
STOCHASTIC METHODS FOR FINANCE

PROF. MARTINO GRASSELLI
UNIVERSITY OF PADOVA

WRITTEN BY: MATTEO BORTOLETTO

ACADEMIC YEAR 2019/20

Compiled on July 4, 2020



Contents

1	Introduction	1
1.1	Forward contract	1
1.2	Options	2
1.2.1	Long call option	3
1.2.2	Long put option	3
2	Discrete time finance	5
2.1	Static binomial model	5
2.1.1	Implementation	8
2.2	Hedging	9
2.3	Dividends	10
2.3.1	Put-call parity	11
2.4	Market completeness	12
2.5	The multi period case	12
2.6	Binomial lattice	14
2.6.1	Pricing options	14
2.6.2	Hedging	15
2.6.3	Dynamic interest rates	16
2.6.4	Binary call at the money	17
2.7	General binomial model	18
2.8	Convergence of the binomial model	19
2.8.1	The Black & Scholes formula	22
2.9	American options	24
2.9.1	Call option with no dividends and $R > 0$	24
2.9.2	General discrete time case with no dividends and $R = 0$	25
2.9.3	Possible strategies at time n	26
2.9.4	Snell envelope	27
2.10	General one period model	29
2.10.1	Absence of arbitrage	30
2.10.2	Market completeness	33
3	Continuous time finance	37
3.1	Stochastic integrals	37
3.1.1	Black & Scholes market	41
3.1.2	The Girsanov Theorem	42
3.2	Portfolio dynamics	43
3.3	Option pricing in the Black & Scholes model	45
3.4	The martingale approach	53
3.4.1	Application to the Black & Scholes model	53
3.5	Call option price in the Black & Scholes model	55
3.6	The Greeks	57

3.6.1	Delta	57
3.6.2	Gamma	59
3.6.3	Rho	60
3.6.4	Theta	60
3.6.5	Vega	60
3.6.6	Put option pricing in the Black & Scholes model	62
3.7	Portfolio immunization	63
3.8	Digital options	65
3.8.1	Pricing	65
3.8.2	Hedging	65
3.9	Black & Scholes model with continuous dividends	68
3.10	Multi-dimensional market	70
3.10.1	Absence of arbitrage	72
3.10.2	Completeness	73
3.10.3	Pricing	73
3.11	Change of numéraire	75
3.11.1	Application to the exchange option	76
3.11.2	Forward Measures	78
3.12	Black formula	85
3.12.1	Forward contracts	85
3.12.2	Futures contracts	86
3.12.3	Black formula	88
3.13	Quantity-adjusted options	89
3.14	Barrier options	94
3.14.1	Mathematical background	94
3.14.2	Barrier options: a classic approach	102
3.14.3	Barrier options: a modern approach	103
3.15	Bonds and interest rates	113
3.15.1	The crisis of 2007	116
3.16	Relations between $f(t, T)$, $p(t, T)$ and $r(t)$	118
3.17	Linear fixed income products	122
3.18	Floating rate bonds	122
3.19	Yield and duration	124
3.20	Interest rate swaps	125
3.21	Non linear contracts: caps and floors	126
3.22	Short interest rates	127
3.23	Affine term structures	129
3.24	Stochastic volatility	132
3.24.1	Review of Bessel squared processes	133
3.24.2	The pricing problem with stochastic volatility	134
3.24.3	Characteristic functions	137
3.25	The Carr-Madan approach	140
3.26	Implementation of the Heston model	142
3.27	Multi-dimensional Riccati equation	142
3.27.1	Matrix exponential	143
3.28	Multifactor stochastic volatility models	144
3.28.1	2-Heston model for one asset	144
3.28.2	Multiple assets Heston model	145
3.28.3	Wishart processes	146
3.28.4	A multi-factor Wishart volatility model	157
3.28.5	Comparison between Wishard and $A_2(3)$ -Heston model	160

3.29 Wishart affine stochastic correlation model	163
3.29.1 Pricing in a multiple asset model	165
3.29.2 Leverage in single name vanillas	167
3.29.3 Asymmetric conditional correlation effect	168
Bibliography	171

If you find any mistake or typo, please write at matteo.bortoletto.2@studenti.unipd.it.

Matteo Bortoletto, 07/03/2020

1 | Introduction

1.1 Forward contract

Let's assume that there are two counterparts interested in entering in a contract in the future. This means that at time $t = 0$ they want to fix the price K – called *strike price* – of whatever (e.g. an asset, an interest rate, an exchange rate, the price of some good) at a future time T , called *maturity* of the contract. The two counterparts are called:

- *long position*, which at time T has to buy the asset at the price K fixed at $t = 0$, regardless the market price;
- *short position*, which has to sell the asset at maturity

Lesson 1 of
10/03/2020

The price of the underlying asset at the current time t is denoted as S_t and it is known/measurable, whereas the price at time S_T is a random variable. This type of contract is called *forward contract*, since we are looking forward in the future. Because it costs nothing to enter into a forward contract, the payoff is also the trader's total gain or loss from the contract.

What can we say about the payoff of the forward contract? If at time T $S_T > K$ the long position is happy because it has the possibility to buy at K and immediately sell at S_T in the market, getting the payoff $S_T - K > 0$. Conversely, if at time T $S_T < K$ the long position is not happy because it buys at K what is worth less in the market. For the short position the situation is symmetric. There are two important considerations:

1. For the long position the loss is bounded and the gain is unbounded, whereas for the short position the loss is unbounded and the gain is bounded;
2. If K changes, for example if $K_2 > K_1$, then the long position will not be happy because it wants K to be as small as possible. So, it will not agree to enter in the contract.

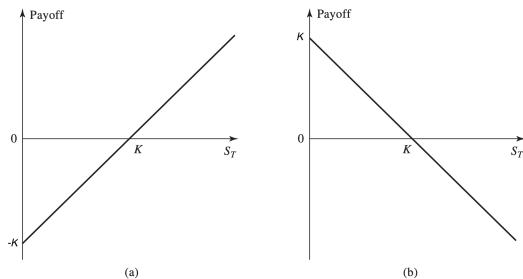


Figure 1.1: Payoffs from forward contracts: (a) long position, (b) short position.

So, there is a trade-off between the two counterparts. Our purpose is to find the value of K such that both parts will agree in entering the contract. To do that we have to avoid the presence of an *arbitrage opportunity*, i.e. a possibility of making money for free, without risk (if there is an arbitrage opportunity then fixing a price is meaningless). Quantitatively, an arbitrage opportunity is defined as:

$$V(t) = \begin{cases} 0 & t = 0 \\ V_T \geq 0 \text{ with } \mathbb{P}(V_T \geq 0) = 1 \text{ and } \mathbb{P}(V_T > 0) > 0 & t = T \end{cases} \quad (1.1)$$

In order to avoid arbitrage opportunity we can construct the following portfolio:

1. Take a long position in the forward contract. Entering the contract doesn't cost nothing and the payoff is $S_T - K$.
2. Now we need a contract which gives $-S_T$ at $t = T$. One possibility is selling short: even if we don't have the underlying we borrow it from the broker and we sell it at the current price on the market, promising to re-buy it at time $t = T$ at the price S_T .
3. Now we want to eliminate the $-K$ of step 1, so at $t = T$ we need $+K$ in any case, whatever happens to S_T . One possibility is to invest in the riskless market, which basically is the interest rate market¹. There are two classical notions of interest rate:
 - *simple compounding*, which grows linearly in time: $1 \rightarrow 1 + TR$;
 - *exponential compounding*, which grows exponentially in time: $1 \rightarrow e^{RT}$.

So we invest $-Ke^{-RT}$ in the riskless market and at the end we get $-Ke^{-RT}e^{RT} = -K$ changed of sign because at the beginning we spend and at the end we get.

Action	$t = 0$	$t = T$
Take a long position in the forward contract	0	$S_T - K$
Sell short	$+S_0$	$-S_T$
Invest in the riskless market	$-Ke^{-RT}$	$+K$
	$S_0 - Ke^{-RT}$	0

In the end the sum at $t = 0$ is $S_0 - Ke^{-RT}$ and the sum at $t = T$ is zero. Since we want to avoid the presence of no arbitrage opportunity also the sum at $t = 0$ has to be zero, otherwise there is an immediate arbitrage opportunity:

$$S_0 - Ke^{-RT} = 0 \Rightarrow K = S_0 e^{RT} \quad (1.2)$$

This is the value of K such that both parts will agree in entering the contract.

1.2 Options

Options are derivatives based on the value of underlying securities, such as stocks. An option contract offers the buyer the opportunity to buy or sell – depending on the type of contract they hold – the underlying asset. There are two types of option:

- a *call option* gives the holder the right to buy the underlying asset by a certain date for a certain price;
- a *put option* gives the holder the right to sell the underlying asset by a certain date for a certain price.

¹^To be honest, it is not completely riskless, there is the default risk.

1.2.1 Long call option

In *long call options* we have the possibility to buy the underlying at time $t = T$ at price K decided at $t = 0$. So, at maturity we will buy only if it is convenient.

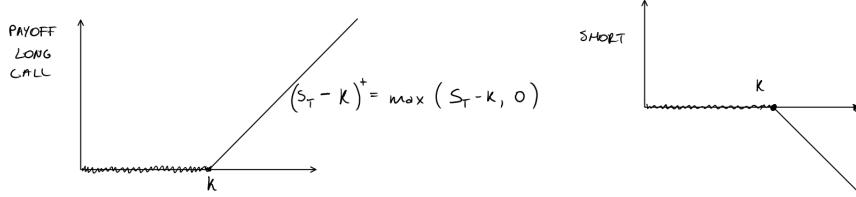


Figure 1.2: Call option payoffs.

Notice that now the payoff is no more linear. If we are in the short position at most we hope to get a zero payoff: nobody will accept this position. This means that the contract will have a price that the long position has to pay to the short position in order to enter the contract. In other words, it is as if the short position sells to the long position the possibility to buy.

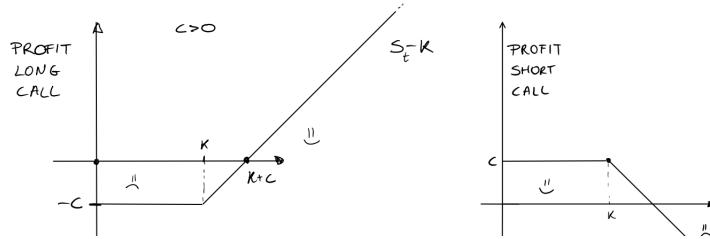


Figure 1.3: Call option profits.

If we have different strike prices the corresponding costs will be different. For example, if we have $K_1 < K_2$, since the probability to exercise the option corresponding to K_2 is smaller than the probability to exercise the option corresponding to K_1 , then c_1 will be larger than c_2 :

$$K_1 < K_2 \Rightarrow c_1 > c_2$$

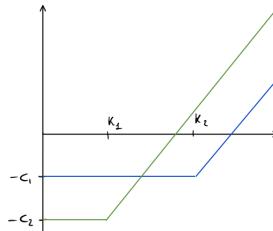


Figure 1.4: Costs for different strike prices.

1.2.2 Long put option

In *long put options* we have the possibility to sell the underlying at time $t = T$ at price K decided at $t = 0$. So, at maturity we will sell to the counterpart only if it is convenient, i.e. if $S_T < K$ ². In this case the short position has the obligation to

² ^This is used, for example, by arabian countries which want to avoid that the price of crude oil goes below a certain level.

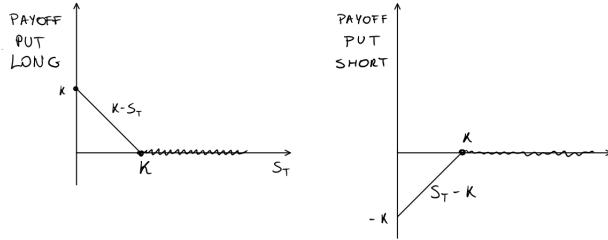


Figure 1.5: Put option payoffs.

buy. But there is no convenience in entering such a contract, so again there will be a cost to enter the contract as long position. If we have different strike prices, for

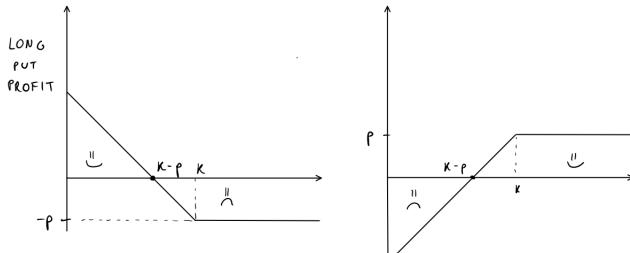


Figure 1.6: Put option profits.

example $K_1 < K_2$, since the probability to exercise the option corresponding to K_2 is bigger than the probability to exercise the option corresponding to K_1 , then c_1 will be smaller than c_2 :

$$K_1 < K_2 \Rightarrow c_1 < c_2$$

*Lesson 2 of
12/03/2020* Now we ask ourselves: which is the fair value of the price of this kind of contracts? In order to find it we need to introduce a probabilistic model.

2 | Discrete time finance

2.1 Static binomial model

The simplest probabilistic model is the *static binomial model*. Suppose that we have an underlying with initial price S_0 at $t = 0$ and which at time $t = T$ can assume two values, S_u and S_d ($u > d$). The randomness arises from the fact that at time $t = 0$ we don't know which will be the value at maturity, so we can say that S_T is a two states random variable.

Now we would like to find which is the payoff of the call option, i.e.

$$(S_T - K)^+ = \begin{pmatrix} (S_u - K)^+ \\ (S_d - K)^+ \end{pmatrix} \quad (2.1)$$

Let's consider the following toy model: suppose to have an underlying which initial price is $S_0 = 100$ and $S_T = (S_u, S_d) = (110, 90)$ and consider a strike price $K = 100$ ¹. The call payoff at the end will be

$$\begin{pmatrix} (110 - 100)^+ \\ (90 - 100)^+ \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \end{pmatrix}$$

Which is the price of the call option at time $t = 0$? Without knowing anything we can assume that the probability to get S_d or S_u is the same, i.e. $1/2$, so the expected payoff at time $t = T$ is

$$10 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = 5$$

Suppose that $T = 1$ year. We are interested in the price today, so the natural way to get it is to discount this amount of money today. Assuming that the interest rate is exponential compounding with $R = 2\%$ per year we have that the expected payoff at time $t = 0$ is

$$e^{-RT} \mathbb{E}[\text{payoff}_T] = e^{-0.02 \cdot 1} \left(10 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} \right)$$

Financially speaking, **this is completely wrong** because we are giving just our personal evaluation of the option price. Even using different probabilities for S_d and S_u we would not have a price, because we are using a personal evaluation of the probability. A price has to be objective, not subjective. So, we have to completely forget this approach.

We have to find an objective way to determine, if not the price, at least an interval for the price. The idea is to perform the same reasoning we did for the price of the forward contract, i.e. to apply the no arbitrage principle. Let's write the payoff at time T as

$$\begin{pmatrix} f^u \\ f^d \end{pmatrix}$$

The strategy is the following

¹ If $K = S_0$ we say that the strike price is *at the money*. If $K < S_0$ we say that it is *in the money* and if $K > S_0$ it is *out of the money*.

1. We build up a strategy which replicates state of the nature by state of the nature all the values assumed by the random variable payoff. The idea is that the portfolio strategy has a value V and the value of the portfolio at time T is the payoff, i.e.

$$V_T = \text{payoff}_T = \begin{pmatrix} f^u \\ f^d \end{pmatrix}$$

Define α as the number of units of the underlying we have and β as the number of obligations we have in the market (the possibility to lend or borrow money), for example the *zero coupon bond*, i.e. the contract that – whatever happens to the market or to the underlying – it gives the payoff equal to 1. There is no intermediate cash flow in the meanwhile. The terminal amount of money is denoted as *notional*, which is typically computed in one hundred. So, a zero coupon bond is a contract that gives 100 at the end and at the beginning the value of the notional is $100e^{-RT}$.

Example 2.1.1

If the price of the zero coupon bond that gives 100 at the end is 99.5 today, and the time period between today and maturity is 3 months, we have

$$99.5 = 100e^{-R \cdot 3 \text{ months}} \frac{3}{12}$$

where $3/12$ is due to the fact that typically the interest rate is computed in years. Inverting we get

$$R = -\frac{12}{3} \ln \left(\frac{99.5}{100} \right)$$

So we can invest in the risky market (α) on in the interest rate/riskless market (β). Defining the portfolio strategy

$$V = \alpha S + \beta B$$

at the beginning we have

$$V_0 = \alpha S_0 + \beta \cdot 1$$

and at the end we get

$$V_T = \alpha S_T + \beta e^{RT} = \alpha \begin{pmatrix} Su \\ Sd \end{pmatrix} + \beta \begin{pmatrix} e^{RT} \\ e^{RT} \end{pmatrix} \quad (2.2)$$

which is a random variable. This leads to a system of two equations:

$$\begin{cases} \alpha Su + \beta e^{RT} = f^u \\ \alpha Sd + \beta e^{RT} = f^d \end{cases} \quad (2.3)$$

which has solution (just subtract the second by the first)

$$\alpha = \frac{f^u - f^d}{Su - Sd}, \quad \beta = e^{RT} \left(f^d - Sd \frac{f^u - f^d}{Su - Sd} \right) \quad (2.4)$$

2. At the end $V_T = \text{payoff}_T$ with probability 1, because state of the nature by state of the nature they are the same. But if in the market there are two contracts with the same value at a certain time then necessarily they must share the same value also at the beginning:

$$V_0 = \text{price}_0(\text{payoff}_T)$$

This is due to the absence of arbitrage opportunity. In fact, suppose for example that $V_0 > \text{price}_0(\text{payoff}_T)$. Then we can construct an arbitrary strategy. In order to do that we have to understand which is the anomaly in this situation. If at the beginning the portfolio is more expensive than the option, this means that it is too expensive, because with the option – which is cheaper – we can get the same value at the end. So, V is the loser and the option is the winner and we can exploit the arbitrage by going long (buying) in the winner and going short (selling) in the loser.

	$t = 0$	$t = T$
Long option	- price_0	αpayoff
Short V	$+V_0$	$-V_T$
	$V_0 - \text{price}_0 > 0$	0

If $V_0 < \text{price}_0(\text{payoff}_T)$ we go short in the option and long in the portfolio. In the end, we get:

$$V_0 = \text{price}_0(\text{payoff}_T) = \alpha S_0 + \beta = \frac{f^u - f^d}{S_u - S_d} S_0 + e^{RT} \left(f^d - S_d \frac{f^u - f^d}{S_u - S_d} \right) \quad (2.5)$$

Notice that if using this simple model we end up with a pretty complicate expression, this means that this methodology doesn't lead to a technology which can be exploited in a realistic and more sophisticated framework. So we have to find a way to manipulate this expression in order to get something more readable.

Let's rewrite (2) as:

$$\begin{aligned} V_0 &= \alpha S_0 + \beta = \frac{f^u - f^d}{S_u - S_d} S_0 + e^{RT} \left(f^d - S_d \frac{f^u - f^d}{S_u - S_d} \right) \\ &= e^{-RT} \left[\frac{f^u - f^d}{u - d} + f^d - d \frac{f^u - f^d}{u - d} \right] \\ &= e^{-RT} \left[f^u \underbrace{\frac{e^{RT} - d}{u - d}}_q + f^d \underbrace{\frac{u - e^{RT}}{u - d}}_{1-q} \right] \\ &= e^{-RT} [f^u q + f^d (1 - q)] \end{aligned} \quad (2.6)$$

We need $q > 0$, otherwise there is an arbitrage opportunity. In fact, having a risky asset

$$S = \begin{cases} S_u \\ S_d \end{cases} \xrightarrow{\text{normalize}} 1 = \begin{cases} u \\ d \end{cases}$$

and a riskless asset

$$1 = \begin{cases} e^{RT} \\ e^{-RT} \end{cases}$$

if $e^{RT} \leq d$ it means that we can invest in the risky asset and at worst get the return d , but in any case – even in the worst case – the return is still greater than what we hope to get investing in the riskless asset. The risk asset becomes the winner! So we can exploit the arbitrage opportunity by taking long position in the winner and short position in the loser.

	$t = 0$	$t = T$
Long S	$-S_0$	$+(S_0u, S_0d)$
Short #S riskless	$+S_0$	$-(S_0e^{RT}, S_0e^{RT})$
	0	$(S_0(u - e^{RT}), S_0(d - e^{RT})) > 0$

We need also that $1 - q > 0$, which leads to $u < e^{RT}$. Again, if this does not hold we have an arbitrage opportunity (where this time the winner is the riskless asset). So in the end I get eq. (2.6), which is way more readable now, because what I get is that the price price of the payoff is equal to the discounted value of a convex combination with positive probability weights, which is exactly an expected value:

$$\text{price}_0(\text{payoff}_T) = e^{-RT} \mathbb{E}[\text{payoff}_T] \quad (2.7)$$

This expression looks the same as the one we said was completely wrong. Not exactly: the crucial difference is that I don't use my subjective probability but only elements which are present in the market. In particular, the probability weight q is not chosen by us, but it depends on market parameters, i.e. the fact that the underlying goes up and down without touching any probability or interest rate. To stress the fact that the probability is not subjective, we write

$$\text{price}_0(\text{payoff}_T) = e^{-RT} \mathbb{E}^Q[\text{payoff}_T] \quad (2.8)$$

where

$$Q \rightarrow \begin{pmatrix} q \\ 1-q \end{pmatrix} \quad (2.9)$$

is known as *risk neutral probability measure*.

So, in the end, we found a methodology for the pricing:

1. Which is the payoff $_T$ I'm interested to price?
2. Compute $q = \frac{e^{RT}-d}{u-d}$;
3. Compute the price using (2.8).

2.1.1 Implementation

Let's talk about implementation. Even if the model is very stylized, it would be nice to understand if it is possible to apply it concretely in some cases. We need to specify u and d , which values are related to the volatility of the price, i.e. to the variability of the underlying. If $u \gg d$ we are confident on the fact that future realizations of the underlying will increase, if $u \sim d$ we are not confident on the fact that future realizations of the underlying will increase.

