** (the Black–Scholes value using a volatility of 30% and an interest rate of 6%).

This spread is much larger than that in the market. The market prices may, for example, be based on an interest rate of 5.5% with a volatility between 24% and 26%.

Unless the model can produce narrower spreads the model will be useless in practice.

The spreads *can* be tightened by ‘static hedging.’

This means the purchase and sale of traded option contracts so as to improve the marginal value of our original position.

This only works because we have a *non-linear* governing equation: the price of a contract depends on what else is in the portfolio.

This static hedge can be optimized so as to give the original contract its best value, we can squeeze even more value out of our contract with the best hedge.

Static Hedging

Delta hedging is a wonderful concept. It leads to preference-free pricing (risk neutrality) and a risk-elimination strategy that can be used in practice. There are quite a few problems, though, on both the practical and the theoretical side.

- In practice, hedging must be done at discrete times and is costly. Sometimes one has to buy or sell a prohibitively large number of the underlying in order to follow the theory.

- On the theoretical side, we have to accept that the model for the underlying is not perfect, at the very least we do not know parameter values accurately.

Many of these problems can be reduced or eliminated if we follow a strategy of static hedging as well as delta hedging: buy or sell more liquid contracts to reduce the cashflows in the original contract.

Static hedging: non-linear governing equation

Many pricing models are non linear:

- Transaction costs: Purchase and sale of the underlying for delta hedging when there are costs leads to non-linear equations for the option value.

- Uncertain parameters: When parameters, such as volatility, dividend rate, interest rate, are permitted to lie in a range, options can be valued in a worst-case scenario.

Non-linear equations

Nonlinearity has many important consequences.

Because of the nonlinearity,

- the value of a portfolio of options is not necessarily the same as the sum of the values of the individual components

- the value of a contract depends on what else is in the portfolio

These two points are key to the importance of non-linear pricing equations: they give us a bid-offer spread on option prices, and they allow *optimal* static hedging.

But the most important point about the non-linear model we have here is that the value of a portfolio of options is at least as valuable as the sum of the individual components:

$$\text{Value}(A+B+C+D+\cdots)$$

$$\geq \text{Value}(A)+\text{Value}(B)+\text{Value}(C)+\text{Value}(D)+\cdots.$$

With these models always price a portfolio rather than option by option.

This way you will get cancelation of gamma and so less exposure to volatility, and so better prices!

For the rest of this lecture we discuss the pricing and hedging of options when the governing equation is non linear.

The ideas are applicable to any non-linear models.

We use the notation $V_{NL}(S,t)$ to mean the solution of the model in question, whichever model it may be.

Pricing with a non-linear equation

One of the interesting points about non-linear models is the prediction of a spread between long and short prices.

If the model gives different values for long and short then this is in effect a spread on option prices.

This can be seen as either a good or a bad point.

- It is good because it is realistic, spreads exist in practice.

- It only becomes bad when this spread is too large to make the model useful.

Motivation

Suppose that we want to sell an option with some payoff that does not exist as a traded contract, i.e. an exotic or OTC contract.

We want to determine how low a price can we sell it for, with the constraint that we guarantee that we will not lose money as long as our range for volatility is not breached.

Using the uncertain parameter model we will expect to get best-worst ranges that are too big. This means that we will lose the deal.

We will now see how to use traded options to decrease that bid-offer spread.

To make the example as simple to follow as possible, we are going to 'pretend' that a vanilla call is 'exotic' i.e. not traded.

This is easier to visualize, to gain intuition.

Example:

- European call
- strike price \$100
- today's asset price is \$100
- there are six months to expiry
- no dividends
- interest rate of 5%
- volatility lies between 20% and 30%

Long call position is worth **\$6.89**.

Short call is worth **\$9.63**.

This spread is too large and we will be unable to either buy or sell the contract!

Static hedging

Suppose that call options on this particular stock are traded with strikes of \$90 and \$110 and with six months to expiry.