It would be nice to be able to estimate u and d starting from real data. In order to do that we have to consider some historical series of daily closing prices (S_{-n}, \dots, S_0) and switch to returns (since we want u and d):

$$\text{Return}_t = R_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \ln \left(\frac{S_t}{S_{t-1}} \right) \quad (2.10)$$

So we get a vector (R_{-n-1}, \dots, R_0) . Then we have to compute the standard deviation

$$\sigma_{\text{daily}} = \sqrt{\frac{\sum_{i=0}^{n-1} (R_i - \bar{R})^2}{n-1}} \quad (2.11)$$

which is called *daily volatility*. u and d are given by

$$u = e^{\sigma_{year}\sqrt{T}}, \quad d = e^{-\sigma_{year}\sqrt{T}} \quad (2.12)$$

where $\sigma_{year} = \sigma_{daily}\sqrt{250}$ (250 are the opening days in the market) and T is measured in years (so if we have $T = 3$ months we have to put $T = 3/12$). Once we have u and d we can implement our binomial model

Remark 2.1.1. In principle we can price whatever. If we consider that the payoff is the underlying itself

$$\text{payoff}_T = S_T \quad (2.13)$$

i.e. the payoff gives one unit of underlying at the end, the corresponding price of this particular payoff is given by (2.8):

$$\begin{aligned} e^{-RT}\mathbb{E}^{\mathbb{Q}}[(S_T)] &= e^{-RT}(qSu + (1-q)Sd) = e^{-RT}S(qu + (1-q)d) \\ &= e^{-RT}S\left(u\frac{e^{RT}-d}{u-d} + d\frac{u-e^{RT}}{u-d}\right) = S_0 \end{aligned} \quad (2.14)$$

which makes sense, because the market value of the underlying today is exactly what we observe today, i.e. the quoted price of the underlying. So we can say the methodology is consistent. But there is something more, in the sense that if we write

$$\mathbb{E}^{\mathbb{Q}}[e^{-RT}(S_T)] = S_0e^{-R\cdot 0} = S_0 \quad (2.15)$$

we see that the expected value is the same as the value that we observe. This is a very stylized case of the property of a *martingale*. However, this is not properly a martingale (since there is no stochasticity), but just a preliminary intuition of the fact that “the process S_te^{-RT} , i.e. the discounted asset, will be a \mathbb{Q} -martingale”.

2.2 Hedging

We saw that the price of a given payoff is given by (1.1), so we solved the problem of pricing. This completely solve the problem for the long position, who wants to know which is the correct price to pay to the short position in order to get the payoff at the end.

Lesson 3 of
13/03/2020

Now we have to solve the problem for the short position. One possibility is to sell the option (possibility) and hope that the market goes in the right direction, but this it is not a good strategy. For the short position we have to build up a portfolio which value at time T is exactly the payoff we have to give to the long position. This problem is known as *hedging*. The strategy is the following:

1. Solve the pricing problem
2. Take α shares, i.e. the number of shares that we have to hold in our portfolio in order to replicate the payoff:

$$\alpha = \frac{f^u - f^d}{Su - Sd} = \frac{\Delta \text{payoff}}{\Delta S} \equiv \Delta \quad (2.16)$$

This quantity is know in general as *delta* and it represents the amount of underlying we have to get in the portfolio in order to replicate the payoff of the option.

3. Invest into the riskless market, taking β according to (2.5):

$$\beta = \text{price}_0 - \Delta S_0 \quad (2.17)$$

Example 2.2.1

Consider again the toy model where we have an underlying which initial price is $S_0 = 100$, $T = 1$ year, $S_T = (Su, Sd) = (110, 90)$, a strike price $K = 100$ and an interest rate of 1% per year. Then we have

$$u = \frac{110}{100} = 1.1 \quad d = \frac{90}{100} = 0.9$$

The corresponding q is

$$q = \frac{e^{RT} - d}{u - d} = \frac{e^{0.01} - 0.9}{1.1 - 0.9} = 0.555$$

Now we are able to compute the price of the call:

$$\text{price}_0(\text{call}) = e^{-RT}(10q + 0(1 - q)) = 10qe^{-RT} = 5.44 \quad (2.18)$$

This is the market price coherent with the no arbitrage principle of a promise that can give a payoff between 10 and 0. So, in order to use the replicating portfolio strategy, the short position has to sell the option for a price $\text{price}^{mkt} > 5.44$. The income of $\text{price}^{mkt} - 5.44$ can be put in the pocket, but we have to use 5.44 in order to build up the hedging strategy, which consists in taking

$$\Delta = \frac{f^u - f^d}{Su - Sd} = \frac{10 - 0}{110 - 90} = 0.5$$

shares in the risky market and

$$\beta = \text{price}_0 - \Delta S_0 = 5.44 - 0.5 \cdot 100 = -44.56$$

For example, if the underlying is worth 100€, I have to invest 50€ in the risky market. But if we are in the short position we don't have 50€, because we receive only 5.44€ from the long position. So we take the 5.44€ and we borrow 44.56€ from the riskless market in order to invest in the risky market^a.

In this kind of business we usually get slightly more than price_0 . For example, if $\text{price}_0 = 5.44$ € we would expect to get something like 5.54€. Notice that for a personal business of 0.10€ we are obliged to construct a trading strategy which involves an investment of 50€ in the risky market and 44.56€ in the riskless market, so we have to move a huge amount of money compared to what we get at the end. That's why it is very important to be precise in the hedging strategy.

^aIn the real market you don't buy one $\Delta = 0.5$ but you consider a package of two options and the global hedging.

2.3 Dividends

A *dividend* is a payment made by a corporation to its shareholders. The idea is that if we own one asset sometimes the corresponding corporate decides to give us – in addition to the market value of our asset – some extra bonus.

Consider a binomial model in which we have an underlying with initial value S and

final values given by the terminal values S_u, S_d plus a bonus:

$$S \rightarrow \begin{cases} S_u + S e^{\delta T} \\ S_d + S e^{\delta T} \end{cases} \quad (2.19)$$

where δ is called *dividend rate* and in general is a random variable. In such cases we say that the price is *cum dividend*, whereas if there is no dividend the price is *ex dividend*. We can repeat the same arguments we did before for the price ex dividend introducing the dividend as a shift of the price of the underlying. The probability weight becomes

$$q = \frac{e^{RT} - (d + e^{\delta T})}{u - d} \quad (2.20)$$

Remark 2.3.1. Recall that in Remark 2.1.1 we found a stylized case of martingale. Now we have

$$\mathbb{E}^{\mathbb{Q}}[S_T] = S_0(e^{RT} - e^{\delta T}) \quad (2.21)$$

so it is not true anymore that discounted assets are constant in expected value.

2.3.1 Put-call parity

One interesting thing is that it is possible to deduce the implicit dividend policy of a corporate – which usually is private – from the prices quoted in the market. In order to do that we use the *put-call parity*, which is a link between the put and the call prices.

Start by considering no dividends. Notice that any number can be written as positive part minus negative part:

$$x = x^+ - x^- = x^+ - (-x)^+$$

So, we can write

$$S_T - K = \underbrace{(S_T - K)^+}_{call} - \underbrace{(K - S_T)^+}_{put} \quad (2.22)$$

Now we use the linearity of expected value:

$$\mathbb{E}^{\mathbb{Q}}[S_T - K] = \mathbb{E}^{\mathbb{Q}}[(S_T - K)^+] - \mathbb{E}^{\mathbb{Q}}[(K - S_T)^+]$$

and multiply both sides for the discounted value:

$$e^{-RT} \mathbb{E}^{\mathbb{Q}}[S_T - K] = \underbrace{e^{-RT} \mathbb{E}^{\mathbb{Q}}[(S_T - K)^+]}_{call_0} - \underbrace{e^{-RT} \mathbb{E}^{\mathbb{Q}}[(K - S_T)^+]}_{put_0}$$

where we recognized the initial price of the call and of the put. Then we use the property $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$:

$$e^{-RT} \mathbb{E}^{\mathbb{Q}}[S_T - K] = e^{-RT} \mathbb{E}^{\mathbb{Q}}[S_T] - K e^{-RT} = \mathbb{E}^{\mathbb{Q}}[e^{-RT} S_T] - K e^{-RT} \stackrel{(2.15)}{=} S_0 - K e^{-RT}$$

In the end, we get:

$$S_0 - K e^{-RT} = call_0 - put_0 \quad \Rightarrow \quad put_0 = call_0 - S_0 + K e^{-RT} \quad (2.23)$$

Now consider the presence of dividends the put-call parity becomes

$$put_0 = call_0 - S_0 + K e^{-RT} + S_0 e^{(\delta - R)T} \quad (2.24)$$

that means that if for the same asset there is a market of put options and call options with a certain maturity, then we consider the list of the quoted prices of the put and of the call and use eq. (2.24) to find δ .

2.4 Market completeness

Let's consider the following assumptions:

1. short positions and fractional holdings of the underlying are allowed;
2. there is no *bid-ask spread*, i.e. there is no difference between the prices quoted for an immediate sale (ask) and an immediate purchase (bid) for an underlying;
3. there are no transaction costs;
4. there is a perfectly liquid market, i.e. there is the possibility to buy or sell an unlimited quantity of an underlying.

Definition 2.4.1 (Reachable contingent claim). *A contingent claim² $x \in \mathbb{R}^2$ is reachable if there exists a portfolio $h(\alpha, \beta)$ with terminal value $V_T^h = x$ almost surely (i.e. with probability 1). In this case h is called hedging portfolio.*

Definition 2.4.2 (Complete market). *If all the claims are reachable, the market is complete.*

So, in a complete market it is possible to replicate all the possible random variables, i.e. all the possible contracts.

Proposition 2.4.1. *If a claim x is reachable with h and there is absence of arbitrage opportunity, then the price of x is*

$$\text{price}_0(x) = V_0^h. \quad (2.25)$$

Remark 2.4.1. If a claim x is reachable than Prop. 2.4.1 holds also in an incomplete market.

Proposition 2.4.2. *The binomial model with $d < e^{RT} < u$ is arbitrary free and complete.*

So, in a static binomial model it is possible to price everything. The next step is to consider a dynamic model.

2.5 The multi period case

We want to generalize the binomial model

$$S \rightarrow \begin{cases} S_u \\ S_d \end{cases}$$

For example we can introduce a third possible realization of the underlying, obtaining a trinomial model:

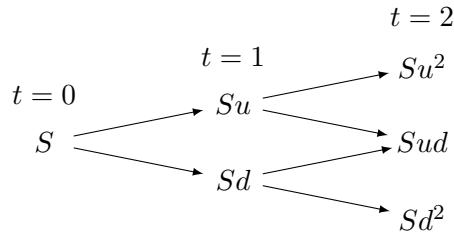
$$S \rightarrow \begin{cases} S_u \\ S_m \\ S_d \end{cases}$$

² A contingent claim is a derivative whose future payoff depends on the value of another “underlying” asset, or more generally, that is dependent on the realization of some uncertain future event.

In this case, since we have only two tools in our investment strategy (the risky and riskless markets) if, for example, we want to replicate a call option, the payoff is

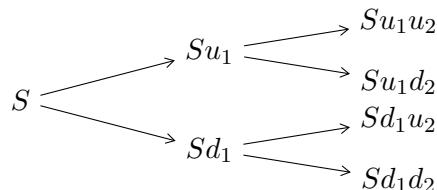
$$(S_T - K)^+ = \begin{pmatrix} (Su - K)^+ \\ (Sm - K)^+ \\ (Sd - K)^+ \end{pmatrix}$$

and we have to solve three equations with two unknowns, which – apart from special cases – is impossible to solve. So, just adding one more possible outcome leads to uncomfortable consequences in terms of pricing methodology, because our model is no more complete. We have the possibility to invest in the risky market and in the riskless market in order to follow the fluctuations of something which can take three values, so the underlying fluctuates too much. We need another instrument to follow the fluctuations of the underlying, for example an extra underlying S' , but things become complicated. It is much more convenient to introduce a dynamic framework. For example, consider the following tree.



In this case the tree is stationary, i.e. u and d are always the same for every t . This tree is called *recombining tree* or *lattice*. Computationally speaking this is an advantage because if we have n steps then the terminal values are $n + 1$, so the number of all the possible realizations remains linear with respect to time.

Conversely, if we consider a non-recombining tree, like



for n steps we have 2^n possible realizations, which for large n is impossible to manage. The decision between combining and non-combining tree is crucial, and it should be motivated by some financial argument. Why should we use non-recombining trees? What we observe in the market is that when S goes up the volatility is low, whereas when the market decreases there is an increase of the volatility. This is the reason

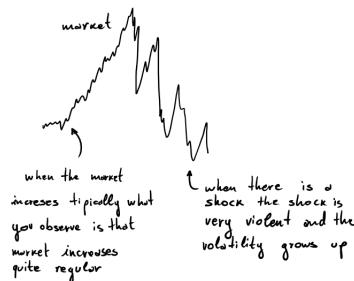


Figure 2.1: Volatility behavior.

why it could be interesting to consider a non-recombining tree. However, if we choose this kind of model there is a problem in the implementation. Recall that in the binomial model u and d are given by $(u, d) \sim e^{\pm\sigma\sqrt{T}}$: the problem is we have only one historical series, and if the parameters vary step by steps it is impossible to estimate them. So we drop this model and we use a recombining tree.

Remark 2.5.1. Sometimes it is not possible to use a recombining tree because there are same contracts which payoff depends on the whole past history of the underlying. So, in these cases it is not enough to take a picture of the underlying at the maturity. For example, in the 70s some manipulations have been used in order to rise the price of crude oil immediately before the maturity and then exercise the call option, buying at $K < S_T$ and getting $S_T - K > 0$. In order to avoid this kind of manipulations, contracts which payoff depends on the whole past history of the underlying – called *asian options* – have been introduced. Asian options payoff is given by

$$\text{payoff}_T = \left(\frac{1}{T} \int_0^T S_s ds - K \right)^+ \quad (2.26)$$

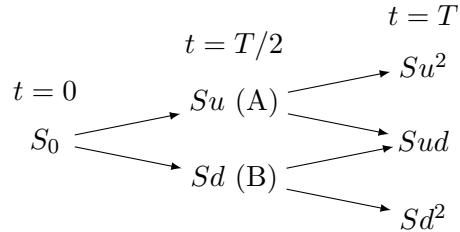
Lesson 4 of
18/03/2020

Since we are interested in pricing simple options, we consider a *stationary binomial recombining tree* or *lattice*, i.e. a model in which the way the underlying goes up and down increases or decreases is always the same.

2.6 Binomial lattice

2.6.1 Pricing options

Consider the following binomial lattice:



and a call option, whose payoff is the random vector

$$\text{payoff}_T = \begin{pmatrix} (Su^2 - K)^+ \\ (Sud - K)^+ \\ (Sd^2 - K)^+ \end{pmatrix} = \begin{pmatrix} f^{u^2} \\ f^{ud} \\ f^{d^2} \end{pmatrix}$$

The problem is always the same: how much do we have to pay at time $t = 0$ in order to get this payoff at time $t = T$?

Suppose that we are in node (A). According to the risk neutral methodology, the value of the relative sub-payoff is given by

$$\begin{aligned} \text{Value}_A \left(\begin{pmatrix} f^{u^2} \\ f^{ud} \end{pmatrix} \right) &= e^{-R\frac{T}{2}} \mathbb{E}^Q \left(\left. \begin{pmatrix} f^{u^2} \\ f^{ud} \end{pmatrix} \right| (A) \right) \\ &= e^{-R\frac{T}{2}} (qf^{u^2} + (1-q)f^{ud}) \end{aligned} \quad (2.27)$$

where

$$q = \frac{e^{R\frac{T}{2}} - d}{u - d} \quad (2.28)$$

Notice that in this case the expected value is a conditional expected value. In analogy, in the node (B) we have

$$Value_B \left(\begin{pmatrix} f^{ud} \\ f^{d^2} \end{pmatrix} \right) = e^{-R\frac{T}{2}} (q f^{ud} + (1 - q) f^{d^2}) \quad (2.29)$$

Now, we can compute the initial price of the whole payoff:

$$\begin{aligned} price_0 \left(\begin{pmatrix} f^{u^2} \\ f^{ud} \\ f^{d^2} \end{pmatrix} \right) &= price_0 \left(\begin{pmatrix} Value_A \\ Value_B \end{pmatrix} \right) \\ &= e^{-R\frac{T}{2}} \mathbb{E}^{\mathbb{Q}} \left[\begin{pmatrix} Value_A \\ Value_B \end{pmatrix} \right] \\ &= e^{-R\frac{T}{2}} (q Value_A + (1 - q) Value_B) \\ &= e^{-R\frac{T}{2}} \left[q \left(e^{-R\frac{T}{2}} (q f^{u^2} + (1 - q) f^{ud}) \right) + (1 - q) \left(e^{-R\frac{T}{2}} (q f^{ud} + (1 - q) f^{d^2}) \right) \right] \\ &= e^{-RT} [q^2 f^{u^2} + 2q(1 - q) f^{ud} + (1 - q)^2 f^{d^2}] \end{aligned} \quad (2.30)$$

Notice that the three coefficients are positive and their sum gives 1, so again we obtained a convex combination of elements, i.e. an expected value:

$$price_0 \left(\begin{pmatrix} f^{u^2} \\ f^{ud} \\ f^{d^2} \end{pmatrix} \right) = e^{-RT} \mathbb{E}^{\mathbb{Q}} \left[\begin{pmatrix} f^{u^2} \\ f^{ud} \\ f^{d^2} \end{pmatrix} \right] = e^{-RT} \mathbb{E}^{\mathbb{Q}} [\text{payoff}_T] \quad (2.31)$$

The formula has the same form as before but it must be contextualized as the expectation under the risk neutral probability measure

$$Q \sim \begin{pmatrix} q \\ 2q(1 - q) \\ q^2 \end{pmatrix} \quad (2.32)$$

2.6.2 Hedging

We now want to implement the corresponding hedging portfolio, i.e. the replicating portfolio which allows the short position to construct the payoff for the long position. First we have to solve the pricing procedure for all the intermediate nodes, in fact hedging is a forward procedure. Then, the number of shares we have to put in our portfolio is

$$\Delta_0 = \frac{Value_A - Value_B}{Su - Sd} \quad (2.33)$$

and the corresponding investment in the riskless market is

$$\beta_0 = p_0 - \Delta_0 S \quad (2.34)$$

where $p_0 = V_0 = price_0$. This is the initial step. Now we have to consider the intermediate step, according to the possible values of the underlying:

$$\Delta_A = \frac{f^{u^2} - f^{ud}}{Su^2 - Sud}, \quad \Delta_B = \frac{f^{ud} - f^{d^2}}{Sud - Sd^2} \quad (2.35)$$

Now we have

$$Value_A = \Delta_A Su + \beta_A e^{R\frac{T}{2}}, \quad Value_B = \Delta_B Sd + \beta_B e^{R\frac{T}{2}}$$

so the investment in the riskless market is

$$\beta_A = (Value_A - \Delta_A S_u) e^{-R \frac{T}{2}}, \quad \beta_B = (Value_B - \Delta_B S_d) e^{R \frac{T}{2}} \quad (2.36)$$

So, in a dynamic binomial model, also the hedging is dynamic. This is obvious, since we can exploit the possibility to change our strategy according to the fluctuations of the underlying.

2.6.3 Dynamic interest rates

Remark 2.6.1. There are two possible riskless assets:

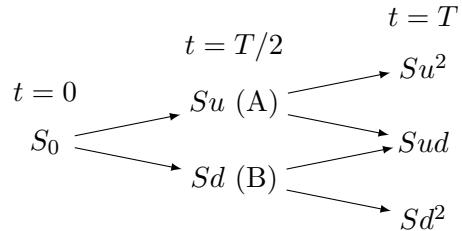
- the one which consider the forward evolution of the asset, used for example in savings accounts
- the zero coupon bond, which consider the backward evolution of the asset

If the interest rate is constant, the two are equivalent. Moreover, the two are equivalent even if the interest rate evolves and its dynamic is deterministic.

If we consider different maturity times we can find which is the corresponding value of one unit of cash, i.e. 1€. It is intuitive to expect that for a larger maturity there will be a larger interest rate. However, there are some cases in which the interest rate decreases in time. This typically happens after a financial crisis. In other cases the curve will assume other shapes.

The input to find which is the interest rate corresponding to a certain maturity are the prices of the zero coupon bound. So, starting from now we have to take into account different rates with respect to different maturities.

Recall the binomial lattice



and focus on node (A). We have two different interest rates, $R_{T/2} = R(0, T/2)$ and $R_T = R(0, T)$, called *spot interest rates* because they start from $t = 0$. The problem is to compute q : we would like to find an expression similar to (2.28) but we need $R = R(0, T/2, T)$ – i.e. the interest rate decided at $t = 0$ that starts to capitalize at $t = T/2$ until $t = T$ – which is not defined yet. $R = R(0, T/2, T)$ is called *forward interest rate*. We would like to deduce the value of the forward interest rate starting from the spot interest rates. If we put 1€ at $t = 0$ we have

$$1 \longrightarrow e^{R(0, \frac{T}{2}) \frac{T}{2}} \longrightarrow e^{R(0, \frac{T}{2}) \frac{T}{2}} e^{R(0, \frac{T}{2}, T) \frac{T}{2}}$$

which has to be consistent with the one shot capitalization using

$$1 \longrightarrow e^{R(0, T)T}$$

otherwise there is an arbitrage opportunity. We obtain the following equation:

$$e^{R(0, \frac{T}{2}) \frac{T}{2}} e^{R(0, \frac{T}{2}, T) \frac{T}{2}} = e^{R(0, T)T} \quad (2.37)$$

which has solution

$$R\left(0, \frac{T}{2}, T\right) = \frac{R(0, T)T + R\left(0, \frac{T}{2}\right)\frac{T}{2}}{\frac{T}{2}} \quad (2.38)$$

Now we can define the risk neutral probability in the node (A):

$$q_A = \frac{e^{R\left(0, \frac{T}{2}, T\right)\frac{T}{2}} - d}{u - d} \quad (2.39)$$

The same argument holds for node (B), so the expression for q_B is exactly the same of q_A . The interest rate at $t = 0$ will be

$$q_0 = \frac{e^{R\left(0, \frac{T}{2}, T\right)\frac{T}{2}} - d}{u - d} \quad (2.40)$$

So, denoting $q_A = q_B \equiv q_1$, the price at $t = 0$ is given by

$$\text{price}_0 \begin{pmatrix} f^{u^2} \\ f^{ud} \\ f^{d^2} \end{pmatrix} = e^{-R(0,T)T} \mathbb{E}^{\mathbb{Q}}[\text{payoff}_T] \quad (2.41)$$

where

$$Q \sim \begin{pmatrix} q_0 q_1 \\ q_0(1 - q_1) + (1 - q_0)q_1 \\ (1 - q_0)(1 - q_1) \end{pmatrix} \quad (2.42)$$

Expanding (2.41) we get

$$\text{price}_0 \begin{pmatrix} f^{u^2} \\ f^{ud} \\ f^{d^2} \end{pmatrix} = e^{-R(0,T)T} (f^{u^2} q_0 q_1 + f^{ud} (q_0(1 - q_1) + (1 - q_0)q_1) + f^{d^2} (1 - q_0)(1 - q_1)) \quad (2.43)$$

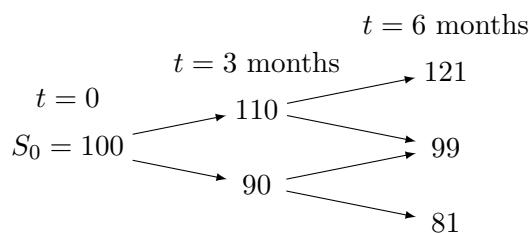
2.6.4 Binary call at the money

We consider, as an example, the binary call at the money. In this case, the payoff is

$$\text{payoff}_T = \mathbb{1}_{S_T \geq K} \quad (2.44)$$

Lesson 5 of
19/03/2020

and $K = S_0$ (at the money). So, when it is possible to exercise the contract, the cash flow is always equal to 1. Consider a toy model with $u = 1.1$, $d = 0.9$ and $S_0 = K = 100$:



The payoff is

$$\text{payoff}_T = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad (2.45)$$

We would like to price this option. The input are the zero coupon bonds at these maturities:

1. The first has notional value 1 at 3 months and its price today is 1.001.

$$1.001 = e^{-R(0,3m)\frac{3}{12}} \cdot 1 \quad \Rightarrow \quad R(0,3m) = -0.001 = -0.1\%$$

This means that the interest rate is negative and it is as if we pay in order to keep our money safe.

2. The second has notional value 1 at 6 months and its price today is 1, so the interest rate is zero.

In order to find which is the forward interest rate, we use eq. (2.37):

$$e^{R(0,3m)\frac{3}{12}} e^{R(0,3m,6m)\frac{3}{12}} = e^{R(0,6m)\frac{6}{12}} \quad (2.46)$$

which has solution

$$R(0,3m,6m) = \underbrace{2R(0,6m)}_{=0} - R(0,3m) = 0.1\% \quad (2.47)$$

Therefore, the risk neutral probability weights are given by

$$q_1 = \frac{e^{R(0,3m,6m)\frac{3}{12}} - d}{u - d} = 4.501, \quad q_0 = \frac{e^{R(0,3m)\frac{3}{12}} - d}{u - d} = 4.499 \quad (2.48)$$

Now we are able to price the corresponding binary option:

$$price_0^{BinCall} = e^{-RT} \mathbb{E}^{\mathbb{Q}}[\text{payoff}_T] = e^{-RT} \mathbb{E}^{\mathbb{Q}} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (2.49)$$

where

$$Q \sim \begin{pmatrix} q_0 q_1 \\ (1 - q_0)q_1 + q_0(1 - q_1) \\ (1 - q_0)(1 - q_1) \end{pmatrix} \quad (2.50)$$

which gives

$$\begin{aligned} price_0^{BinCall} &= e^{-R(0,6m)\frac{6}{12}} (1 \cdot q_0 q_1 + 0 \cdot (1 - q_0)q_1 + q_0(1 - q_1) + 0 \cdot (1 - q_0)(1 - q_1)) \\ &= e^{-R(0,6m)\frac{6}{12}} q_0 q_1 \end{aligned} \quad (2.51)$$

For the hedging the story is the same. [scrivi](#)

2.7 General binomial model

Now we consider a generic number of time steps, so that the underlying is $S_T \in \mathbb{R}^{t+1}$ and has terminal value

$$\begin{pmatrix} Su^t \\ Su^{t-1}d \\ \vdots \\ Su^j d^{t-j} \\ \vdots \\ Sud^{t-1} \\ Sd^t \end{pmatrix} \quad (2.52)$$

where j represents the number of up moves in the lattice. The payoff (for the call option) we want to compute is

$$\text{payoff}_T = (S_T - K)^+ \quad (2.53)$$

Then the value of the corresponding replicating portfolio is

$$\begin{aligned} P_t(j) &= V_t(j) \stackrel{(a)}{=} e^{-R \cdot 1} (qV_{t+1}(j+1) + (1-q)V_{t+1}(j)) \\ &= e^{-R} \mathbb{E}^{\mathbb{Q}}[V_{t+1}|j] \end{aligned} \quad (2.54)$$

where in (a) we used the backward induction procedure using unitary time steps and constant $q = \frac{e^R - d}{u - d}$. At the end, the corresponding value is

$$P_T(j) = (S_0 u^j d^{T-j} - K)^+ \quad (2.55)$$

Considering the hedging, we have

$$\Delta_t(j) = \frac{V_{t+1}(j+1) - V_{t+1}(j)}{S_t u - S_t d}, \quad \beta_t(j) = e^{-Rt} (V_t(j) - \Delta_t(j) S_t) \quad (2.56)$$

So the initial price can be written as:

$$p_0(\text{call}) = e^{-RT} \sum_{j=0}^T \binom{T}{j} q^j (1-q)^{T-j} (S u^j d^{T-j} - K)^+ \quad (2.57)$$

Remark 2.7.1. Notice that the probability distribution

$$Q \sim \begin{pmatrix} q^T \\ \vdots \\ \binom{T}{j} q^j (1-q)^{T-j} \\ \vdots \\ (1-q)^T \end{pmatrix} \quad (2.58)$$

is normalized, in fact

$$\sum_{j=0}^T \binom{T}{j} q^j (1-q)^{T-j} = (q+1-q)^T = 1 \quad (2.59)$$

where we used the fact that $(a+b)^n = \sum_{j=0}^n \binom{n}{j} a^j b^{n-j}$.

2.8 From the binomial model to a geometric brownian motion

Now we want to see if the binomial model leads to convergence if we take a number of steps which tends to infinity. Let's consider a binomial model with unitary time steps, $n < T$ and $T = m$. The underlying at time $n+1$ is defined as

$$S_{n+1} = S_n Z_{n+1} \quad (2.60)$$

where Z_n are i.i.d. random variables with binomial distribution and risk neutral probability

$$Q(Z_{n+1} = u_n) = q_n, \quad Q(Z_{n+1} = d_n) = 1 - q_n \quad (2.61)$$

with

$$q_n = \frac{e^R - d_n}{u_n - d_n} \quad (2.62)$$

In principle, if we consider this particular binomial model and we consider the discounted asset $S_n e^{-Rn}$, it turns out to be a Q -martingale, i.e. a martingale under the measure \mathbb{Q} . This means that $S_n e^{-Rn}$ is constant in expected value and that

$$\mathbb{E}^{\mathbb{Q}}[Z_n] = e^{R\frac{T}{n}} \quad (2.63)$$

In order to prove it, it is sufficient to start from eq. (2.60) to do (6:50).

We would like to shift from the multiplicative structure of (2.60) to an additive structure. This can be done just by taking the logarithm of eq. (2.60):

$$\ln S_n = \ln S_0 + \sum_{i=1}^n \ln Z_i \quad (2.64)$$

The sum in (2.64) is a sum of bounded i.i.d random values, so we can use the Central Limit Theorem in order to show that it converges to a gaussian random variable. However, the convergence is not well defined as $\text{Low}(\ln Z_i)$ depends on n and when the random variable depends on the number of random variables considered we have to be careful in considering the limit. So we can not use a simple version of the CLT. The idea is to show that

$$\text{Low}(S_n) \rightarrow \text{Low}(S_T) \quad (2.65)$$

where S_T is the terminal value of a process S_t satisfying

$$\frac{dS_t}{dt} = rdt + \sigma dW_t^{\mathbb{Q}} \quad (2.66)$$

where $W_t^{\mathbb{Q}}$ is a brownian motion under the probability measure \mathbb{Q} . In order to study the convergence it is better to work with the corresponding characteristic function, which is the Fourier transform of the pdf:

$$\begin{aligned} \phi_{\ln S_n}(\lambda) &= \mathbb{E}^{\mathbb{Q}}[e^{i\lambda \ln S_n}] \stackrel{(2.64)}{=} e^{i\lambda \ln S_0} \left(\mathbb{E}^{\mathbb{Q}}[e^{i\lambda \ln Z}] \right)^n \\ &= S_0^{i\lambda} \left(qe^{i\lambda \ln n} + (1-q)e^{i\lambda \ln d} \right)^n \end{aligned} \quad (2.67)$$

Now we define the relation between (u, d) and (r, σ) . Let's take

$$u_n = e^{\sigma \sqrt{\frac{T}{n}}}, \quad d_n = e^{-\sigma \sqrt{\frac{T}{n}}} \quad (2.68)$$

which is in line with what we did in sec. 2.1.1. Now we have to link the interest rate for the period in the binomial model with the corresponding interest rate in the continuous time version:

$$e^R = e^{r\frac{T}{n}} \quad (2.69)$$

where R is the discrete time interest rate and r is the continuous time interest rate. Plugging this mapping in (2.67) we get:

$$\begin{aligned} \phi_{\ln S_n}(\lambda) &= S_0^{i\lambda} \left(qe^{i\lambda \sigma \sqrt{\frac{T}{n}}} + (1-q)e^{-i\lambda \sigma \sqrt{\frac{T}{n}}} \right)^n \\ &\stackrel{(a)}{=} S_0^{i\lambda} \left[q \left(1 + i\lambda \sigma \sqrt{\frac{T}{n}} - \frac{1}{2} \lambda^2 \sigma^2 \frac{T}{n} + \mathcal{O}(n^{-1}) \right) + \right. \end{aligned}$$

$$\begin{aligned}
& + (1-q) \left(1 - i\lambda\sigma \sqrt{\frac{T}{n}} - \frac{1}{2}\lambda^2\sigma^2 \frac{T}{n} + \mathcal{O}(n^{-1}) \right)^n \\
& = S_0^{i\lambda} \left(1 + i\lambda\sigma \sqrt{\frac{T}{n}} (2q-1) - \frac{1}{2}\lambda^2\sigma^2 \frac{T}{n} + \mathcal{O}(n^{-1}) \right)^n
\end{aligned} \tag{2.70}$$

where in (a) we used the Taylor expansion. Now, recall that $q = \frac{e^R - d}{u - d}$, so using (2.68), (2.69) and Taylor expanding we get

$$\begin{aligned}
q &= \frac{e^{r\frac{T}{n}} - e^{-\sigma\sqrt{\frac{T}{n}}}}{e^{\sigma\sqrt{\frac{T}{n}}} - e^{-\sigma\sqrt{\frac{T}{n}}}} \\
&= \frac{1 + r\frac{T}{n} - 1 + \sigma\sqrt{\frac{T}{n}} - \frac{1}{2}\sigma^2\frac{T}{n} + \mathcal{O}(n^{-1})}{1 + \sigma\sqrt{\frac{T}{n}} + \frac{1}{2}\sigma^2\frac{T}{n} - 1 + \sigma\sqrt{\frac{T}{n}} - \frac{1}{2}\sigma^2\frac{T}{n} + \mathcal{O}(n^{-1})} \\
&= \frac{\sigma + \left(r - \frac{\sigma^2}{2}\right)\sqrt{\frac{T}{n}} + \mathcal{O}(n^{-1/2})}{2\sigma + \mathcal{O}(n^{-1/2})} \\
&= \frac{1}{2} + \frac{1}{2\sigma} \left(r - \frac{\sigma^2}{2}\right) \sqrt{\frac{T}{n}} + \mathcal{O}(n^{-1/2})
\end{aligned} \tag{2.71}$$

Then we have

$$2q - 1 = \frac{1}{\sigma} \left(r - \frac{\sigma^2}{2}\right) \sqrt{\frac{T}{n}} + \mathcal{O}(n^{-1/2}) \tag{2.72}$$

and we can put it into (2.70):

$$\begin{aligned}
\phi_{\ln S_n}(\lambda) &= S_0^{i\lambda} \left(1 + i\lambda\sigma \left(r - \frac{\sigma^2}{2}\right) \frac{1}{\sigma} \frac{T}{n} - \frac{1}{2}\lambda^2\sigma^2 \frac{T}{n} + \mathcal{O}(n^{-1}) \right)^n \\
&= S_0^{i\lambda} \left[1 + \frac{1}{n} \left(i\lambda T \left(r - \frac{\sigma^2}{2}\right) - \frac{1}{2}\lambda^2\sigma^2 T\right) \right]^n
\end{aligned} \tag{2.73}$$

Recall that

$$\lim_{n \rightarrow +\infty} \left(1 + \frac{a}{n}\right)^n = e^a$$

so if in (2.73) we define $a = \left(i\lambda T \left(r - \frac{\sigma^2}{2}\right) - \frac{1}{2}\lambda^2\sigma^2 T\right)$, we have that

$$\lim_{n \rightarrow +\infty} \phi_{\ln S_n}(\lambda) = S_0^{i\lambda} \exp \left[i\lambda T \left(r - \frac{\sigma^2}{2}\right) - \frac{1}{2}\lambda^2\sigma^2 T \right] \tag{2.74}$$

Now we have to understand how (2.74) is related to the characteristic function of the terminal value of $\ln S_T$ in the continuous time framework, $\mathbb{E}^{\mathbb{Q}} [e^{i\lambda \ln S_T}]$. We will find that the solution of (2.66) is

$$S_T = S_0 e^{\left(r - \frac{\sigma^2}{2}\right)T + \sigma W_T^{\mathbb{Q}}} \tag{2.75}$$

$W_T^{\mathbb{Q}} \sim \mathcal{N}(0; T)$ so the brownian motion has a gaussian distribution with variance equal to the time window we are considering. This means that S_T is a random variable distributed as

$$S_T \sim S_0 \exp \left[\underbrace{\ln S_0 + \left(r - \frac{\sigma^2}{2}\right) T}_{deterministic} + \underbrace{\sigma \mathcal{N}(0; T)}_{stochastic} \right] \tag{2.76}$$

The deterministic part represents just a shift in the expected value, so we can rewrite (2.76) as

$$S_T \sim S_0 \exp \left[\sigma \mathcal{N} \left(\ln S_0 + \left(r - \frac{\sigma^2}{2} \right) T; \sigma^2 T \right) \right] \quad (2.77)$$

So S_T follows a log-normal distribution. Now recall the characteristic function of a gaussian variable

$$\mathbb{E}^{\mathbb{Q}} \left[e^{i\lambda \mathcal{N}(\mu; \sigma^2)} \right] = e^{i\lambda \mu - \frac{\sigma^2 \lambda^2}{2}} \quad (2.78)$$

In our case we have:

$$\mathbb{E}^{\mathbb{Q}} \left[e^{i\lambda \ln S_T} \right] = \mathbb{E}^{\mathbb{Q}} \left[e^{i\lambda \mathcal{N}(\ln S_0 + \left(r - \frac{\sigma^2}{2} \right) T; \sigma^2 T)} \right] \quad (2.79)$$

$$= \exp \left(i\lambda \ln S_0 + i\lambda \left(r - \frac{\sigma^2}{2} \right) T - \frac{\sigma^2 \lambda^2}{2} T \right) \quad (2.80)$$

$$= S_0^{i\lambda} \exp \left(i\lambda \left(r - \frac{\sigma^2}{2} \right) T - \frac{\sigma^2 \lambda^2}{2} T \right) \quad (2.81)$$

This is the characteristic function of the geometric brownian motion. Going back to (2.74) we find exactly the same expression, so the binomial model converges to the characteristic function of the corresponding solution of the geometric brownian motion.

2.8.1 Call option price in binomial model: the Black & Scholes formula

*Lesson 6 of
20/03/2020*

The call option price in a n steps binomial model is given by the one shot discounted value of the risk neutral expectation of the payoff:

$$p_0(\text{call}) = e^{-RT} \sum_{j=0}^n \binom{n}{j} q^j (1-q)^{n-j} (Su^j d^{n-j} - K)^+ \quad (2.82)$$

We would like to remove the non-linearity in the positive part. In order to do that we have to find when $(Su^j d^{n-j} - K)^+ \geq 0$, so that we can remove the positive part. So we look for the smallest index j such that $(Su^j d^{n-j} - K)^+ \geq 0$:

$$\hat{j} = \inf \{j : S_0 u^2 d^{n-j} \geq K\} \quad (2.83)$$

Taking the logarithm of $S_0 u^2 d^{n-j} \geq K$ we get:

$$\ln S_0 + j \ln n + (n-j) \ln d \geq \ln K \quad (2.84)$$

$$j(\ln n - \ln d) \geq \ln \frac{K}{S_0} - n \ln d \quad (2.85)$$

$$\hat{j} = \left\lfloor \frac{\ln \frac{K}{S_0} - n \ln d}{\ln n - \ln d} \right\rfloor \stackrel{(2.68)}{=} \left\lfloor \frac{\ln \frac{K}{S_0} + n\sigma \sqrt{\frac{T}{n}}}{2\sigma \sqrt{\frac{T}{n}}} \right\rfloor = \left\lfloor \frac{n}{2} - \frac{\ln \frac{S_0}{K}}{2\sigma \sqrt{\frac{T}{n}}} \right\rfloor \quad (2.86)$$

Now we can rewrite (2.82) as:

$$p_0(\text{call}) = e^{-RT} \sum_{j=\hat{j}}^n \binom{n}{j} q^j (1-q)^{n-j} Su^j d^{n-j} - K e^{-RT} \sum_{j=\hat{j}}^n \binom{n}{j} q^j (1-q)^{n-j}$$

$$= S \sum_{j=\hat{j}}^n \binom{n}{j} \left(e^{-\frac{RT}{n}} qu \right)^j \left(e^{-\frac{RT}{n}} (1-q)d \right)^{n-j} - K e^{-RT} \sum_{j=\hat{j}}^n \binom{n}{j} q^j (1-q)^{n-j} \quad (2.87)$$

where we took S out of the sum, e^{RT} inside it and split the resulting sum in two components. Notice that

$$e^{-\frac{RT}{n}} qu + e^{-\frac{RT}{n}} (1-q)d = 1, \quad \Rightarrow \quad q + 1 - q = 1$$

so we can interpret them as probability weights. In particular we can see them as the expansion of the binomial distribution, in fact the cumulated risk neutral probability of a binomial random variable parametrized by p and n is³

$$Q(\mathcal{B}(p, n) \geq a) = \sum_{j=a}^n \binom{n}{j} p^j (1-p)^{n-j} \quad (2.88)$$

and the terms in eq. (2.87) have the same structure. So we get the intuition that applying carefully the CLT we can say that

$$e^{-\frac{RT}{n}} qu \xrightarrow[n \rightarrow \infty]{\text{CLT}} \mathcal{N}(ne^{-\frac{RT}{n}} qu; ne^{-\frac{RT}{n}} qu(1 - e^{-\frac{RT}{n}} qn)) \quad (2.89)$$

$$e^{-\frac{RT}{n}} (1-q)d \xrightarrow[n \rightarrow \infty]{\text{CLT}} \mathcal{N}(nq; nq(1-q)) \quad (2.90)$$

Now (2.87) becomes

$$p_0(\text{call}) = SQ(\mathcal{B}(e^{-\frac{RT}{n}} qu; n) \geq \hat{j}) - KQ(\mathcal{B}(q; n) \geq \hat{j}) \quad (2.91)$$

This is the so called *Black & Scholes formula*. Then we consider the second part. Recall that

$$\begin{aligned} \text{Prob}(\mathcal{N}(\mu, \sigma^2) > \hat{j}) &= \text{Prob}\left(\mathcal{N}(0, 1) > \frac{\hat{j} - \mu}{\sigma}\right) \\ &\stackrel{(a)}{=} \text{Prob}\left(\mathcal{N}(0, 1) < \frac{\mu - \hat{j}}{\sigma}\right) \\ &\stackrel{(b)}{=} \Phi\left(\frac{\mu - \hat{j}}{\sigma}\right) \end{aligned} \quad (2.92)$$

where in (a) we used the symmetry of the standardized gaussian and in (b) we recognized the *cumulative probability function*

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (2.93)$$

In our case $\mu = nq$ and $\sigma = \sqrt{nq(1-q)}$, so we get:

$$\begin{aligned} \Phi\left(\frac{nq - \hat{j}}{\sqrt{nq(1-q)}}\right) &= \Phi\left(\left(nq - \frac{n}{2} - \frac{\ln \frac{S_0}{K}}{2\sigma\sqrt{\frac{T}{n}}}\right) \frac{1}{\sqrt{nq(1-q)}}\right) \\ &= \Phi\left(\frac{\sqrt{n}(q - \frac{1}{2})}{\sqrt{q(1-q)}} + \frac{\ln \frac{S_0}{K}}{2\sigma\sqrt{T}\sqrt{q(1-q)}}\right) \end{aligned}$$

³Recall that a binomial random variable is the sum of n Bernoulli independent random variables with parameter p .

$$= \Phi \left(\frac{\ln \frac{S_0}{K} + \left(r - \frac{\sigma^2}{2} \right) T}{\sigma \sqrt{T}} \right) \quad (2.94)$$

where in the first step we used the definiton of \hat{j} (2.86) and in the last we used the following limits:

$$q(1-q) \xrightarrow{n \rightarrow \infty} \frac{1}{4}, \quad \sqrt{n} \left(q - \frac{1}{2} \right) \xrightarrow{n \rightarrow \infty} \frac{\left(r - \frac{\sigma}{2} \sqrt{T} \right)}{2\sigma} \quad (2.95)$$

dimostra. So the gaussian cumulative probability function can be written in terms of the continuous time variables. The same argument holds for the the first term of (2.87), so that with this limit procedure we recovered the Black & Scholes formula (2.91).

This is just an intuition in order to understand that not only the underlying but also the call option price is quite robust in this procedure. So the binomial model approximates very well the continuous time pricing model and this is the reason why nowadays this model is widely used in the banking industry.

We will come back to the Black & Scholes formula after introducing the stochastic calculus in continuous time.

2.9 American options

So far we considered the cases in which the payoff was $(S_T - K)^+$ and in which we can exercise the option to buy/sell only at the maturity. These options are called *European options*. On the contrary, in *American options* we can exercise the option – i.e. the possibility – at any time $t \leq T$. It is now obvious to compute the payoff $(S_\tau - K)^+$ in American options, since we don't know the time τ at which we will be interested to exercise. So, in order to compute the initial price of an American call we need an information involving future time, which is not measurable. The time τ is a random variable called *stopping time*.