Strike	Expiry	Bid	Ask	Quantity
90	180 days	14.42	14.42	?
110	180 days	4.22	4.22	?

Can we take advantage of these contracts for pricing and/or hedging our 100 call?

Suppose that the market prices the 90 and 110 calls with an implied volatility of 25%.

The market prices, i.e. the Black–Scholes prices, are therefore 14.42 and 4.22 respectively. These numbers are shown in the table.

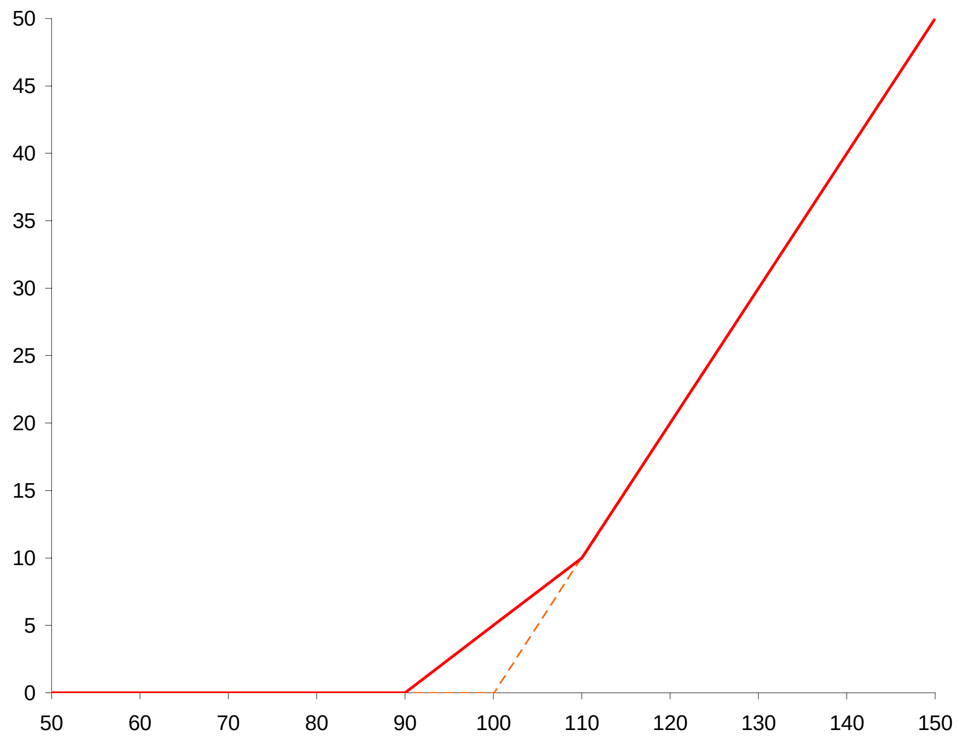
The question marks are to emphasize that we can buy or sell as many of these contracts as we want, but in a fashion which will be made clear shortly.

Shouldn't our quoted prices for the 100 call reflect the availability of contracts with which we can hedge?

Let's try hedging the 100 call with the traded options and see what happens. . .

Let's buy 0.5 of the 90 calls and 0.5 of the 110s.

This 'hedging portfolio' has payoff as shown:



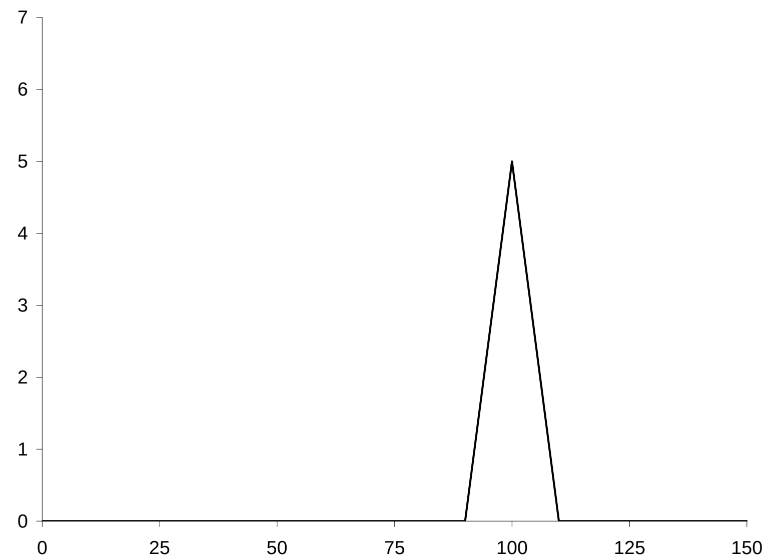
This payoff is very close to the payoff for our 100 call, so the values should be 'similar.'

And we know the value of the hedge from the market!

Hold that thought. . .

The values won't be the same, however, because of the difference in payoffs.

If we sell the 100, and ‘statically hedge’ it by buying 0.5 of the 90 and 0.5 of the 110, then we have a residual payoff as shown in the figure.



We call this a static hedge because we put it in place now and do not expect to change it. This contrasts with the delta hedge, for which we expect to hedge frequently.

The statically-hedged portfolio has a much smaller payoff than the original unhedged call.

It is this *new portfolio* that we value by solving the non-linear pricing equation, and that we must delta hedge; this last point must be emphasized, *the residual payoff must be delta hedged*.

And because the payoff is much smaller, the bid-offer or worst/best spread will be smaller.

To value the residual payoff in our uncertain parameter framework we solve the non-linear pricing equation with final condition

$$V_{NL}(S,T) = -\max(S-100,0)+$$

$$\frac{1}{2}(\max(S-90,0)+\max(S-110,0)).$$

Let us see what effect this has on the price at which we would *sell* the 100 call.

First of all, observe that we have paid

$$0.5 \times \$14.42 + 0.5 \times \$4.22 = \$9.32$$

for the static hedge, 0.5 of each of the 90 and 110 calls.

Now, solve the equation using the residual payoff as the final condition. The solution gives a value for the residual contract today of \$0.61.

The net value of the call is therefore

$$\$9.32 - \$0.61 = \$8.71.$$

To determine how much we should pay to *buy* the 100 call we take as our starting point the sale of 0.5 of each of the 90 and 110 calls.

This nets us \$9.32.

But now we solve the equation using the *negative* of the previous residual payoff as the final condition; the effect of the nonlinearity is different from the previous case because $\sigma(\Gamma)$ takes different values in different places in S, t-space.

We get a value of \$1.65 for the hedged position.

Thus we find that we would pay

$$\$9.32 - \$1.65 = \$7.67.$$

Note how the use of a very simple static hedge has reduced the spread from \$6.89–9.63 to \$7.67–8.71.

This is a substantial improvement, as it represents a volatility range of 23–27%; while our initial estimate for the volatility range was 20–30%

The reason for the smaller spread is that the residual portfolio has a smaller absolute spread, only \$0.61–1.65 and this is because it has a much smaller payoff than the unhedged 100 call.

In the above example we decided to hedge the 100 call using 0.5 of each of the 90 and 110 calls.

What prompted this choice?

There was no reason why we should not choose other numbers. Since our problem is non linear *the value of our OTC option depends on the combination of options with which we hedge*.

So, generally speaking, we expect a different OTC option value if we choose a different static hedge portfolio.

Of course, we now ask 'If we get different values for an option depending on what other contracts we hedge it with then is there a *best* static hedge?'

Optimal static hedging

Continuing with this example, what are the optimal static hedges for long and short positions?

Buying 0.5 each of the two calls to hedge the short 100 call we find a marginal value of $8.71 for the 100 call.