2.9.1 Call option with no dividends and $R > 0$

Now, consider a special case in which we have a call option on a non-dividend paying. In this case the American call price at any time t will be always greater than the price of an European contract:

$$C^{AM}(t, S_t) \geq C^{EU}(t, S_t)$$

because the American option includes the opportunity to exercise at the end but also at any time t , so it is better than the European one. Furthermore, we have that

$$C^{AM}(t, S_t) \geq S_t - Ke^{-R(T-t)}$$

This is due to the absence of an arbitrage opportunity. In fact, if we consider a long call we buy the call option at price C_t and at the end we get the corresponding payoff $(S_T - K)^+$. Then we go short today, receiving the corresponding market price S_t and engaging ourselves to re-buy at the end at price $-S_T$. Finally we go long on the riskless asset investing $Ke^{-R(T-t)}$ and getting K at the end.

Action	$t = 0$	$t = T$
Long call	$-C_t$	$(S_T - K)^+$
Sell short	$+S_t$	$-S_T$
Invest in the riskless market	$-Ke^{-R(T-t)}$	$+K$
	$S_t - C_t - Ke^{-RT}$	$(S_T - K)^+ - (S_T - K) \geq 0$

In order to avoid an arbitrage opportunity it must be:

$$S_t - C_t - Ke^{-RT} \leq 0 \quad \Rightarrow \quad C^{AM}(t, S_t) \geq S_t - Ke^{-RT}$$

Of course, if the interest rate is positive – which is not always the case – then

$$C^{AM}(t, S_t) \geq S_t - K \quad \forall t < T$$

Since $C^{AM}(t, S_t) \geq 0$, then

$$C^{AM}(t, S_t) \geq \max(0, S_t - K) = (S_t - K)^+$$

This means that if there are no dividends and the interest rate is positive it is always better to wait because what we get if we decide to exercise today is always less than the value of the American option. In other words, in this case the European option is better, so there will be no differences between American and European options:

$$C^{AM}(t, S_t) = C^{EU}(t, S_t)$$

2.9.2 General discrete time case with no dividends and $R = 0$

Now we consider the general discrete time case with no dividends and $R = 0$. Let's start with some definitions.

Definition 2.9.1 (Stopping time). *A non-negative random variable τ is called (optional) stopping time with respect to the filtration⁴ \mathcal{F} if*

$$\{\tau \leq n\} \in \mathcal{F}_n \quad \forall n \geq 0 \tag{2.96}$$

Since we are considering a discret time framework, this is equivalent to

$$\{\tau \leq n\} \in \mathcal{F}_n \quad \forall n \geq 0 \tag{2.97}$$

So if a random variable is a stopping time we are able to state that the event $\tau \leq n$ is true or false on the basis of the information available at time n . This means that τ is a *non-anticipative* random value.

Definition 2.9.2 (Optimal stopping time). *Given a real number $T < \infty$, consider a family $Z_n \in L^1$ of random variables. If $\hat{\tau}$ is a stopping time such that*

$$\sup_{0 \leq \tau \leq T} \mathbb{E}[Z_\tau] = \mathbb{E}[Z_{\hat{\tau}}] \tag{2.98}$$

then $\hat{\tau}$ is called optimal stopping time.

Now we ask ourselves: does $\hat{\tau}$ exist? Is $\hat{\tau}$ unique? Even if $\hat{\tau}$ exists and is unique, how can we compute it? We try to answer these questions using a preliminary result.

Proposition 2.9.1. *Consider a random variable Z and $T < \infty$.*

⁴ A family $\{G_t : t \geq 0\}$ of sub- σ -algebras is called a *filtration* if $s < t$ implies $G_s \subseteq G_t$.

1. If Z is a sub-martingale, i.e. its expected value increases in time, then $\hat{\tau} = T$.
2. If Z is a super-martingale, i.e. its expected value decreases in time, then $\hat{\tau} = 0$.
3. If Z is a martingale, i.e. its expected value is constant in time, then all times τ are optimal.

Definition 2.9.3 (Value process). Fix $n \leq T$ and consider a stopping time $n \leq \tau \leq T$. The value process is defined as

$$J_n(\tau) := \mathbb{E}[Z_\tau | \mathcal{F}_n] \quad (2.99)$$

and the optimal value process is

$$V_n := \text{ess sup}_{0 \leq \tau \leq T} \mathbb{E}[Z_\tau | \mathcal{F}_n] \quad (2.100)$$

In eq. (2.100) we are taking the essential supremum, which intuitively is the value that is larger or equal than the function values everywhere when allowing for ignoring what the function does at a set of points of measure zero.

Definition 2.9.4. The optimal stopping time at time n is $\hat{\tau}_n$ such that $V_n = V_{\hat{\tau}_n}$.

Example 2.9.1

Consider

$$f(x) = \begin{cases} 5 & x = 1 \\ -4 & x = -1 \\ 2 & \text{otherwise} \end{cases}$$

In this case $\sup f(x) = 5$ and $\inf f(x) = -4$, but notice that they are points of zero measure (with respect to the Lebesgue measure), since the function is almost everywhere equal to 2. So, in this case

$$\text{ess sup } f(x) = \text{ess inf } f(x) = 2$$

Example 2.9.2

Consider

$$f(x) = \begin{cases} \frac{1}{x} & x \in \mathbb{Q} \setminus \{0\} \\ 0 & x = 0 \\ 2 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

In this case $\sup f(x) = \infty$ but $\text{ess sup } f(x) = 0$ since the Lebesgue measure of the set of rational points is zero.

2.9.3 Possible strategies at time n

Lesson 7 of
25/03/2020

We want to study which are the possible strategies to use at time n , dealing with American options. Essentially, we have three possibilities:

1. Use $\hat{\tau}_n$ and get the corresponding optimal value process V_n ;
2. Stop now, getting the corresponding payoff Z_n ;

3. Wait until $n+1$ and adopt the corresponding stopping time $\hat{\tau}_n$, getting nothing at time n and $Z_{\hat{\tau}_{n+1}}$ at time $n+1$. According to the neutral risk methodology, the value of $Z_{\hat{\tau}_{n+1}}$ at time n is given by

$$\mathbb{E}^{\mathbb{Q}}[Z_{\hat{\tau}_{n+1}} \mid \mathcal{F}_n] \stackrel{(a)}{=} \mathbb{E}^{\mathbb{Q}}[\mathbb{E}^{\mathbb{Q}}[Z_{\hat{\tau}_{n+1}} \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n] = \mathbb{E}^{\mathbb{Q}}[V_{n+1} \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n \quad (2.101)$$

where in (a) we used the so called tower property of the conditional expected value⁵. So the market value at time n is given by the conditional expected value up to the current available information \mathcal{F}_n of what will become optimal at time $n+1$

Since option 1 is optimal, it is better than options 2 and 3:

$$\begin{cases} V_n \geq Z_n \\ V_n \geq \mathbb{E}^{\mathbb{Q}}[V_{n+1} \mid \mathcal{F}_n] \end{cases} \quad (2.102)$$

This represents a simple control problem in which we have only two possibilities at time n : stop or wait.

- If we stop it means that it is optimal to stop, so $\hat{\tau}_n = n$ and $V_n = Z_n$;
- If we wait it means that it is not optimal to stop, so $\hat{\tau}_n = \hat{\tau}_{n+1}$ and $V_n = \mathbb{E}^{\mathbb{Q}}[V_{n+1} \mid \mathcal{F}_n]$

According to these possibilities we can characterize the *optimal rule*, which maximizes them.

Theorem 2.9.1.

- i. The optimal value process V solves the backward equation

$$\begin{cases} V_n = \max\{Z_n, \mathbb{E}^{\mathbb{Q}}[V_{n+1} \mid \mathcal{F}_n]\} \\ V_T = Z_T = \text{payoff}_T \end{cases} \quad (2.103)$$

ii. $\hat{\tau}_n = n \Leftrightarrow V_n = Z_n$;

iii. If $\hat{\tau} > n$ then $V_n > Z_n$ and $V_n = \mathbb{E}^{\mathbb{Q}}[V_{n+1} \mid \mathcal{F}_n]$.

Corollary 2.9.1.1.

- i. The optimal stopping rule is given by $\hat{\tau} = \min\{n \geq 0 : V_n = Z_n\}$;
- ii. For n fixed, $\hat{\tau}_n = \min\{k \geq n : V_k = Z_k\}$;

2.9.4 Snell envelope

Let X and Y be two adopted processes and let \mathcal{F}_n , $n \geq 0$, be a given information structure.

Definition 2.9.5 (Domination). X dominates Y if $X_n \geq Y_n$ almost surely for all $n \geq 0$.

Definition 2.9.6 (Snell envelope). If $Y \in L^1$, i.e. if $\mathbb{E}^{\mathbb{Q}}[Y_n] < \infty \forall n \leq T$, the Snell envelope of Y is the smallest super-martingale dominating Y .

⁵ The tower property of the conditional expected value states that if $\mathcal{F}_1 \subset \dots \subset \mathcal{F}_n \subset \mathcal{F}_m \subset \dots \subset \mathcal{F}_N$ is a family of sub- σ -algebras, we have $E[X \mid \mathcal{F}_n] = E[E[X \mid \mathcal{F}_{n+m}] \mid \mathcal{F}_n]$.

So, the Snell envelope is the smallest super-martingale dominating a stochastic process.

Lemma 2.9.2. *Let $\{D^\alpha\}_{\alpha \in A}$ be a family of super-martingales. Then $X_n := \inf_{\alpha \in A} D_n^\alpha$ is a super-martingale.*

Proof.

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[X_{n+1} | \mathcal{F}_n] &= \mathbb{E}^{\mathbb{Q}}[\inf_{\alpha \in A} D_{n+1}^\alpha | \mathcal{F}_n] \\ &\leq \inf_{\alpha \in A} \mathbb{E}^{\mathbb{Q}}[D_{n+1}^\alpha | \mathcal{F}_n] \\ &\leq \inf_{\alpha \in A} D_n^\alpha = X_n\end{aligned}$$

where the last inequality is true because $\{D^\alpha\}_{\alpha \in A}$ is a family of super-martingales. \square

Proposition 2.9.2. *Consider a process $Y \in L^1$. Then there exists a Snell envelope of Y .*

Proof. If we take the family of super-martingales $\{D^\alpha\}_{\alpha \in A}$ dominating Y . Then $\forall n$

$$D_n^\alpha \stackrel{a.s.}{\geq} Y_n$$

Now define $S_n := \inf_{\alpha} D_n^\alpha$, which – according to the lemma (2.9.2) – is a super-martingale and it is minimal. But S_n dominates Y , so S_n is a Snell envelope. \square

Theorem 2.9.3 (Snell envelope theorem). *The optimal value process V is the Snell envelope of the payoff process Z .*

Proof. From (2.102) we have that V_n is a super-martingale. Then we have to show that V_n is minimal. Suppose that X is a super-martingale dominating Z . Being X a super-martingale, we have that

$$V_T = Z_T \Rightarrow X_T \geq Z_T = V_T, \quad X_{n+1} \geq V_{n+1}$$

But for the same reason we also have

$$X_n \geq \mathbb{E}^{\mathbb{Q}}[X_{n+1} | \mathcal{F}_n] \geq \mathbb{E}^{\mathbb{Q}}[V_{n+1} | \mathcal{F}_n] \tag{2.104}$$

X dominates Z , so we have

$$X_n \geq Z_n \tag{2.105}$$

Using (2.104) together with (2.105) we have that

$$X_n \geq \max\{Z_n, \mathbb{E}^{\mathbb{Q}}[V_{n+1} | \mathcal{F}_n]\} \stackrel{def.}{=} V_n$$

So $X_n \geq V_n$, meaning that V is minimal. \square

Remark 2.9.1. If $R \neq 0$ then $V_n e^{-Rn}$, with

$$V_n = \max\{Z_n, e^{-R} \mathbb{E}^{\mathbb{Q}}[V_{n+1} | \mathcal{F}_n]\} \tag{2.106}$$

is a Snell envelope for the process $e^{Rn} Z_n$.⁶

Now let's see a example in which we see that without dividends and positive interest rate it is never optimal to exercise a call option before maturity.

⁶For more insight see [3], page 22.

Example 2.9.3

Consider an American call option with no dividends and interest rate $R \geq 0$.

The payoff function is given by

$$Z_n = e^{-Rn}(S_n - K)^+ = (S_n e^{-Rn} - K e^{-Rn})^+$$

By absence of arbitrage opportunity we have that $S_n e^{-Rn}$ is a \mathbb{Q} -martingale. Then, since $R \geq 0$ and $K = \text{constant}$ we have that $-K e^{-Rn}$ is an increasing term. The sum of a \mathbb{Q} -martingale with an increasing process is a \mathbb{Q} -sub-martingale. But if $S_n e^{-Rn} - K e^{-Rn}$ is a sub-martingale, then also $Z_n = (S_n e^{-Rn} - K e^{-Rn})^+$ is a sub-martingale (the positive part is a convex transformation) and so – according to prop. 2.9.1 – it is never optimal to exercise before T .

2.10 General one period model

We want to provide a general multi-asset model under the condition of absence of arbitrage opportunity in order to understand the relation between completeness and uniqueness of the risk neutral probability measure.

We consider a financial market with N different financial assets. The market only exists at the two points in time $t = 0$ and $t = 1$, and the price per unit of asset at time t will be denoted by S_t . We thus have a price vector process

$$S_t = \begin{pmatrix} S_t^1 \\ \vdots \\ S_t^N \end{pmatrix}$$

The randomness in the system is modeled by assuming that we have a finite sample space

$$\Omega = \{\omega_1, \dots, \omega_M\}$$

and that the probabilities $P(\omega_j)$, $j = 1, \dots, M$, are all strictly positive. The price vector S_0 is assumed to be deterministic and known to us, but the price vector at time $t = 1$ depends upon the outcome $\omega \in \Omega$, and $S_1^i(\omega_j)$ denotes the price per unit of asset at time $t = 1$ if ω_j has occurred. We may therefore define the matrix D by

$$D = \begin{pmatrix} S_1^1(\omega_1) & S_1^1(\omega_2) & \dots & S_1^1(\omega_M) \\ S_1^2(\omega_1) & S_1^2(\omega_2) & \dots & S_1^2(\omega_M) \\ \vdots & \vdots & & \vdots \\ S_1^N(\omega_1) & S_1^N(\omega_2) & \dots & S_1^N(\omega_M) \end{pmatrix} \quad (2.107)$$

Here we assume that the first asset is always positive

$$S_0^1 > 0, \quad S_1^1(\omega_j) > 0 \quad \forall j$$

so that it is possible to discount any asset with respect to the first one.

We now define a portfolio as a N -dimensional row vector $h = (h_1, \dots, h_N)$, where h_i is the number of units of the asset that we buy at time $t = 0$ and keep until time $t = 1$. Since we are buying the assets with deterministic prices at time $t = 0$ and selling them at time $t = 1$ at stochastic prices, the value process of our portfolio will be a stochastic process V , defined by

$$V_0^h = h S_0 = \sum_{i=1}^N h^i S_0^i \quad (2.108)$$

$$V_1^h = hS_1(\omega_j) = \sum_{i=1}^N h^i S_1^i(\omega_j) = (h^T D)_j \quad (2.109)$$

We now define the concept of arbitrage portfolio.

Definition 2.10.1. *The portfolio h is an arbitrage portfolio if it satisfies the conditions:*

- $V_0^h < 0$;
- with probability 1, $V_1^h(\omega_j) \geq 0$ for $i = 1, \dots, M$;
- with probability > 0 , $V_1^h(\omega_j) > 0$ for some j .

Now from the nominal prices S^1, \dots, S^N we normalize in terms of the numéraire⁷ S^1 :

$$Z^i = \frac{S^i}{S^1} \quad (2.110)$$

that is

$$Z_0^i = \frac{S_0^i}{S_0^1}, \quad Z_1^i(\omega_j) = \frac{S_1^i(\omega_j)}{S_1^1(\omega_j)}$$

The corresponding normalized matrix is

$$D^Z = \begin{pmatrix} 1 & 1 & \dots & 1 \\ Z_1^2(\omega_1) & Z_1^2(\omega_2) & \dots & Z_1^2(\omega_M) \\ \vdots & \vdots & & \vdots \\ Z_1^N(\omega_1) & Z_1^N(\omega_2) & \dots & Z_1^N(\omega_M) \end{pmatrix} \quad (2.111)$$

The economic interpretation of the fact that $Z^1 = 1$ is that in the normalized price system, the numéraire asset is risk free corresponding to a bank with zero interest rate. This is the main reason why the normalized prices system is so much easier to analyze than the nominal one. The normalized portfolio will be:

$$V_t^{h,Z} = \frac{V_t^h}{S_t^1}, \quad t = 0, 1 \quad (2.112)$$

2.10.1 Absence of arbitrage

Lesson 8 of
26/03/2020

We now go on to investigate when the market model above is free of arbitrage possibilities, and the main technical tool for this investigation is the Farkas' Lemma.

Lemma 2.10.1 (Farkas' Lemma). *Suppose that x_0, x_1, \dots, x_K are column vectors in \mathbb{R}^N . Then exactly one of the two following problems possesses a solution:*

- P1: *Find non-negative numbers $\lambda_1, \dots, \lambda_M$ such that*

$$x_0 = \sum_{j=1}^K \lambda_j x_j \quad (2.113)$$

- P2: *Find a row vector $h \in \mathbb{R}^N$ such that*

$$h^T x_0 < 0, \quad h^T x_j \geq 0 \quad \forall j = 1, \dots, K \quad (2.114)$$

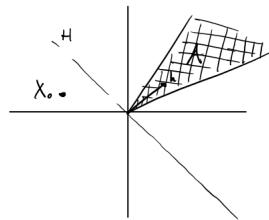
⁷A numéraire is a tradable economic entity in terms of whose price the relative prices of all other tradables are expressed.

Proof. Let A be the set of all non-negative linear combinations of x_1, \dots, x_K :

$$A = \{y \in \mathbb{R}^n : y = \lambda_1 x_1 + \dots + \lambda_K x_K; \lambda_i > 0 \ \forall i\}$$

It is easy to see that A is a closed convex cone containing the origin. Exactly one of the following cases can hold:

- $x_0 \in A$. This means that P1 above has a solution.
- $x_0 \notin A$. Then, by the separation theorem for convex sets (Hahn-Banach theorem), there exists a hyperplane H such that x_0 is strictly on one side of H whereas A is on the other side. Letting h be defined as a normal vector to H pointing in the direction where A lies, this means that P2 has a solution.



□

We can now formulate our first result.

Theorem 2.10.2. *The market is arbitrage free if and only if there exist non-negative normalized numbers q_1, \dots, q_M , $q_1 + \dots + q_M = 1$, such that the following vector equality holds:*

$$Z_0 = \sum_{i=1}^M Z_1(\omega_i) q_i = Z_1(\omega_1) q_1 + \dots + Z_1(\omega_M) q_M \quad (2.115)$$

Proof. From the definition of arbitrage it is clear that the market is arbitrage free if and only if the following system of equations has no solution:

$$\begin{aligned} V_0^{h,Z} &= h^T Z_0 = 0, \\ V_1^{h,Z} &= h^T Z_1(\omega_j) = (h^T D^Z)_j \geq 0, \quad \forall j \\ V_1^{h,Z} &> 0, \quad \text{for some } j \end{aligned}$$

We now want to apply Farkas' Lemma to this system and in order to do this we define the column vector $p = (p_1, \dots, p_M)^T \in \mathbb{R}^M$. We can now rewrite the system above as:

$$\begin{aligned} h^T Z_0 &= 0, \\ h^T D^Z &\geq 0, \\ (h^T D^Z) \cdot p &> 0, \end{aligned}$$

where the second inequality is interpreted component wise. (We remark that we could in fact have replaced the vector p by any vector in \mathbb{R}^M with strictly positive components.) We finally rewrite the first equality as a double inequality to obtain

$$h^T Z_0 \geq 0,$$

$$\begin{aligned} h^T(-Z_0) &\geq 0, \\ h^T D^Z &\geq 0, \\ h^T(-D^Z p) &< 0. \end{aligned}$$

The point of all this is that if we define the vector d_0 and the block matrix \hat{D} by

$$d_0 = (-D^Z p) \in \mathbb{R}^N, \quad \hat{D} = [D^Z, Z_0, -Z_0]_{N \times (M+2)}$$

we see that the market is free of arbitrage if and only if the system

$$\begin{aligned} h^T \hat{D} &\geq 0, \\ h^T d_0 &< 0, \end{aligned}$$

has no solution $h \in \mathbb{R}^N$, where the last inequality is interpreted component wise. Applying Farkas' Lemma to this system we thus see that absence of arbitrage is equivalent to the existence of non-negative real numbers $\lambda_1, \lambda_2, \dots, \lambda_M, \lambda_{N+1}, \lambda_{N+2}$ such that (with obvious notation)

$$d_0 = \hat{D} \cdot \lambda = -D^Z p = (D^Z, Z_0, -Z_0) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_M \\ \lambda_{N+1} \\ \lambda_{N+2} \end{pmatrix}$$

If we now define the vector $\beta = (\lambda_1, \dots, \lambda_M)$ and the real number $\alpha = \lambda_{N+2} - \lambda_{N+1}$ we can write the previous relation as

$$-D^Z p = D^Z \beta + \alpha Z_0, \quad \Rightarrow \quad \alpha Z_0 = D^Z(p + \beta) \quad (*)$$

Here $\beta \geq 0$ but we do not know the sign of α . If we focus on the first component of the vector equality and recall that $Z_0^1 = 1$ and that the first row of D^Z has the unit 1 in all positions, we obtain

$$\alpha = \sum_{j=1}^M (p_j + \beta_j).$$

From this we see that $\alpha > 0$, and if we define the column vector $q \in \mathbb{R}^M$ by

$$q = \frac{p + \beta}{\alpha}$$

we see that

$$q_j > 0, \quad j = 1, \dots, M \quad \sum_{j=1}^M q_j = 1.$$

So, rewriting $(*)$ as

$$Z_0 = D^Z q$$

gives us (2.115). □

Eq. (2.115) can be written as:

$$\begin{pmatrix} 1 \\ Z_0^1 \\ \vdots \\ Z_0^N \end{pmatrix} = \mathbb{E}^{\mathbb{Q}}[Z_1] \quad (2.116)$$

So the market is arbitrage free if and only if there exists a probability distribution Q such that the initial value of the assets can be written in terms of the risk neutral expectation of the terminal value of the assets, once normalized.

We can restate theorem 2.10.2 as the following result, which in its far reaching generalizations is known as the first fundamental theorem of mathematical finance.

Theorem 2.10.3 (First Fundamental Theorem). *Given a fixed numéraire, the market is free of arbitrage possibilities if and only if there exists a martingale measure \mathbb{Q} , i.e. a probability measure such that:*

1. $Q \sim \mathbb{P}$, i.e. Q is equivalent to \mathbb{P} ($Q(A) = 0 \Leftrightarrow \mathbb{P}(A) = 0, \forall A$);
2. $Z_0 = \mathbb{E}^{\mathbb{Q}}[Z_1]$.

Remark 2.10.1. Consider the special case in which S^1 is a risk-free asset. Using the usual convention we have that $S_1 = e^{RT}$ and then

$$S_0^i = e^{-RT} \mathbb{E}^{\mathbb{Q}}[S_1^i], \quad i = 1, \dots, N$$

so we can recover the initial value of the assets as discounted values of the corresponding risk neutral expectations.

2.10.2 Market completeness

Recall that the market is complete if for any random variable $X \in (\Omega, \mathcal{F}, \mathbb{P})$ there exists a portfolio V^h such that $\mathbb{P}(V_1^h = X) = 1$.

Proposition 2.10.1. *The market is complete if and only if $\text{rank}(D) = M$.*

Proof. For any portfolio h , we view the random variable V_1^h as a row vector $V_1^h = (V_1^h(\omega_1), \dots, V_1^h(\omega_M))$ and with this notation we have $V_1^h = h^T D$. The market is thus complete if and only if, for every random variable X (viewed as a row vector in \mathbb{R}^M) there exists the corresponding portfolio, i.e. if there exists h such that the equation $h^T D = X$ has a solution. But

$$h^T D = h^1(S^1) + \dots + h^N(S^N)$$

is exactly a linear combination of the rows of D with the components of h as coefficients. \square

Remark 2.10.2. If X is obtainable (it is possible to span X) and S^1 is risk-free then

$$\text{price}_0(X) = e^{-RT} \mathbb{E}^{\mathbb{Q}}[X].$$

In fact, there exist h such that $X = h^T D$ and from previous results we know that $S_0 = e^{-RT} \mathbb{E}^{\mathbb{Q}}[S_1]$. But we also know that

$$X = h^T D = \sum_i h^i S_1^i,$$

where $S_1^i = \text{price}_0(S^1) = e^{-RT} \mathbb{E}^{\mathbb{Q}}[S^i] = S_0^i$. So

$$\text{price}_0(X) = \sum_i h^i S_0^i = V_0^h.$$

In other words, if X is obtainable we can write the corresponding replicating portfolio. Remember that we cannot speak about pricing without speaking about hedging: here we have an implicit hedging procedure.

In Proposition 2.10.1 we obtained one characterization of complete markets. There is another characterization which connects completeness to martingale measures. This result is known as the second fundamental theorem of mathematical finance.

Theorem 2.10.4 (Second Fundamental Theorem). *Assume that the model is arbitrage free. Then the market is complete if and only if the martingale measure is unique.*

Proof. From Proposition 2.10.1 we know that the market is complete if and only if $\text{rank}(D) = M$, i.e. if and only if

$$\text{Im}\{D^T\} = \mathbb{R}^M,$$

where we view the transpose matrix D^T as a mapping from \mathbb{R}^N to \mathbb{R}^M . On the other hand, from Theorem 2.10.2 and the assumption of absence of arbitrage we know that there exists a solution (even a strictly positive one) to the equation

$$Z_0 = D^Z q.$$

This solution is unique if and only if the kernel of D^Z is trivial, i.e. if and only if

$$\ker(D^Z) = 0,$$

and it is easy to see that this is equivalent to the condition

$$\ker(D) = 0.$$

We now recall the following well known duality result:

$$(\text{Im}\{D^*\})^\perp = \ker(D).$$

Thus $\ker(D) = 0$ if and only if $\text{Im}(D^*) = \mathbb{R}^M$, i.e. the market is complete if and only if the martingale measure is unique. \square

In the case the market is not complete but still arbitrage free, in order to price something we have to consider different probability measures, so the price is not uniquely determined. Instead of numbers we will have intervals, which can be very large and thus not useful to characterize the market value of something.

In conclusion, if S^1 is risk free we have that:

- There is no arbitrage opportunity if and only if there exists a martingale measure \mathbb{Q} under which discounted assets are constant in expected value;
- The market is complete if and only if \mathbb{Q} is unique;
- For any random variable X

$$\text{price}_0(X) = e^{-RT} \mathbb{E}^{\mathbb{Q}}[X].$$

However, \mathbb{Q} may be not unique and so $\text{price}_0(X)$ may be an interval.

- If the market is incomplete – i.e. if it is not possible to reach all the contingent claims uniquely with a replicating portfolio – prices are intervals and the price coherent with the no arbitrage principle is the interval

$$\text{price}_0(X) = \left[\inf_Q e^{-RT} \mathbb{E}^{\mathbb{Q}}[X], \sup_Q e^{-RT} \mathbb{E}^{\mathbb{Q}}[X] \right]$$

If we want a number instead of an interval we have to add subjective information (e.g. utility functions, risk aversion);

- If X is replicable, the price of X does not depend on the particular \mathbb{Q} , even in an incomplete market.

Example 2.10.1: American put option pricing

Consider a toy model with $u = 1.1$, $d = 0.9$ and $S_0 = K = 100$:



Assume that $R = 1\%$ per year and that $T = 2$ years. The payoff of the corresponding European put option is $(K - S)^+ = (0, 1, 19)$. The tree is stationary, so the risk neutral probability weight is given by

$$q = \frac{e^{RT} - d}{u - d} = \frac{e^{0.01 \cdot 1} - 0.9}{1.1 - 0.9}$$

The price in node (A) is given by

$$\begin{aligned} P_{(A)}^{AM} &= \max \left\{ (100 - 110)^+, e^{-0.01} \mathbb{E}^{\mathbb{Q}} \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] \right\} \\ &= \max \{0, e^{-0.01} (0 \cdot q + 1 \cdot (1 - q))\} \\ &= e^{-0.01} (1 - q) \end{aligned}$$

So it is better to wait when we arrive in node (A). If we are in node (B) we have:

$$\begin{aligned} P_{(B)}^{AM} &= \max \{(100 - 90)^+, e^{-0.01} (0 \cdot q + 1 \cdot (1 - q))\} \\ &= 10 \end{aligned}$$

So in node (B) it is better to exercise the option. Finally, the initial value of the American put option is:

$$\begin{aligned} P_0^{AM} &= \max \{(100 - 100)^+, e^{-0.01} (P_{(A)}^{AM} \cdot q + P_{(B)}^{AM} \cdot (1 - q))\} \\ &= e^{-0.01} [e^{-0.01} q (1 - q) + 10 (1 - q)] = \text{conti} \end{aligned}$$

If we compute the initial price of the European option we will find that $P_0^{AM} > P_0^{EU}$. This makes sense because for the American option it was convenient to exercise before the maturity (at least for one state of the nature), so we can really exploit this possibility.

In terms of hedging, we only have to understand which are $\Delta_{0,(A)}$ and $\beta_{0,(A)}$. Since in node (B) the underlying decreases, there is no need to consider hedging in node (B).

3 | Continuous time finance

3.1 Stochastic integrals

Lesson 9 of
27/03/2020

The starting point for finance in continuous time is a probability space

$$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}),$$

where $\{\mathcal{F}_t\}_{t \geq 0}$ is an increasing family of complete σ -algebras. Here, in order to avoid measurability problems, we assume also that $\{\mathcal{F}_t\}_{t \geq 0}$ has right continuity:

$$\bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon} = \mathcal{F}_{t+} = \mathcal{F}_t$$

This means that the information available immediately after t coincides with the one at time t , so there can be shocks only in the left side of the limit, which can be discontinuous. This is a quite heavy assumption, because it involves the future, but it is compulsory. In this probability space we define the underlying stochastic process X_t and we assume measurability with respect to the probability infrastructure and regularity in terms of the trajectory. In particular, we assume that X_t is a *RCLL* function, also called *càdlàg*, i.e. a function defined on the real numbers (or a subset of them) that is right-continuous everywhere and has left limits everywhere.

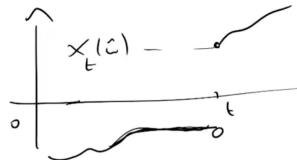


Figure 3.1: Càdlàg.

Processes like this define a *Brownian motion* W_t , i.e. a gaussian process with distribution at fixed time \hat{t} $W_{\hat{t}} \sim \mathcal{N}(0, \hat{t}) \sim \sqrt{\hat{t}} \mathcal{N}(0, 1)$. Then we can define the stochastic integral

$$\int_0^t H_s "dW_s"$$

where $H \in \mathcal{H}$, being \mathcal{H} the class of adapted processes such that $\mathbb{E}[\int_0^t H_s^2 ds] < \infty$. Since the Brownian motion is not differentiable it is not possible to define a differential. However, if we are able to define the integral with respect to the Brownian motion then we can define the differential operator as the derivative of the integral:

$$d \left(\int_0^t H_s dW_s \right) \rightarrow H_t "dW_t"$$

Example 3.1.1

Consider a process given by the deterministic integral

$$X_t = \int_0^t W_s ds$$

where W_s is a Brownian motion. We want to show that X_t behaves like a gaussian random variable $\mathcal{N}(\mu, \sigma)$ where μ, σ have to be determined. To solve the integral we can use *Itô's Lemma*:

$$d(tW_t) = W dt + t dW_t + \underbrace{d\langle \cdot, W \rangle}_0$$

where t is deterministic and the last term – the quadratic covariance – can be neglected because it leads to higher order terms in dt (recall that $dtdW = dt\sqrt{dt}$). Integrating by parts we get:

$$tW_t = \int_0^t W_s ds + \int_0^t s dW_s$$

So we can say that the integral

$$\int_0^t W_s ds = tW_t - \int_0^t s dW_s$$

is Gaussian, being the sum of two gaussian terms.

Alternatively, recall that it is always possible to consider the Brownian motion as the integral of its differential, in fact we can define the integral operator to be compatible with the differential one:

$$W_t = \int_0^t dW_s = W_t - W_0$$

So, we can write:

$$\begin{aligned} \int_0^t W_s ds &= \int_0^t \left(\int_0^s dW_u \right) ds \stackrel{(a)}{=} \int_0^t \left(\int_u^t ds \right) dW_u \\ &= \int_0^t (t-u) dW_u = tW_t - \int_0^t dW_u \end{aligned}$$

where in (a) we used the Stochastic Fubini Theorem to change the order of integration.

Now we look for μ and σ . We have that

$$\mathbb{E}[X_t] = \mathbb{E} \left[\int_0^t W_s ds \right] \stackrel{(a)}{=} \int_0^t \mathbb{E}[W_s] ds \stackrel{(b)}{=} 0$$

where in (a) we moved the expectation value inside the integral (we can do it because it is a deterministic integral) and in (b) we used the fact that by definition $\mathbb{E}[W_s] = 0$. Then:

$$\begin{aligned} \text{Var}(X_t) &= \mathbb{E} \left[\left(\int_0^t W_s ds \right)^2 \right] = \mathbb{E} \left[\int_0^t \int_0^t W_s W_u ds du \right] \\ &= \int_0^t \int_0^t \mathbb{E}[W_s W_u] ds du = \int_0^t \int_0^t \text{Cov}(W_s, W_u) ds du \end{aligned}$$

The increments of the Brownian motion are independent but the brownian motion itself is a strongly autocorrelated process, with covariance $\text{Cov}(W_t, W_u) = \min\{s, u\}$. So we can write:

$$\begin{aligned} &= \int_0^t \left(\int_0^u s \, ds + \int_u^t u \, ds \right) du = \int_0^t \left(\frac{u^2}{2} + u(t-u) \right) du \\ &= \frac{t^3}{6} + \frac{t^3}{2} - \frac{t^3}{3} = \frac{t^3}{3} \end{aligned}$$

So $X_t \sim \mathcal{N}(0, t^3/3)$.

Warning! In this example we are able to do the computations – in particular we can switch expected value and integration – because we are dealing with deterministic integrals. If we have a stochastic integral, e.g. $\int_0^t H_s dW_s$, we cannot move the expected value inside the integral because we get something which is random:

$$\underbrace{\mathbb{E} \left[\int_0^t H_s dW_s \right]}_{\text{deterministic}} \neq \underbrace{\int_0^t \mathbb{E}[H_s] dW_s}_{\text{stochastic}}.$$

An important property is that if $H_s \in \mathcal{H}$, then the stochastic process

$$\left(\int_0^t H_s dW_s \right)_{t \geq 0}$$

is a martingale:

$$\mathbb{E} \left[\int_0^t H_s dW_s \right] = \int_0^0 H_s dW_s = 0. \quad (3.1)$$

Luckily, \mathcal{H} is a very rich class which contains also Brownian motions. In fact:

$$\mathbb{E} \left[\int_0^t W_s^2 ds \right] = \int_0^t \mathbb{E}[W_s^2] ds = \int_0^t \text{Var}(W_s) ds = \int_0^t s ds = \frac{t^2}{2} < \infty.$$

We can extend the notion of stochastic integral even when $H_s \notin \mathcal{H}$, requiring a weaker condition, i.e.

$$H_s \in \tilde{\mathcal{H}} = \left\{ H \text{ adapted} : \mathbb{P} \left(\int_0^t H_s^2 ds < \infty \right) = 1 \right\},$$

but in this case X_t is a *local martingale*, meaning that it is a martingale only for a specific set of stopping times. This means, in principle, that $\mathbb{E}[X_t] \neq 0$.

Now, consider a Geometric Brownian Motion (GBM)

$$\begin{cases} dX_t = \alpha X_t dt + \sigma X_t dW_t \\ X_0 = x_0, \end{cases} \quad (3.2)$$

which also can be written as

$$X_t - x_0 = \alpha \int_0^t X_s ds + \sigma \int_0^t X_s dW_s. \quad (3.3)$$

Since we are dealing with stochastic integrals, trivial integration is not possible and we have to find an other way to compute them.

First, let's consider the case in which $\sigma = 0$. We get a linear deterministic equation:

$$dX_t = \alpha X_t dt \quad (3.4)$$

which can be written as

$$\dot{X}_t = \frac{dX_t}{dt} = \alpha X_t \quad (3.5)$$

and has solution

$$X_t = x_0 e^{\alpha t}. \quad (3.6)$$

If $\sigma \neq 0$ it is as we introduce a white noise which disturbs the evolution of the deterministic solution. The naive approach is to write the solution as

$$X_t = "x_0 e^{\alpha t + \sigma W_t}" \quad (3.7)$$

This is not correct, because it means that we are considering the exponential in a deterministic way, which is not the case. We have to use Itô's Formula.

Theorem 3.1.1 (Itô's Formula). *Assume that the process X has a stochastic differential given by*

$$\begin{cases} dX_t = K_t dt + H_t dW_t \\ X_0 = x_0 \end{cases} \quad (3.8)$$

where α and σ are adapted processes ($\alpha, \sigma \in \tilde{\mathcal{H}}$), and let f be a $C^{1,2}$ function. Define the process $Z(t) = f(t, X(t))$. Then Z has a stochastic differential given by

$$df = \left(\frac{\partial f}{\partial t} + K_t \frac{\partial f}{\partial x} + \frac{H_t^2}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + H_t \frac{\partial f}{\partial x} dW_t. \quad (3.9)$$

In our case, we take f as the function which simplifies the deterministic part, which is an exponential. So, we take:

$$f(t, X_t) = \ln X_t. \quad (3.10)$$

According to (3.9), the differential is:

$$\begin{aligned} d \ln X_t &= \frac{1}{X_t} dX_t + \frac{1}{2} \left(-\frac{1}{X_t^2} \right) d \langle X_t \rangle_t \\ &= \alpha dt + \sigma dW_t - \frac{1}{2X_t^2} (\sigma^2 X_t^2 dt) \\ &= \left(\alpha - \frac{1}{2}\sigma^2 \right) dt + \sigma dW_t \end{aligned} \quad (3.11)$$

and integrating both sides we get:

$$\ln X_t - \ln X_0 = \left(\alpha - \frac{1}{2}\sigma^2 \right) t + \sigma(W_t - W_0) = \left(\alpha - \frac{1}{2}\sigma^2 \right) t + \sigma W_t \quad (3.12)$$

By taking the exponential of both sides we get:

$$X_t = x_0 e^{(\alpha - \frac{1}{2}\sigma^2)t + \sigma W_t} \quad (3.13)$$

Notice that this solution is log-normal:

$$X_t \sim e^{\mathcal{N}(\ln x_0 + (\alpha - \frac{1}{2}\sigma^2)t, \sigma^2 t)}$$

and that it fluctuates around the corresponding deterministic solution (3.6). We can demonstrate that averaging (3.13) we get exactly (3.6), in fact – according to the moment generating function formula $M(t) = \mathbb{E}[e^{tX}] = e^{\mu t} e^{\frac{1}{2}\sigma^2 t^2}$ – the expected value of $e^{-\frac{1}{2}\sigma^2 t + \sigma W_t}$ is:

$$\mathbb{E} \left[e^{-\frac{1}{2}\sigma^2 t + \sigma W_t} \right] = \mathbb{E} \left[e^{\mathcal{N}(-\frac{1}{2}\sigma^2 t, \sigma^2 t)} \right] = e^{-\frac{1}{2}\sigma^2 t + \frac{1}{2}\sigma^2 t} = 1$$

Alternatively, one may recognize that $e^{-\frac{1}{2}\sigma^2 t + \sigma W_t}$ is a famous exponential martingale Y_t , and so $\mathbb{E}[Y_t] = Y_0 = e^0 = 1$.

3.1.1 Black & Scholes market

In the early 1970s, Fischer Black, Myron Scholes and Robert Merton achieved a major breakthrough in the pricing of European stock options. This was the development of what has become known as the Black–Scholes–Merton (Black & Scholes) model. This model describes the asset price S_t in terms of the infinitesimal return $\frac{dS_t}{S_t}$, assuming that percentage changes in the stock price in a very short period of time are normally distributed. If we define a drift μ as the *expected return of a stock per year* and σ as the *volatility of the stock price per year* (a measure of our uncertainty about the returns provided by the stock), we have that

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad (3.14)$$

where dW_t represents the noise given by this variability. The mean and standard deviation of the return in time dt are approximately μdt and $\sigma\sqrt{dt}$, so that

$$\frac{dS_t}{S_t} \sim \mathcal{N}(\mu dt, \sigma^2 dt) \quad (3.15)$$

Recalling what we saw in the previous section, this implies that the return $Y_t = \ln S_t$ is normally distributed:

$$\ln S_t \sim \mathcal{N}\left(\ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right) \quad (3.16)$$

and so S_t is log-normal distributed.



Figure 3.2: Log-normal distribution

While it is possible to estimate the historical volatility of the market, it is not possible to estimate the historical drift because it is not observable and also not stable. In fact, according to the particular time window we are considering, μ can assume different values.

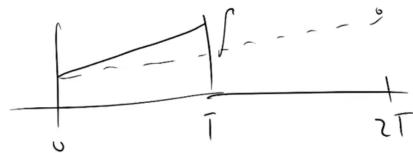


Figure 3.3: Different values of μ .

In conclusion, we cannot jointly estimate μ and σ from only one time series and we would like to get results which are independent of the drift (which plays the same role as the subjective probability in the static binomial model).

3.1.2 The Girsanov Theorem

We want to find a way to switch from the subjective probability measure \mathbb{P} to \mathbb{Q} in continuous time. The tool that allows us to do that is the Girsanov Theorem.

Consider a Brownian motion $W_t^{\mathbb{P}}$ under the probability measure $(\Omega, \mathcal{F}, \mathbb{P})$ and a *finite variation process* $\int_0^t H_s ds$ (its differential has not a diffusion part). Then

$$W_t^{\mathbb{Q}} := W_t^{\mathbb{P}} - \int_0^t H_s ds \quad (3.17)$$

is not a Brownian motion under the measure \mathbb{P} , because it is not a martingale. However, it is a Brownian motion under the measure \mathbb{Q} , where

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_t = \exp \left(\int_0^t H_s dW_s^{\mathbb{P}} - \frac{1}{2} \int_0^t H_s^2 ds \right) := Z_t \quad (3.18)$$

is a process that satisfies the stochastic differential equation (SDE)

$$dZ_t = H_t Z_t dW_t^{\mathbb{P}} \quad (3.19)$$

and can be written as a constant plus a stochastic integral, which at least is a local martingale:

$$Z_t = Z_0 + \int_0^t H_s Z_s dW_s^{\mathbb{P}} \quad (3.20)$$

Eq. (3.18) is called *Radon-Nikodym derivative*: the choice of notation and the name of the function reflects the fact that it is analogous to a derivative in calculus, in the sense that it describes the rate of change of density of one measure with respect to another¹.

One possible issue is that in the change from \mathbb{P} to \mathbb{Q} there can be a loss of probability mass, ending up with a \mathbb{Q} which is not normalized to 1. In order to avoid that, $\int_0^t H_s Z_s dW_s^{\mathbb{P}}$ must be a global martingale. This is the assumption of Girsanov Theorem.

Theorem 3.1.2 (Girsanov Theorem). *Consider a Brownian motion under a probability measure \mathbb{P} . Then, a candidate Brownian motion under a different probability measure \mathbb{Q} is a true \mathbb{Q} -Brownian motion if the change of measure characterized by Z_t is a true martingale.*

One consequence of this theorem is that changing measure is equivalent to shift the Brownian motion by a finite variation process. If we want to move from a probability measure to another, the effect on the SDE involves the drift, not the volatility.

How can we check the true martingality condition on Z_t ? We know that if $H_s Z_s \in \mathcal{H}$ then $\int H_s Z_s dW_s$ is a martingale, but we cannot use this information because it involves Z_s itself, which is unknown. In order to do that we use the *Novikov condition*.

Lemma 3.1.3 (Novikov condition). *If*

$$\mathbb{E} \left[e^{\frac{1}{2} \int_0^t H_s^2 ds} \right] < \infty \quad (3.21)$$

then Z_t is a true martingale.

¹See https://en.wikipedia.org/wiki/Radon%20%93Nikodym_theorem

Summary. In conclusion, suppose that X is a random variable and we want to compute

$$\mathbb{E}[X] = \int_{\Omega} X \, d\mathbb{P} = \int_{\Omega} X(\omega) \mathbb{P}(d\omega).$$

If we don't know the distribution of X under the probability measure \mathbb{P} but we know it under an other probability measure \mathbb{Q} , it is possible to switch from \mathbb{P} to \mathbb{Q} using the change of measure

$$\mathbb{E}^{\mathbb{P}}[X] = \int X \frac{d\mathbb{P}}{d\mathbb{Q}} \, d\mathbb{Q} = \mathbb{E}^{\mathbb{Q}} \left[X \frac{d\mathbb{P}}{d\mathbb{Q}} \right]$$

This holds for all X . In particular, if we take $X = 1$ then

$$1 = \mathbb{E}^{\mathbb{P}}[1] = \mathbb{E}^{\mathbb{Q}} \left[\frac{d\mathbb{P}}{d\mathbb{Q}} \right]$$

So, in order to avoid loss of probability mass, $\frac{d\mathbb{P}}{d\mathbb{Q}}$ must be a true martingale.