Slightly better, and *optimal*, is to buy 0.62 of the 90 call and 0.53 of the 110 call giving a marginal value of **$8.66**.

The optimal hedge for a long position is different. We should sell 0.67 of the 90 call and sell 0.55 of the 110 call. The marginal value of the 100 call is then **$7.76**.

$q_i =$

$q_{19} = -1$

With symbols

Here is the full definition of 'value' in this model, in words first:

$-\text{Vanilla}_{19}$

$$\text{Value}(\cancel{\text{Exotic}}) =$$

Van_{19}

$$\max_{q_i}\left(V_{NL}\left(\cancel{\text{Exotic}} + \sum_{i=1}^{N} q_i \ \text{Vanilla}_i\right) - \sum_{i=1}^{N} q_i \ \text{CostVanilla}_i\right).$$

q_i

The value of an exotic depends on what it is hedge with. . . and after it has been optimally hedged.

Let's do the maths.

$\text{Value}(\text{Van}_{19}) = \text{CostVanilla}_{19}$

Suppose that we want to find the lowest price for which we can sell a particular OTC or 'exotic' option with payoff $\Lambda(S)$.

Suppose that we can hedge our exotic with a variety of traded options, of which there are n.

These options will have payoffs (at the same date as our exotic to keep things simple for the moment) which we call

$$\Lambda_i(S).$$

At this point we can introduce bid and offer prices for the traded options: C_i^+ is the ask price of the ith option and C_i^- the bid, with $C_i^- < C_i^+$.

Now we set up our statically hedged portfolio: we will have λ_i of each option in our hedged portfolio.

The cost of setting up this static hedge is

$$\sum_i \lambda_i C_i(\lambda_i),$$

where

$$C_i(\lambda_i) = \begin{cases} C_i^+ & \text{if } \lambda_i > 0 \\ C_i^- & \text{if } \lambda_i < 0 \end{cases}.$$

If $\lambda_i > 0$ then we have a positive quantity of option i at the offer price in the market, C_i^+; if $\lambda_i < 0$ then we have a negative quantity of option i at the bid price in the market, C_i^-.

We let $V^{-}(S,t)$ be the pessimistic value of our *hedged* position.

The residual payoff for our statically hedged option is

$$V^{-}(S,T) = \Lambda(S) + \sum_i \lambda_i \Lambda_i(S).$$

Now we solve the pricing equation with this as final data, to find the *net* value of our position (today, at time $t = 0$, say) as

$$V^{-}(S(0),0) - \sum_i \lambda_i C_i(\lambda_i) = F(\lambda_1, \ldots, \lambda_n).$$

This is a mathematical representation of the type of problem we solved in our first hedging example.

Our goal now is to choose the λ_i to minimize $F(\cdots)$ if we are selling the exotic, and maximize if buying. (Thus the best hedge in the two cases will usually be different.) This is what we mean by 'optimal' static hedging.

Calibration?

What does all this mean for calibration?

Do our theoretical prices look anything like market prices?

In the trivial case where the option on which we are quoting is also traded then we would find that our quoted price was the same as the market price.

This is because we would hedge one for one and the residual payoff, which we would normally delta hedge, would be identically zero.

What does this say about calibration?

Calibration is perfect. . . by definition!

What about calibrating to both bid and offer? And liquidity?

Pros and cons of the uncertain volatility model

Pros:

- It is robust
- It values, and dynamically hedges, and statically hedges at the same level, consistently within the model
- Portfolios can be optimized
- Calibration is automatic

Cons:

- Everything has to be solved numerically, by finite-difference methods
- Cannot be used by the buy side

Summary

Please take away the following important ideas

- The most parsimonious volatility model that can be made consistent with market prices is that of deterministic volatility, $\sigma(S, t)$

- 'Uncertainty' is when you do not have either a deterministic or probabilistic description

- Volatility can be modeled as uncertain, which leads to a non-linear pricing equation and the possibility of static hedging