Lesson 10 of
01/04/2020

3.2 Portfolio dynamics

Let us consider a financial market consisting of different assets such as stocks, bonds with different maturities, or various kinds of financial derivatives. In this section we will take the price dynamics of the various assets as given, and the main objective is that of deriving the dynamics of (the value of) a so-called *self financing portfolio*. In continuous time this turns out to be a fairly delicate task, so we start by studying a model in discrete time. We will then let the length of the time step tend to zero, thus obtaining the continuous time analogs.

Let us thus study a financial market, where time is divided into periods of length Δt , and where trading only takes place at the discrete points in time $n\Delta t$, $n = 0, 1, \dots$. We consider a fixed period $[t, t + \Delta t)$. This period is henceforth referred to as “period t ”.

Definition 3.2.1. Let's define the following notation:

- $N =$ the number of different types of stocks.
- $h_i(t) =$ number of shares of type i held during the period $[t, t + \Delta t)$.
- $h(t) =$ the portfolio $[h_1(t), \dots, h_N(t)]$ held during period t .
- $c(t) =$ the amount of money spent on consumption per unit time during the period $[t, t + \Delta t)$.
- $S_i(t) =$ the price of one share of type i during the period $[t, t + \Delta t)$.
- $V(t) =$ the value of the portfolio h at time t .

The information and the decisions in the model are structured as follows:

- At time t , i.e. at the start of period t , we bring with us an “old” portfolio $h(t - \Delta t) = \{h_i(t - \Delta t), i = 1, \dots, N\}$ from the previous period $t - \Delta t$.
- At time t we can observe the price vector $S(t) = (S_1(t), \dots, S_N(t))$.
- At time t , after having observed $S(t)$, we choose a new portfolio $h(t)$, to be held during period t . At the same time we also choose the consumption rate $c(t)$ for the period t . Both $h(t)$ and $c(t)$ are assumed to be constant over the period t .

Remark 3.2.1. We only consider non-dividend paying assets.

To start the analysis we observe that the wealth at the start of period t , $V(t)$, equals the value of the old portfolio $h(t - \Delta t)$. Thus we have:

$$V(t) = \sum_{i=1}^N h_i(t - \Delta t)S_i(t) = h(t - \Delta t)S(t). \quad (3.22)$$

This equation simply says that at the beginning of period t our wealth equals what we get if we sell our old portfolio at today's prices. We may now use the proceeds of this sale for two purposes:

1. Reinvest in a new portfolio $h(t)$.
2. Consume at the rate $c(t)$ over the period t

The cost of the new portfolio $h(t)$, which has to be bought at today's prices, is given by

$$\sum_{i=1}^N h_i(t)S_i(t) = h(t)S(t), \quad (3.23)$$

whereas the cost for the consumption rate $c(t)$ is given by $c(t)\Delta t$. The *budget equation* for period t thus reads

$$h(t - \Delta t)S(t) = h(t)S(t) + c(t)\Delta t, \quad (3.24)$$

$$(h(t) - h(t - \Delta t))S(t) + c(t)\Delta t = 0, \\ \Delta h(t)S(t) + c(t)\Delta t = 0. \quad (3.25)$$

Since our goal is to obtain the budget equation in continuous time, the naive approach is to let $\Delta t \rightarrow 0$ in eq. (3.25), obtaining the formal expression

$$S(t)dh(t) + c(t)dt = 0. \quad (3.26)$$

This procedure is, however, **not correct**. In fact, all the stochastic differentials are to be interpreted in the Itô sense and the Itô integral $\int H(t)dW(t)$ was defined as the limit of sums of the type $\sum H(t_n)[W(t_{n+1}) - W(t_n)]$ where it was essential that the W -increments were forward differences. But in eq. (3.25) we have a backward h -difference.

In order to get Itô differentials we thus have to reformulate eq. (3.25). This is done by adding and subtracting the term $S(t - \Delta t)\Delta h(t)$ to the left-hand side:

$$\Delta h(t)S(t) + c(t)\Delta t + S(t - \Delta t)\Delta h(t) - S(t - \Delta t)\Delta h(t) = 0 \\ S(t - \Delta t)\Delta h(t) + \Delta S(t)\Delta h(t) + c(t)\Delta t = 0 \quad (3.27)$$

Now, at last, we may let $\Delta t \rightarrow 0$ in the budget eq. (3.27), giving us

$$S(t)dh(t) + dh(t)dS(t) + c(t)dt = 0. \quad (3.28)$$

Letting $\Delta t \rightarrow 0$ in eq. (3.22) gives us

$$V(t) = h(t)S(t), \quad (3.29)$$

and if we take the Itô differential of this expression we get

$$dV(t) = h(t)dS(t) + S(t)dh(t) + dS(t)dh(t)$$

$$\begin{aligned} &\stackrel{(3.28)}{=} h(t)dS(t) + S(t)dh(t) - c(t)dt - S(t)dh(t) \\ &= S(t)dh(t) - c(t)dt. \end{aligned} \quad (3.30)$$

In particular, in a situation without any consumption we have the following V-dynamics:

$$dV(t) = h(t)dS(t). \quad (3.31)$$

This is the so called *self-financing portfolio*, i.e. a portfolio with no exogenous infusion or withdrawal of money. In other words, the purchase of a new portfolio, as well as all consumption, must be financed solely by selling assets already in the portfolio.

The natural financial interpretation of eq. (3.31) is that in a model without any exogenous income, all change of wealth is due to changes in asset prices. In other words, the fact that the portfolio fluctuates does not depend on our strategy h but only on the fact that in a infinitesimal interval there has been a fluctuation in the market due to a shock in the risky asset.

3.3 Option pricing in the Black & Scholes model

Let us consider a financial market consisting of only two assets:

1. A savings account (risk free asset) with price process B that evolves according to the deterministic equation

$$\frac{dB_t}{B_t} = rdt; \quad (3.32)$$

2. A stock with price process (risky asset) S , evolving according to the Brownian motion SDE

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \quad (3.33)$$

where $\sigma \neq 0$, otherwise if $\mu \neq r$ there is an arbitrage opportunity.

In the Black & Scholes model μ, σ and r are constant. We make the following assumptions:

- The derivative can be bought or sold without any restriction;
- There is no arbitrage opportunity;
- The price can be written as a function of the underlying;

$$price_t = p(t) = F(t, S(t)) \in C^{1,2}. \quad (3.34)$$

This is a Markovian-like assumption, which is quite strong because it says that the price depends only on the current value of the underlying $S(t)$, which is not true in general.

The idea is to repeat the same argument we used in the discrete time framework. So we want to construct the replicating portfolio of the payoff at time T and it must be self-financing, otherwise the initial value of the option may be different from the value of our portfolio if in the meanwhile we receive something else. We write the value of the replicating portfolio at time t as

$$V(t) = \alpha(t)S(t) + \beta(t)B(t) = F(t, S(t)). \quad (3.35)$$

At maturity we have

$$V(T) = \text{payoff}_T. \quad (3.36)$$

Then, by absence of arbitrage opportunity it must be that

$$V(0) = \text{price}_0(\text{payoff}_T) = p(0) = F(0, S(0)) \quad (3.37)$$

and this must hold also for all the intermediate times

$$V(t) = F(0, S(0)). \quad (3.38)$$

So our constructed portfolio and the option have the same value at any time. This means that also the fluctuations, i.e. their differentials, are the same:

$$dV(t) = dF(t, S(t)). \quad (3.39)$$

According to (3.31) we have that

$$\begin{aligned} dV(t) &= h(t)dS(t) = (\alpha, \beta) \begin{pmatrix} dS \\ dB \end{pmatrix} \\ &= \alpha dS + \beta dB \\ &= \alpha(\mu S dt + \sigma dW) + \beta r B dt \end{aligned} \quad (3.40)$$

and $dF(t, S(t))$ is given by the Itô's Formula:

$$\begin{aligned} dF &= \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial S} dS + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} d\langle S \rangle \\ &\stackrel{(3.33)}{=} \left(\frac{\partial F}{\partial t} + \frac{\partial F}{\partial S} \mu S + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial F}{\partial S} \sigma S dW. \end{aligned} \quad (3.41)$$

Comparing the two terms in pink we find that

$$\alpha(t) = \frac{\partial F}{\partial S} = \Delta(t), \quad (3.42)$$

so $\alpha(t)$ is the equivalent of the difference in the value of the option divided by the difference of the underlying (Δ) in the discrete time framework.

Comparing the remaining terms we get:

$$\alpha(t)\mu S + \beta r B = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial S} \mu S + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} \sigma^2 S^2. \quad (3.43)$$

Multiply both sides of eq. (3.35) for r :

$$rV = r\alpha S + r\beta B$$

and then substitute in eq. (3.43), recalling that $\alpha(t) = \frac{\partial F}{\partial S}$ and $V(t) = F(t, S(t)) \forall t$:

$$\cancel{\frac{\partial F}{\partial S}} \mu S + rF - r \cancel{\frac{\partial F}{\partial S}} S = \frac{\partial F}{\partial t} + \cancel{\frac{\partial F}{\partial S}} \mu S + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} \sigma^2 S^2.$$

We obtain the so called *Black & Scholes partial differential equation*:

$$\begin{cases} \frac{\partial F(t, S(t))}{\partial t} + r \frac{\partial F(t, S(t))}{\partial S(t)} S(t) + \frac{1}{2} \frac{\partial^2 F(t, S(t))}{\partial S(t)^2} \sigma^2 S(t)^2 - rF(t, S(t)) \\ F(T, S(T)) = \text{payoff}_T \end{cases} \quad (3.44)$$

We started with a generic derivative and we assumed that the price was a function $F(t, S(t))$ of the time and the current value of the underlying. Now we have found

that – in order to satisfy the absence of arbitrage opportunity – F obeys to the Black & Scholes PDE. The only thing that characterizes a specific payoff is the terminal condition

$$F(T, S(T)) = \text{payoff}_T \quad (3.45)$$

So, for example, if we want to price a call option we have to solve the PDE (3.44) with the terminal condition $F(T, S(T)) = (S(T) - K)^+$.

Remark 3.3.1. Notice that eq. (3.44) is a deterministic equation. Of course, the solution fluctuates according to the Brownian motion, but at any fixed time it is deterministic.

By a change of variable, there is the possibility to reduce (3.44) to the heat equation. However, we will use a different approach, which involves the *Feynman-Kac formula*.

Theorem 3.3.1 (Feynman-Kac formula, v1). *Assume that $F \in C^{1,2}$ is a solution to the boundary value problem*

$$\begin{cases} \frac{\partial F(t,x)}{\partial t} + \mu \frac{\partial F(t,x)}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 F(t,x)}{\partial x^2} = 0 \\ F(T, x) = \Phi(x) \end{cases} \quad (3.46)$$

where $x \in \mathbb{R}$ and $t \in [0, T]$. Then, provided that $\sigma \frac{\partial F(t,x)}{\partial x} \in \mathcal{H}$, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we can define a Brownian motion $(W_t)_{t \geq 0}$ such that

$$F(t, x) = \mathbb{E}_{x,t}[\Phi(X_T)] \quad (3.47)$$

where X satisfies the SDE

$$\begin{cases} dX_s = \mu(s, X_s)ds + \sigma(s, X_s)dW_s \\ X_t = x. \end{cases} \quad (3.48)$$

Proof. We have that F evolves according the Itô's Formula:

$$\begin{aligned} dF(t, x) &= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial x}dx + \frac{1}{2} \frac{\partial^2 F}{\partial x^2}d\langle x \rangle \\ &\stackrel{(3.48)}{=} \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial x} + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} \sigma^2 dt. \end{aligned}$$

Taking the integral from t to T of both sides we get:

$$\begin{aligned} F(T, x) - F(t, x) &= \int_t^T \left(\frac{\partial F}{\partial u} + \mu \frac{\partial F}{\partial x} + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} \sigma^2 \right) du + \int_t^T \frac{\partial F}{\partial x} \sigma dW_u \\ &\stackrel{(3.46)}{=} \int_t^T \frac{\partial F}{\partial x} \sigma dW_u. \end{aligned}$$

Now we want to remove the stochastic integral, so we take the expected value:

$$\mathbb{E}_{t,x}[F(T, x)] - \mathbb{E}_{t,x}[F(t, x)] = \mathbb{E}_{t,x} \left[\int_t^T \frac{\partial F}{\partial x} \sigma dW_u \right],$$

$$\mathbb{E}_{t,x}[F(T, x)] - F(t, x) = \mathbb{E}_{t,x} \left[\int_t^T \frac{\partial F}{\partial x} \sigma dW_u \right].$$

We know that if $\frac{\partial F}{\partial x}\sigma \in \mathcal{H}$ then the stochastic integral is a \mathbb{P} -martingale, so by definition we have that

$$\mathbb{E}_{t,x} \left[\int_t^T \frac{\partial F}{\partial x} \sigma dW_u \right] = \int_t^t \frac{\partial F}{\partial x} \sigma dW_u = 0.$$

In the end, we get

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

□

We consider also another version which appears over and over again in the study of pricing problems for financial derivatives.

Theorem 3.3.2 (Feynman-Kac formula, v2). *Assume that F is a solution to the boundary value problem*

$$\begin{cases} \frac{\partial F(t,x)}{\partial t} + K \frac{\partial F(t,x)}{\partial x} + \frac{1}{2} H^2 \frac{\partial^2 F(t,x)}{\partial x^2} - rF(t,x) = 0 \\ F(T, x) = \Phi(x) \end{cases} \quad (3.49)$$

where $x \in \mathbb{R}$, $t \in [0, T]$ and r is a given real number. Then, provided that

$$e^{-rt} H \frac{\partial F(t, x)}{\partial x} \in \mathcal{H},$$

there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we can define a Brownian motion $(W_t)_{t \geq 0}$ such that

$$F(t, x) = e^{-r(T-t)} \mathbb{E}_{x,t}[\Phi(X_T)] \quad (3.50)$$

where X satisfies the SDE

$$\begin{cases} dX_s = K(s, X_s)ds + H(s, X_s)dW_s \\ X_t = x. \end{cases} \quad (3.51)$$

Proof. Let's consider the Itô development of the discounted value of F :

$$\begin{aligned} d(e^{-rt} F(t, x)) &= (-re^{-rt} dt)F + e^{-rt} dF + \mathcal{O}(dtdW) \\ &= -re^{-rt} dtF + e^{-rt} dF \\ &= -re^{-rt} dtF + e^{-rt} \left(\frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial x} dx + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} d\langle x \rangle \right) \\ &= -re^{-rt} dtF + e^{-rt} \left(\frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial x} K + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} H^2 \right) dt + e^{-rt} \frac{\partial F}{\partial x} H dW \\ &= e^{-rt} dt \left(-rF + \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial x} K + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} H^2 \right) + e^{-rt} \frac{\partial F}{\partial x} H dW \\ &\stackrel{(3.49)}{=} e^{-rt} \frac{\partial F}{\partial x} H dW. \end{aligned}$$

Taking the integral from t to T of both sides we get:

$$e^{-rT} F(T, x) - e^{-rt} F(t, x) = \int_t^T e^{-rt} \frac{\partial F}{\partial x} H dW_u.$$

Now we want to remove the stochastic integral, so we take the expected value:

$$\mathbb{E}_{t,x}[e^{-rT} F(T, x)] - \mathbb{E}_{t,x}[e^{-rt} F(t, x)] = \mathbb{E}_{t,x} \left[\int_t^T e^{-rt} \frac{\partial F}{\partial x} H dW_u \right],$$

$$\mathbb{E}_{t,x}[e^{-rT}F(T, x)] - e^{-rt}F(t, x) = \mathbb{E}_{t,x}\left[\int_t^T e^{-rs}\frac{\partial F}{\partial x}HdW_s\right],$$

We know that if $e^{-rt}\frac{\partial F}{\partial x}H \in \mathcal{H}$ then the stochastic integral is a \mathbb{P} -martingale, so by the definition we have that

$$\mathbb{E}_{t,x}\left[\int_t^T e^{-rs}\frac{\partial F}{\partial x}HdW_s\right] = 0.$$

In the end, we get

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

□

Let us give also a third formulation.

Theorem 3.3.3 (Feynman-Kac formula, v3). *Assume that F is a solution to the boundary value problem*

$$\begin{cases} \frac{\partial F(t, x)}{\partial t} + K\frac{\partial F(t, x)}{\partial x} + \frac{1}{2}H^2\frac{\partial^2 F(t, x)}{\partial x^2} + g(t, x) = 0 \\ F(T, x) = \Phi(x). \end{cases} \quad (3.52)$$

Then, provided that

$$H\frac{\partial F(t, x)}{\partial x} \in \mathcal{H},$$

there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we can define a Brownian motion $(W_t)_{t \geq 0}$ such that

$$F(t, x) = \mathbb{E}_{x,t}[\Phi(X_T)] + \int_t^T \mathbb{E}_{x,t}[g(s, X_s)]ds, \quad (3.53)$$

where X satisfies the SDE

$$\begin{cases} dX_s = K(s, X_s)ds + H(s, X_s)dW_s \\ X_t = x. \end{cases} \quad (3.54)$$

Proof. Apply Itô's Formula on F :

$$\begin{aligned} dF(t, x) &= \left(\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x}K + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}H^2 \right) dt + \frac{\partial F}{\partial x}HdW \\ &\stackrel{(3.52)}{=} -g(t, x)dt + \frac{\partial F}{\partial x}HdW. \end{aligned}$$

Taking the integral from t to T of both sides we get:

$$F(T, x) - F(t, x) = - \int_t^T g(s, X_s)ds + \int_t^T \frac{\partial F}{\partial x}HdW_s.$$

Now we want to remove the stochastic integral, so we take the expected value:

$$\mathbb{E}_{t,x}[F(T, x)] - F(t, x) = -\mathbb{E}_{t,x}\left[\int_t^T g(s, X_s)ds\right] + \mathbb{E}_{t,x}\left[\int_t^T \frac{\partial F}{\partial x}HdW_s\right].$$

We know that if $H\frac{\partial F}{\partial x}\sigma \in \mathcal{H}$ then the stochastic integral is a \mathbb{P} -martingale, so by definition we have that

$$\mathbb{E}_{t,x}\left[\int_t^T \frac{\partial F}{\partial x}HdW_s\right] = 0.$$

Lesson 11 of
02/04/2020

In the end, we get

$$\begin{aligned}\mathbb{E}_{t,x}[F(T, x)] &= F(t, x) - \left[\int_t^T \mathbb{E}_{t,x}[g(s, x_s)] ds \right], \\ F(t, x) &= \mathbb{E}_{x,t}[\Phi(X_T)] + \left[\int_t^T \mathbb{E}_{t,x}[g(s, x_s)] ds \right].\end{aligned}$$

□

Example 3.3.1

Solve the PDE

$$\begin{cases} \frac{\partial F}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 F}{\partial x^2} = 0 \\ F(T, x) = x^2 \end{cases}$$

where σ is constant.

Solution. In this case $K = 0$ and $H = \sigma$, so the process X follows the differential equation

$$\begin{cases} dX = \sigma dW_t \\ X_t = x. \end{cases} \quad (*)$$

Integrating $(*)$ we get:

$$X_T - x = \sigma(W_T - W_t) \Rightarrow X_T = x + \sigma(W_T - W_t).$$

Using the first version of the Feynman-Kac formula we have the solution:

$$\begin{aligned}F(t, x) &= \mathbb{E}_{x,t}[X_T^2] \\ &= \mathbb{E}_{x,t}[x^2 + \sigma^2(W_T - W_t)^2 + 2x\sigma(W_T - W_t)] \\ &= x^2 + \sigma^2(T - t) + 0 = x^2 + \sigma^2(T - t),\end{aligned}$$

where we used the fact that the Brownian increment $(W_T - W_t)$ is independent of the filtration and that $(W_T - W_t) \sim \mathcal{N}(0, T-t)$, so $dW^2 = dt$ and $(W_T - W_t) = 0$. However, we can not say that it is the solution yet, because we first have to check if F satisfies the PDE

$$\begin{cases} \frac{\partial F}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 F}{\partial x^2} = -\sigma^2 + \frac{\sigma^2}{2}2 = 0 \\ F(T, x) = x^2 + \sigma^2(T - T) = x^2 \end{cases}$$

and then we have to check the regularity condition $\frac{\partial F}{\partial x} H = 2X_t\sigma \in \mathcal{H}$, i.e. it must be that $\mathbb{E} \left[\int_0^T (2X_t\sigma)^2 dt \right] < \infty$:

$$\begin{aligned}\mathbb{E} \left[\int_0^T (2X_t\sigma)^2 dt \right] &= \int_0^T 4\sigma^2 \mathbb{E}[X_t^2] dt \\ &= \int_0^T 4\sigma^2 \mathbb{E}[(X_0 + \sigma W_t)^2] dt \\ &= 4\sigma^2 \int_0^T \mathbb{E}[X_0^2 + \sigma^2 W_t^2 + 2X_0\sigma W_t] dt \\ &= 4\sigma^2 \int_0^T (x_0^2 + \sigma^2 t + 0) dt \\ &= 4\sigma^2 \left(x_0^2 t + \frac{\sigma^2 T^2}{2} \right) < \infty \quad \forall T > 0.\end{aligned}$$

Example 3.3.2

Solve the PDE

$$\begin{cases} \frac{\partial F}{\partial t} + \frac{x^2}{2} \frac{\partial^2 F}{\partial x^2} + x = 0 \\ F(T, x) = \ln x^2. \end{cases}$$

Solution. In this case $K = 0$, $H = x$ and $g(t, x) = x$, so the process X follows the SDE

$$\begin{cases} dX_t = X_t dW_t \\ X_t = x, \end{cases}$$

which can be integrated from t to T , obtaining:

$$X_T = X_t e^{-\frac{1}{2}(T-t)+W_T-W_t}. \quad (\star)$$

Using the Feynman-Kac formula v3 we have the solution

$$\begin{aligned} F(t, x) &= \mathbb{E}_{t,x}[\ln x^2] + \int_t^T \mathbb{E}_{t,x}[X_s] ds \\ &\stackrel{(*)}{=} \mathbb{E}_{t,x}\left[2 \ln x + 2\left(-\frac{1}{2}(T-t) + W_T - W_t\right)\right] + \\ &\quad + \int_t^T \mathbb{E}_{t,x}[X_t e^{-\frac{1}{2}(s-t)+W_s-W_t}] ds \\ &\stackrel{(a)}{=} 2 \ln x - (T-t) + x \int_t^T \mathbb{E}_{t,x}\left[e^{-\frac{1}{2}(s-t)+W_s-W_t}\right] ds \\ &\stackrel{(c)}{=} 2 \ln x - (T-t) + x \int_t^T e^{-\frac{1}{2}(s-t)} \mathbb{E}\left[e^{\mathcal{N}(0,s-t)}\right] ds \\ &\stackrel{(b)}{=} 2 \ln x - (T-t) + x \int_t^T e^{-\frac{1}{2}(s-t)} e^{\frac{1}{2}(s-t)} ds \\ &= 2 \ln x - (T-t) + x(T-t) \\ &= \ln x^2 + (T-t)(x-1), \end{aligned}$$

where in (a) we used the fact that the expected value of $W_T - W_t$ is zero and in (b) we used the fact that since $W_s - W_t$ is independent of \mathcal{F}_t we can remove the conditional expected value. In (c) we used the moment generating function formula $M(t) = \mathbb{E}[e^{tX}] = e^{\mu t} e^{\frac{1}{2}\sigma^2 t^2}$. We also could have used the fact the the exponential in (*) is an exponential martingale.

Now we have to check that F satisfies the PDE:

$$\begin{cases} \frac{\partial F}{\partial t} + \frac{x^2}{2} \frac{\partial^2 F}{\partial x^2} + x = -x + 1 + \frac{x^2}{2} \left(-\frac{2}{x^2}\right) + x = 0 \\ F(T, x) = \ln x^2 + (T-t)(x-1) = \ln x^2. \end{cases}$$

Then we have to check that $\frac{\partial F}{\partial x} H = \left(\frac{2}{X_t} + T-t\right) X_t \in \mathcal{H}$ fai conti :

$$\mathbb{E}\left[\int_0^T (2 + X_t(T-t))^2 dt\right] \stackrel{(*)}{=} \int_0^T e^{-linear(t)} dt < \infty.$$

Let's consider the Black & Scholes PDE

$$\begin{cases} \frac{\partial F(t, S(t))}{\partial t} + r \frac{\partial F(t, S(t))}{\partial S(t)} S(t) + \frac{1}{2} \frac{\partial^2 F(t, S(t))}{\partial S(t)^2} \sigma^2 S(t)^2 - r F(t, S(t)) = 0 \\ F(T, S(T)) = \text{payoff}_T. \end{cases} \quad (3.55)$$

By the Feynman-Kac formula v2 we know that if a function is a solution of the previous equation then we can associate the problem of solving a PDE to the problem of computing an expected value:

$$F(t, S(t)) = e^{-r(T-t)} \mathbb{E}_{t,S_T}^{\hat{\mathbb{P}}} [\text{payoff}_T(S_T)], \quad (3.56)$$

where S_T is the terminal value of the process which satisfies the stochastic equation

$$dS(t) = rS(t)dt + \sigma S(t)dW^{\hat{\mathbb{P}}}. \quad (3.57)$$

Notice that our starting point was a probability space with subjective/historical probability measure \mathbb{P} . Now the F-K formula says that if we start from a deterministic problem we can forget the starting probability space because we have to introduce a new probability space with probability measure $\hat{\mathbb{P}}$. In plain words, under the historical probability measure the dynamics of the risky asset evolves according

$$dS(t) = \mu S(t)dt + \sigma S(t)dW^{\mathbb{P}}, \quad (3.58)$$

but now under the probability $\hat{\mathbb{P}}$ we have a new drift r . Now we would like to give a financial meaning to this fact.

Proposition 3.3.1 (Risk neutral valuation). *The arbitrage free price of a general payoff $\Phi(S(T))$ is given by*

$$F(t, S(t)) = e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}} [\Phi(S(T))], \quad (3.59)$$

where the discounted asset $S_t e^{-rt}$ is a \mathbb{Q} -martingale.

Proof. To prove that $S_t e^{-rt}$ is a \mathbb{Q} -martingale we have to show that the corresponding drift in the differential equation is equal to zero:

$$\begin{aligned} d(S_t e^{-rt}) &= e^{-rt} dS + S_t d(-rt) \\ &= (rS_t dt + \sigma S_t dW^{\mathbb{Q}}) e^{-rt} + \cancel{S_t (-r)e^{-rt} dt} \\ &= \sigma e^{-rt} S_t dW^{\mathbb{Q}}. \end{aligned}$$

In alternative, we know the solution of the corresponding stochastic differential equation:

$$S_t = S_0 \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) t + \sigma W_t^{\mathbb{Q}} \right\}.$$

By multiplying both sides for e^{-rt} we get

$$\begin{aligned} e^{-rt} S_t &= S_0 \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) t + \sigma W_t^{\mathbb{Q}} \right\} e^{-rt} \\ &= S_0 \exp \left\{ -\frac{\sigma^2}{2} t + \sigma W_t^{\mathbb{Q}} \right\} \\ &= \mathbb{Q} - \text{martingale}. \end{aligned}$$

□

There is a natural financial interpretation of the formula (3.59). We see that the price of the derivative, given today's date t and today's stock price s , is computed by taking the expectation of the final payment $\mathbb{E}_t^{\mathbb{Q}} [\Phi(S(T))]$ and then discounting this expected value to present value using the discount factor $e^{-r(T-t)}$. The important point to note is that when we take the expected value we do not use the objective probability measure \mathbb{P} .

Now we will see that we can find the same results taking a completely different approach.

3.4 The martingale approach

The starting point of this approach is the continuous case equivalent of the First Fundamental Theorem we introduced in the discrete time framework.

Theorem 3.4.1 (First Fundamental Theorem). *The model is arbitrage free if and only if there exists a (local) martingale measure \mathbb{Q} equivalent to \mathbb{P} , $\mathbb{Q} \sim \mathbb{P}$, under which*

$$\frac{S^i}{S^0} = \mathbb{Q} - (\text{local}) - \text{martingales}. \quad (3.60)$$

To be honest, this theorem is not perfectly true. A more precise version of the First Fundamental Theorem is the following.

Theorem 3.4.2 (First Fundamental Theorem). *Assume that the asset price process S is bounded. Then there exists an equivalent martingale measure if and only if the model satisfies NFLVR².*

Honestly, this has no real applications in finance (in most applications, the assumption of a bounded S process is far too restrictive), it is only related to mathematical finance.

There is also the equivalent of the Second Fundamental Theorem.

Theorem 3.4.3 (Second Fundamental Theorem). *Assume that the market is arbitrage free and consider a fixed numeraire asset S_0 . Then the market is complete if and only if the martingale measure \mathbb{Q} , corresponding to the numeraire S_0 , is unique.*

3.4.1 Application of the martingale approach to the Black & Scholes model

In the B&S model the dynamics of the asset price is given by

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t^{\mathbb{P}} \quad (3.61)$$

and the evolution of the riskless asset is given by

$$\frac{dB_t}{B_t} = rdt. \quad (3.62)$$

We want to impose the property that characterizes the absence of arbitrage. We know that, under the risk neutral measure, the discounted asset S_t/B_t should be a \mathbb{Q} -martingale. Let's consider its dynamic under the historical probability measure \mathbb{P} :

$$d\left(\frac{S_t}{B_t}\right) = d(S e^{-rt}) = e^{-rt} dS + S d(e^{-rt}) + 0 \quad (3.63)$$

$$= (\mu S dt + \sigma S dW_t^{\mathbb{P}}) e^{-rt} - r S e^{-rt} dt \quad (3.64)$$

$$= e^{-rt} (\mu - r) S dt + \sigma S e^{-rt} dW^{\mathbb{P}}. \quad (3.65)$$

Notice that the drift is $\mu - r \neq 0$, so \mathbb{P} cannot be the pricing measure. So S_t/B_t is not a \mathbb{P} -martingale and then $\mathbb{Q} \neq \mathbb{P}$.

²^No Free Lunch with Vanishing Risk. It is the possibility to exploit an arbitrage not in a finite amount of time but asymptotically. This particular notion of arbitrage is defined ad-hoc in order to establish which is the mathematical condition which ensures the equivalence with the existence of the risk neutral probability measure.

In order to find \mathbb{Q} we impose the no arbitrage condition, i.e. S_t/B_t must be a \mathbb{Q} -martingale. The only possibility to do that is to absorb the drift into the corresponding Brownian motion:

$$\begin{aligned} d\left(\frac{S_t}{B_t}\right) &= e^{-rt}(\mu - r)Sdt + \sigma S e^{-rt} dW^{\mathbb{P}} \\ &= \sigma S e^{-rt} \left(dW^{\mathbb{P}} + \frac{\mu - r}{\sigma} dt \right). \end{aligned} \quad (3.66)$$

Now we have to find a probability measure under which the pink term becomes a Brownian motion $dW^{\mathbb{Q}}$. In the previous approach this was given by Girsanov theorem: if

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_t = \exp \left(\int_0^t \left(\frac{r - \mu}{\sigma} \right) dW_s^{\mathbb{P}} - \frac{1}{2} \int_0^t \left(\frac{r - \mu}{\sigma} \right)^2 ds \right) \quad (3.67)$$

is a true martingale, then

$$W_t^{\mathbb{Q}} := W_t^{\mathbb{P}} - \left(\frac{r - \mu}{\sigma} \right) t \quad (3.68)$$

is a \mathbb{Q} -Brownian motion. But this is true, in fact the Novikov condition holds:

$$\mathbb{E} \left[e^{\frac{1}{2} \int_0^T \left(\frac{r - \mu}{\sigma} \right)^2 ds} \right] \sim e^{constant \cdot T} < \infty \quad (3.69)$$

as the integrand is constant and the whole expression is deterministic. Alternatively, we can use the fact that (3.67) is an exponential martingale.

In the Black & Scholes model σ and μ are constant, but if we consider a more sophisticated model where the volatility is no more constant, $\sigma = \sigma_t$, then it is quite challenging to find a change of measure, because σ_t may be very close to zero and the integral (3.67) may diverge. Moreover, there are some stochastic volatility models in which the Girsanov theorem fails, leading to a loss of probability mass.

Now we can use the shift (3.68) in (3.61):

$$\begin{aligned} \frac{dS}{S} &= \mu dt + \sigma dW_t^{\mathbb{P}} \\ &= \mu dt + \sigma \left(dW^{\mathbb{Q}} + \frac{r - \mu}{\sigma} dt \right) \end{aligned} \quad (3.70)$$

$$= rdt + \sigma dW^{\mathbb{Q}}. \quad (3.71)$$

We end up with the same result we found using the PDEs, i.e. we are able to characterize the dynamics of the risky asset under the risk neutral probability measure \mathbb{Q} . An important difference with respect to the PDE approach is that here we do not make any Markovianity assumption.

In conclusion, we have that under \mathbb{Q} :

- Discounted assets $e^{-rt}S_t$ are \mathbb{Q} -martingales;
- Discounted portfolios $e^{-rt}V_t$ are \mathbb{Q} -martingales. In fact, recall the definition of the portfolio:

$$V_t = \alpha_t S_t + \beta_t B_t \quad (3.72)$$

and then integrate by part the discounted portfolio:

$$d(e^{-rt}V_t) = e^{-rt}dV + V d(e^{-rt}) + 0$$

$$\begin{aligned}
&\stackrel{(a)}{=} (\alpha dS + \beta dB)e^{-rt} - r(\alpha dS + \beta dB)e^{-rt} dt \\
&= e^{-rt} \left(\alpha(rSdt + \sigma SdW^Q) + \beta rBdt - \alpha rSdt + \beta Bdt \right) \\
&= \alpha \sigma S e^{-rt} dW^Q,
\end{aligned} \tag{3.73}$$

where in (a) we used the self financing condition. Since there is no drift, the discounted portfolio is a \mathbb{Q} -martingale. Then, by definition of martingality we have that

$$V_t e^{-rt} = \mathbb{E}_t^{\mathbb{Q}}[V_T e^{-rT}]. \tag{3.74}$$

If we apply this equality to a special case in which V_t is the replicating portfolio – such that $V_T = \text{payoff}_T$ – we get the universal formula

$$\text{price}_t(\text{payoff}_T) = e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}}[\text{payoff}_T]. \tag{3.75}$$

Again, we do not need any Markovianity assumption, so we can consider any payoff, even path-dependent ones.

3.5 Call option price in the Black & Scholes model

Start with the risk neutral dynamics of the assets:

Lesson 12 of
03/04/2020

$$\frac{dS}{S} = \mu dt + \sigma dW^Q, \quad \frac{dB}{B} = r dt. \tag{3.76}$$

The call option payoff is given by

$$\text{Call}_T = (S_T - K)^+ = \text{payoff}_T. \tag{3.77}$$

Let's consider the case in which $t = 0$, so that in the price of the call we have to compute an unconditional expected value:

$$\text{price}_0(\text{Call}) = e^{-rT} \mathbb{E}^{\mathbb{Q}}[(S_T - K)^+]. \tag{3.78}$$

We know that, under \mathbb{Q} , S_T is a Brownian motion:

$$S_T = S_0 \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) T + \sigma W_T^Q \right\}, \tag{3.79}$$

where $W_T^Q = \sqrt{T} \mathcal{N}(0, 1)$. So, we have to compute:

$$\text{price}_0(\text{Call}) = e^{-rT} \int_{\mathbb{R}} \left(S_0 \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) T + \sigma \sqrt{T} y \right\} - K \right)^+ \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy. \tag{3.80}$$

In order to remove the non-linearity given by the positive part, we study the sign of the expression:

$$S_0 \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) T + \sigma \sqrt{T} y \right\} - K \geq 0. \tag{3.81}$$

If we take the logarithm, we get:

$$\left(r - \frac{\sigma^2}{2} \right) T + \sigma \sqrt{T} y \geq \ln \frac{K}{S_0},$$

$$y \geq \frac{\ln \frac{K}{S_0} - \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \equiv \hat{y}.$$

So, if $y \geq \hat{y}$ we can remove the positive part:

$$\begin{aligned} price_0(\text{Call}) &= S_0 e^{-rT} \int_{\hat{y}}^{+\infty} \exp\left\{\left(r - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}y\right\} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy - \\ &\quad - Ke^{-rT} \int_{\hat{y}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &\stackrel{(a)}{=} \frac{S_0}{\sqrt{2\pi}} \int_{\hat{y}}^{+\infty} \exp\left\{-\frac{\sigma^2}{2}T + \sigma\sqrt{T}y - \frac{y^2}{2}\right\} dy - Ke^{-rT} \Phi(-\hat{y}) \\ &\stackrel{(b)}{=} \frac{S_0}{\sqrt{2\pi}} \int_{\hat{y}}^{+\infty} \exp\left\{-\frac{1}{2}(y - \sigma\sqrt{T})^2\right\} dy - Ke^{-rT} \Phi(-\hat{y}) \\ &\stackrel{(c)}{=} \frac{S_0}{\sqrt{2\pi}} \int_{\hat{y} - \sigma\sqrt{T}}^{+\infty} e^{-\frac{z^2}{2}} dz - Ke^{-rT} \Phi(-\hat{y}) \\ &= S_0 \Phi(\sigma\sqrt{T} - \hat{y}) - Ke^{-rT} \Phi(-\hat{y}), \end{aligned} \tag{3.82}$$

where:

- in (a) we used the cumulative distribution function

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy;$$

- in (b) we recognized that the argument of the exponent is a square expansion;
- in (c) we introduced the change of variable $z = y - \sigma\sqrt{T}$, $dz = dy$.

So we end up with the following Black & Scholes formula:

$$price_0(\text{Call}) = S_0 \Phi(d_1) - Ke^{-rT} \Phi(d_2), \tag{3.83}$$

where

$$d_1 = \frac{\ln \frac{S_0}{K} + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}. \tag{3.84}$$

What about the case in which $t \in (0, T)$? In this case

$$price_t(\text{Call}) = e^{-rt} \mathbb{E}_t^\mathbb{Q}[(S_T - K)^+] \tag{3.85}$$

and

$$\begin{aligned} S_T &= S_0 \exp\left\{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma(W_T^\mathbb{Q} - W_t^\mathbb{Q})\right\} \\ &= S_0 \exp\left\{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\sqrt{T-t} \mathcal{N}(0, 1)\right\}. \end{aligned} \tag{3.86}$$

Now S_t is measurable and $(W_T^\mathbb{Q} - W_t^\mathbb{Q})$ is a gaussian random variable independent of the information available at time t , \mathcal{F}_t . This independence allows us to eliminate the conditioning in the expected value. So we only have to repeat the procedure substituting $S_0 \rightarrow S_t$ and $T \rightarrow T-t$. In the end we arrive at the following Black & Scholes formula (1973):

$$price_t(\text{Call}) = S_t \Phi(d_1) - Ke^{-r(T-t)} \Phi(d_2), \tag{3.87}$$

where

$$d_1 = \frac{\ln \frac{S_t}{K} + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t}. \quad (3.88)$$

At the time, no financial journal accepted to publish this result. In order to understand the reason, we consider an example.

Example 3.5.1

Suppose we have two assets, Alitalia and Facebook, and that the volatility and the initial price are the same, $\sigma_A = \sigma_F$ and $S_0^A = S_0^F$. Let's consider a call option at the money ($S_0 = K$) with maturity $T = 1$ year for both of them. Facebook's shares price is expected to increase, while Alitalia's one is expected to decrease. So, the payoffs will be

$$(S_T^F - K)^+ > 0, \quad (S_T^A - K)^+ = 0$$

So, it is reasonable to assume that

$$\text{price}_t^A(\text{Call}) < \text{price}_t^F(\text{Call})$$

The problem is that we know that all the prices satisfy the PDE, which has solution (3.87). But (3.87) contains no information about the drift, because it is written in terms of the risk neutral probability measure. So, if the two assets share the same parameters (volatility, initial price and maturity) they will have the same price, regardless their historical behavior.

$$\text{price}_t^A(\text{Call}) = \text{price}_t^F(\text{Call})$$

This is completely counter-intuitive, and this is why no journal accepted to publish the paper by Black and Scholes. However, even if the academic community did not accept this result^a, traders were already using it in their calculators.

^aIn the end, the paper was published in the *Journal of Political Economy*, which is not a quantitative journal.

3.6 The Greeks

Each Greek letter measures a different dimension of the risk in an option position. The aim of a trader is to manage the Greeks so that all risks are acceptable.

3.6.1 Delta

The most important Greek is the Δ , which represents the sensitivity of the option price (given by the B&S formula) with respect to the price of the underlying asset:

$$\Delta = \frac{\partial \text{price}}{\partial S} = \Phi(d_1). \quad (3.89)$$

Apparently, this differentiation is not correct, because inside d_1 there is a dependence of S , as we can see in (3.88). The correct one should be:

$$\Delta = \frac{\partial \text{price}}{\partial S} = \Phi(d_1) + S_t \frac{\partial \Phi(d_1)}{\partial S} - K e^{-r(t-T)} \frac{\partial \Phi(d_2)}{\partial S}. \quad (3.90)$$

Nevertheless, eq. (3.89) is correct. Let's prove it. Recall that

$$\Phi(x) = \int_{-\infty}^x \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy.$$

In order to compute the derivative of this integral we use the Lebesgue formula, where we can have the dependence of a parameter (even in the extremes of integration):

$$\frac{d}{dt} \int_{\alpha(t)}^{\beta(t)} f(x, t) dx = \beta'(t)f(\beta(t), t) - \alpha'(t)f(\alpha(t), t) + \int_{\alpha(t)}^{\beta(t)} \frac{\partial f(x, t)}{\partial t} dx. \quad (3.91)$$

In our particular case, we have:

$$\frac{d\Phi(d_1)}{dS} = \frac{d}{dS} \int_{-\infty}^{d_1(S)} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy = d'_1(S) \frac{e^{-\frac{d_1^2}{2}}}{\sqrt{2\pi}} = \frac{1}{S_t \sigma \sqrt{T-t}} \frac{e^{-\frac{d_1^2}{2}}}{\sqrt{2\pi}}, \quad (3.92)$$

$$\frac{d\Phi(d_2)}{dS} = \frac{d}{dS} \int_{-\infty}^{d_2(S)} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy = d'_2(S) \frac{e^{-\frac{d_2^2}{2}}}{\sqrt{2\pi}} = \frac{1}{S_t \sigma \sqrt{T-t}} \frac{e^{-\frac{d_2^2}{2}}}{\sqrt{2\pi}}, \quad (3.93)$$

and

$$\begin{aligned} \Delta_t &= \Phi(d_1) + S_t \frac{1}{S_t \sigma \sqrt{T-t}} \frac{e^{-\frac{d_1^2}{2}}}{\sqrt{2\pi}} - K e^{-r(T-t)} \frac{1}{S_t \sigma \sqrt{T-t}} \frac{e^{-\frac{d_2^2}{2}}}{\sqrt{2\pi}} \\ &= \Phi(d_1) + \frac{1}{S_t \sigma \sqrt{T-t} \sqrt{2\pi}} \left(S_t e^{-\frac{d_1^2}{2}} - K e^{-r(T-t)} e^{-\frac{d_2^2}{2}} \right) \\ &\stackrel{(a)}{=} \Phi(d_1) + \frac{e^{-\frac{d_2^2}{2}}}{S_t \sigma \sqrt{T-t} \sqrt{2\pi}} \left(S_t e^{-\frac{\sigma^2}{2}(T-t) - d_2 \sigma(T-t)} - K e^{-r(T-t)} \right) \\ &\stackrel{(b)}{=} \Phi(d_1) + \frac{e^{-\frac{d_2^2}{2}}}{S_t \sigma \sqrt{T-t} \sqrt{2\pi}} \left(S_t e^{-\frac{\sigma^2}{2}(T-t) - \ln \frac{S_t}{K} - \left(r - \frac{\sigma^2}{2}\right)(T-t)} - K e^{-r(T-t)} \right) \\ &= \Phi(d_1) + \frac{e^{-\frac{d_2^2}{2}}}{S_t \sigma \sqrt{T-t} \sqrt{2\pi}} \left(S_t e^{-\ln \frac{S_t}{K} - r(T-t)} - K e^{-r(T-t)} \right) \\ &= \Phi(d_1) + \frac{e^{-\frac{d_2^2}{2} - r(T-t)}}{S_t \sigma \sqrt{T-t} \sqrt{2\pi}} \left(S_t \frac{K}{S_t} - K \right) \\ &= \Phi(d_1), \end{aligned}$$

where in (a) we used the fact that

$$\begin{aligned} d_1 &= d_2 + \sigma \sqrt{T-t} \quad \Rightarrow \quad d_1^2 = d_2^2 - \sigma^2(T-t) + 2d_2 \sigma \sqrt{T-t} \\ &\Rightarrow \quad -\frac{1}{2} d_1^2 = -\frac{1}{2} d_2^2 - \frac{\sigma^2}{2}(T-t) - d_2 \sigma \sqrt{T-t} \end{aligned}$$

and in (b) we used the fact that

$$d_2 \sigma \sqrt{T-t} = \ln \frac{S_t}{K} + \left(r - \frac{\sigma^2}{2} \right).$$

So, the number of shares we have to get in our portfolio in order to replicate the payoff at the end is $\Phi(d_1)$.

The call payoff is equal to zero until K and then it is a straight line with slope 45° .

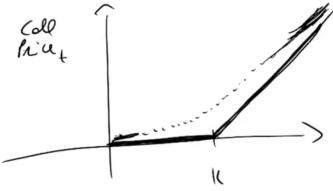
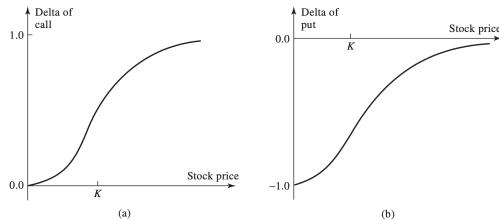
Figure 3.4: $\Phi(d_1)$ for call payoff.

Figure 3.5: Variation of delta with stock price for (a) a call option and (b) a put option on a non-dividend-paying stock.

Being the call price a double derivable function, it must be a smooth function which is very close to zero when the underlying is very low and reaches the behavior of the straight line when the underlying is very big (see Figure 3.4). When we take its derivative with respect to S , i.e. the Δ , we obtain the cumulative gaussian density (see Figure 3.5). The financial intuition is the following. When the underlying is less than the strike we are out of the money and there is no need to keep it in our portfolio. On the other hand, if the underlying is greater than the strike we are in the money and in order to construct the payoff for our long counterpart we have to hold a great amount of underlying (almost 1). The problem arises around the strike K , because for small fluctuations around K there can be huge changes in the hedging portfolio.

So, hedging is more irregular than pricing. This is due to the fact that pricing involves an expectation, which is an integral that increases regularity. On the contrary, by deriving we loose regularity.

Now we would like to know the frequency at which we have to re-balance our portfolio in order to follow the fluctuations of the underlying. This is given by the Greek Γ .

Lesson 13 of
09/04/2020

3.6.2 Gamma

The gamma (Γ) of a portfolio of options on an underlying asset is the rate of change of the portfolio's delta with respect to the price of the underlying asset. It is the second partial derivative of the portfolio with respect to asset price:

$$\Gamma = \frac{\partial^2 \text{price}}{\partial S^2} = \frac{\partial \Delta}{\partial S} = \frac{1}{S\sigma\sqrt{T-t}} \frac{e^{-\frac{d_1}{2}}}{\sqrt{2\pi}}. \quad (3.94)$$

If gamma is small, the delta changes slowly and adjustments to keep a portfolio delta-neutral need to be made only relatively infrequently. If gamma is highly negative or highly positive, the delta is very sensitive to the price of the underlying asset and it is quite risky to leave a delta-neutral portfolio unchanged for any length of time.

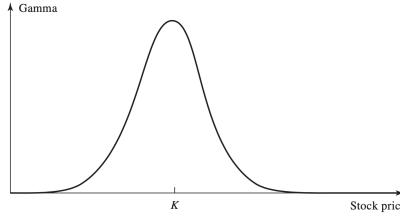


Figure 3.6: Variation of gamma with stock price for an option.

3.6.3 Rho

The rho of a portfolio of options is the rate of change of the value of the portfolio with respect to the interest rate:

$$\rho = \frac{\partial \text{price}}{\partial r} = K(T-t)e^{-r(T-t)}\Phi(d_2). \quad (3.95)$$

It measures the sensitivity of the value of a portfolio to a change in the interest rate when all else remains the same. Notice that it is not normalized as the delta.

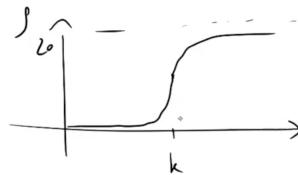


Figure 3.7: Variation of rho with stock price for an option.

Its interpretation is that if we are far from the strike, a bias in the estimation of the interest rate is not a problem. Conversely, if we are around the strike a small bias in the estimation of the interest rate results in a huge change in the price.

3.6.4 Theta

The theta (Θ) of a portfolio of options is the rate of change of the value of the portfolio with respect to the passage of time with all else remaining the same. Theta is sometimes referred to as the *time decay* of the portfolio.

$$\Theta = \frac{\partial \text{price}}{\partial t} = \frac{-S\sigma e^{-\frac{d_1^2}{2}}}{4\sqrt{2\pi(T-t)}} - rKe^{-r(T-t)}\Phi(d_2). \quad (3.96)$$

Theta is usually negative for an option. This is because, as time passes with all else remaining the same, the option tends to become less valuable (we loose the opportunity to get a greater payoff). The variation of Θ with stock price for a call option is shown in Figure 3.8.

When the stock price is very low, theta is close to zero. For an at-the-money call option, theta is large and negative. As the stock price becomes larger, theta tends to $-rKe^{-rT}$.

3.6.5 Vega

Up to now we have implicitly assumed that the volatility of the asset underlying a derivative is constant. In practice, volatilities change over time. This means that the value of a derivative is liable to change because of movements in volatility as well as

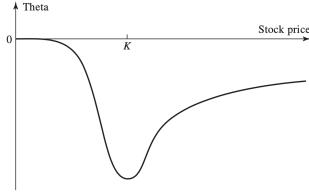


Figure 3.8: Variation of theta of a European call option with stock price.

because of changes in the asset price and the passage of time.

The vega of a portfolio of derivatives, \mathcal{V} , is the rate of change of the value of the portfolio with respect to the volatility of the underlying asset:

$$\mathcal{V} = \frac{\partial \text{price}}{\partial \sigma}. \quad (3.97)$$

If vega is highly positive or highly negative, the portfolio's value is very sensitive to small changes in volatility. If it is close to zero, volatility changes have relatively little impact on the value of the portfolio.

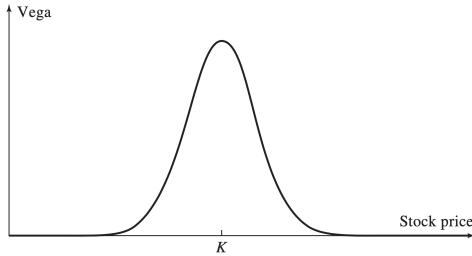


Figure 3.9: Variation of vega with stock price for an option.

The fact that $\mathcal{V} > 0$ means that there is a monotone increasing relationship between price and volatility. In other words, the higher the volatility of the option, the higher the price of the option. We can exploit this relationship to invert the B&S formula and find the volatility in terms of the price. Let us fix the maturity T and suppose that we want to price a set of options indexed by different strikes K_1, \dots, K_n :

$$B\&S(S, T - t, r, K_1, \sigma), \dots, B\&S(S, T - t, r, K_n, \sigma).$$

Now, if we consider the quoted prices

$$mkt(S, T - t, K_1), \dots, mkt(S, T - t, K_n)$$

we can look for the value of the volatility – called *implied volatility* – to plug in the B&S formula in order to obtain these prices. If the B&S model is correct we should get always the same value of σ_{imp} , but indeed in practice we find non-constant values. This distortion is known as *smile effect* and it is the proof that the B&S model is not 100% correct.

However, the B&S is very useful because it allows to quote the options not in dollars but in terms of the implied volatility.

The form of the smile is very important because it tells us how much the B&S model is wrong. If the smile is flat it means that the B&S model is not so wrong, while if it is more pronounced it means that it is dangerous to approximate prices using it. The B&S model is important also because in some contracts there is not a quoted price, so it must be deduced.

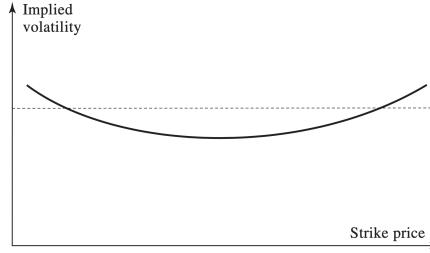


Figure 3.10: Smile effect.

If we consider different values of the maturity T and we introduce the time to maturity as an additional dimension, we can find the so called *implied volatility surface*. Typically, for short times to maturity the smile is very pronounced (the violation of the B&S model is stronger) and for longer times to maturity the smile flattens (because of the central limit theorem).

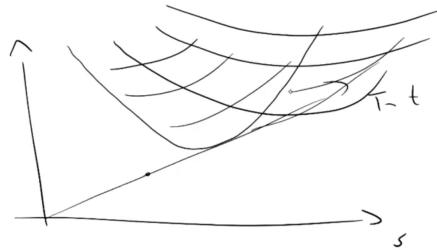


Figure 3.11: Implied volatility surface.

3.6.6 Put option pricing in the Black & Scholes model

From the put-call parity we have that:

$$\begin{aligned} \text{Put}_t &= \text{Call}_t + Ke^{-r(T-t)} - S_t \\ &= S_t \Phi(d_1) - Ke^{-r(T-t)} \Phi(d_2) + Ke^{-r(T-t)} - S_t \\ &= Ke^{-r(T-t)} \underbrace{(1 - \Phi(d_2))}_{>0} - S_t \underbrace{(1 - \Phi(d_1))}_{>0}. \end{aligned} \quad (3.98)$$

Now we want to understand how to hedge a put option, in particular to find the Greeks. We have:

$$\Delta^{\text{Put}} = \frac{\partial \text{Put}_t}{\partial S} = \Phi(d_1) - 1. \quad (3.99)$$

In this case the delta of the put is always negative, as we can see in Figure 3.5 (b). We also have

$$\Gamma^{\text{Put}} = \Gamma^{\text{Call}}, \quad (3.100)$$

$$\gamma^{\text{Put}} = \gamma^{\text{Call}}. \quad (3.101)$$

Eq. (3.101) is very important because it says that the price of the put is an increasing function of the corresponding volatility. This means that we can repeat the procedure of inverting the formula of the price of the put in terms of the volatility in order to get the quoted market price. Even in this case we will have the smile effect.

What happens to the Greeks and to the price if there is an upward shift in the volatility? Consider a call option. In any case there are the constraints to start from zero and to converge to the straight line $(S_t - K)^+$. Then we know that

for higher volatility we will have higher prices. So, for $\sigma_2 > \sigma_1$ the curve will stay above the original price. The delta will change such that it will be convenient to buy the underlying earlier and to sell it immediately after the strike. If the gamma

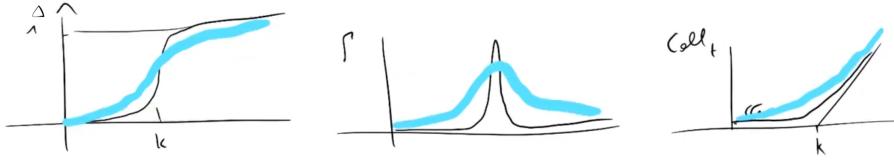


Figure 3.12: Delta, gamma and price of a call option for σ_1 (in black) and $\sigma_2 > \sigma_1$ (in blue).

corresponding to σ_1 was like a Dirac delta, the one for σ_2 will be more spread out. In other words, there is a smoothing effect both on the price and on the Greeks.

3.7 Portfolio immunization

We would like our portfolio to be Δ -neutral, i.e. we would like the global delta of our hedging strategy to match exactly the global delta of the corresponding portfolio derivatives. However, immediately after we balanced the portfolio, there can be a shock in the market. In other words, there is no guarantee for us to stay with the same hedging portfolio for too long. In order to stay safe for a while we need our portfolio to be Γ -neutral, so that the re-balance frequency is low. Moreover, since the volatility is a crucial ingredient we would like our portfolio to be also \mathcal{V} -neutral. For example, from the B&S PDE

$$\frac{\partial F}{\partial t} + rS \frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = rF \quad (3.102)$$

we can replace the derivatives with the corresponding Greeks:

$$\Theta + rS\Delta + \frac{1}{2}\sigma^2 S^2\Gamma = r\text{ price}. \quad (3.103)$$

For a Δ -neutral portfolio ($\Delta = 0$) we have

$$\Theta + \frac{1}{2}\sigma^2 S^2\Gamma = r\text{ price}. \quad (3.104)$$

From this equality we immediately understand that if $\Theta < 0$ then $\Gamma > 0$, which is true for *vanilla options* (call/put). Conversely, if $\Gamma < 0$ then $\Theta > 0$, which typically happens to *exotic options*. In order to understand what does it mean to have a $\Gamma < 0$, let's start from the dynamics of the hedging portfolio:

$$d\text{price}_t = \Theta dt + S\Delta dS + \frac{1}{2}\sigma^2 \Gamma dS^2. \quad (3.105)$$

Notice that this is a quadratic function of dS . If Γ is positive it means that for large variations of the underlying – i.e. shocks in the market – there is no risk (the hedging portfolio leads to a positive P&L³), whereas for small variations we can incur in some potential risk (negative P&L). If $\Gamma < 0$ we have the converse: for small variations of the underlying there is no risk (positive P&L) whereas for large variations we can incur in some risk (negative P&L).

³^The *profit and loss* (P&L) statement is a financial statement that summarizes the revenues, costs, and expenses incurred during a specified period, usually a fiscal quarter or year. The P&L statement is synonymous with the income statement. These records provide information about a company's ability or inability to generate profit by increasing revenue, reducing costs, or both.



Figure 3.13: Hedging portfolio price for (a) $\Gamma > 0$ and (b) $\Gamma < 0$.

Now let's construct an immune portfolio. At the end of the trading day the value of our portfolio is V_t and the corresponding delta is Δ_t^V . If $\Delta_t^V \gg 0$ or $\Delta_t^V \ll 0$ there is a mismatch between the behavior of what we are trying to hedge and our corresponding hedging strategy, known as *delta risk*. We have to look for an instrument in the market that allows us to neutralize the global delta. There are two possibilities:

- Buy/sell options (put, call, ...). In the immunization we have to take into account their prices;
- Consider linear contracts (forward contract, futures, ...). Since we do not have to pay in order to enter these contracts, these are cheaper strategies. These are delta-1 products, i.e. their delta is 1. In fact, if we consider as an example a long position in a forward contract, we have that the payoff is $S_T - K$ and according to the risk neutral methodology its price is

$$\begin{aligned} \text{price}_t^{\text{FWD}} &= e^{r(T-t)} \mathbb{E}^{\mathbb{Q}}[S_T - K] \\ &= e^{rt} \mathbb{E}^{\mathbb{Q}}[e^{-rT} S_T] - K e^{r(T-t)} \\ &= e^{rt} e^{-rt} S_t - K e^{-r(T-t)} \\ &= S_t - K e^{-r(T-t)}, \end{aligned}$$

so the delta is $\Delta_t^{\text{FWD}} = 1$.

Let's consider the situation in which $\Delta_t^V \neq 0$ and we take a position in the forward market entering a contract with maturity T and strike K , so that

$$V_t^{\text{tot}} = V_t + \alpha \text{FWD}(T, K), \quad (3.106)$$

$$\Delta_t^{\text{tot}} = \Delta_t^V + \alpha \Delta^{\text{FWD}} = \Delta_t^V + \alpha \cdot 1. \quad (3.107)$$

If we impose $\Delta_t^{\text{tot}} = 0$ in the last equation we have that the number of forward contracts we have to include in our portfolio is

$$\alpha = -\Delta_t^V. \quad (3.108)$$

Furthermore, if we want also to neutralize the portfolio with respect to the *gamma risk* (related to the frequency of re-balancing) then linear contracts are useless, because their second derivative with respect to S is zero and so $\Gamma = 0$. The idea is to take a position in the options market. The portfolio will be

$$V_t^{\text{tot}} = V_t + \alpha \text{FWD}_t + \beta \text{Call}(T, K) \quad (3.109)$$

and we have to jointly impose the delta and gamma neutrality. In fact, if we take the previous Δ -neutral portfolio and we impose the Γ -neutrality, the involved call option will have $\Delta^{\text{Call}} \neq 0$ so the neutrality is destabilized. In other words, we have to solve

$$\begin{cases} 0 = \Delta_t^{\text{tot}} = \Delta_t^V + \alpha + \beta \Delta_t^{\text{Call}} \\ 0 = \Gamma_t^{\text{tot}} = \Gamma_t^V + 0 + \beta \Gamma_t^{\text{Call}}. \end{cases} \quad (3.110)$$

If we want also vega-neutrality we have to consider a second option, different from the first (for example with different strike).

3.8 Digital options

Digital options are related to a specific type of European options for which the payoff is either some fixed monetary amount or nothing at all:

$$\text{payoff}_T = \mathbb{1}_{S_T \geq K}. \quad (3.111)$$

3.8.1 Pricing

For pricing we could adapt the risk neutral methodology, but there is a much simpler way. In fact, we can write the call payoff as:

$$(S_T - K)^+ = (S_T - K)\mathbb{1}_{S_T \geq K} = S_T\mathbb{1}_{S_T \geq K} - K\mathbb{1}_{S_T \geq K}. \quad (3.112)$$

So, the price of the digital option payoff is:

$$\begin{aligned} \text{price}_t^{DIG} &= e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}}[\mathbb{1}_{S_T \geq K}] \\ &= e^{-r(T-t)} \mathbb{Q}_t(S_T \geq K) \\ &= e^{-r(T-t)} \Phi(d_2), \end{aligned} \quad (3.113)$$

where \mathbb{Q} is the conditional risk neutral probability. We only have to compute the gaussian CDF. The problems appear when we consider the hedging: we saw that in the call option the payoff is continuous (it is the solution of a PDE) and differentiable. Here the payoff is not continuous, so we expect additional irregularities in the Greeks.

3.8.2 Hedging

Recall that

$$\Phi(d_2) = \int_{-\infty}^{d_2} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \quad \Rightarrow \quad \frac{\partial \Phi(d_2)}{\partial S} = \frac{e^{-\frac{d_2^2}{2}}}{\sqrt{2\pi}} \frac{1}{S \sigma \sqrt{T-t}},$$

so the delta is

$$\Delta_t^{DIG} = \frac{e^{-r(T-t)} e^{-\frac{d_2^2}{2}}}{\sqrt{2\pi} S \sigma \sqrt{T-t}} \quad (3.114)$$

and in this case its shape is like a Dirac delta.

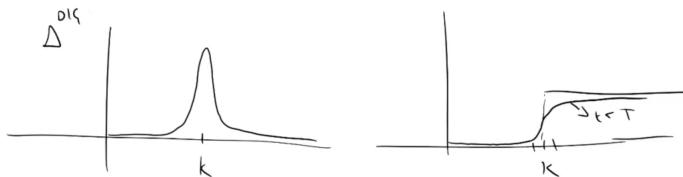


Figure 3.14: (a) Delta and (b) pricing function for the digital options.

With such a delta, if the underlying fluctuates in a small interval around the strike, it is impossible to hedge a digital option. In fact, the price function is such that it is impossible to buy or sell without incurring in a huge risk (in correspondence of the unitary jump).

Unfortunately, the traders' portfolios are typically full of digital options because they think in a myopic way: three years and three years plus a week are almost the same. So, if they have to match some conditions in a horizon of three years, they look in the

market for some products with maturity quite close to the one they have to hedge. But as time goes by the maturity approaches and if for example we hedge with an option which maturity is a bit shorter there will be a small additional increment of maturity (one week) in which we are uncovered (or in some other cases overcovered, i.e. we are over-hedging). The way traders hedge digital options is like a proxy, in the sense that instead of considering the true payoff of the option (the step function) they consider a combination of different options. For example, they consider a long call with strike K_1 and a short call with strike K_2 (Figure 3.15). This combination is called *bull spread* (we have limited profits but limited losses).

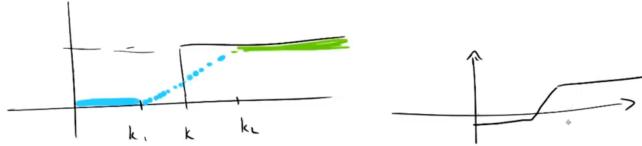


Figure 3.15: Bull spread.

Let's consider the gamma:

$$\Gamma^{DIG} = \frac{\partial \Delta^{DIG}}{\partial S} \quad (3.115)$$

The explicit expression is not very interesting, the important thing is that it can be both positive or negative. If it is negative there is the possibility to incur in very heavy losses for large fluctuations of the underlying. This is an example of *exotic options*, because it is very difficult to hedge and it has a potential negative gamma.

Example 3.8.1

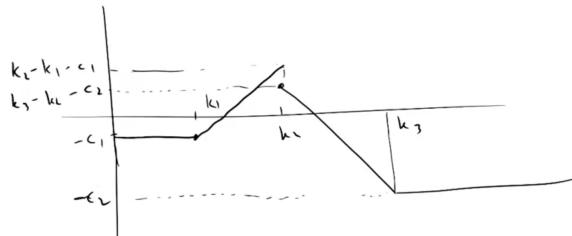
Consider the following payoff:

$$\text{payoff}_T = \begin{cases} -c_1 & S_T \in [0, K_1) \\ S_T - K_1 - c_1 & S_T \in [K_1, K_2) \\ K_3 - S_T - c_2 & S_T \in [K_2, K_3) \\ -c_2 & S_T \geq K_3 \end{cases} \quad (3.116)$$

where $0 < c_1 < c_2$ and $0 < K_1 < K_2 < K_3$.

1. How can we price a product which leads to this payoff?
2. Which is the form of delta and gamma?
3. Show how the price changes for an upward shift of the volatility $\sigma_1 \rightarrow \sigma_2 > \sigma_1$.

Solution. 1. First, let's visualize the payoff function.



In principle, we can always write the price as the discounted value of the risk neutral expectation of the payoff:

$$\text{price}_t(\text{payoff}_T) = e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}}[\text{payoff}_T]$$

$$= e^{-r(T-t)} \left(\int_{-\infty}^{K_1} \dots + \int_{K_1}^{K_2} \dots + \int_{K_2}^{K_3} \dots + \int_{K_3}^{+\infty} \dots \right) \quad (3.117)$$

However, we would like to avoid brute force integration. Instead, we would like to write this expression in terms of some building blocks we already know how to price. We can start from the left or from the right. If we start from the left, the idea is to start with some product and add something else such that what is on the left remains unchanged and in case of need we can change what is in the right. But we know something that leaves the left unchanged: the call option. In fact, when the underlying is smaller than the strike, the call option has no effect. If we start from the right, the opposite holds and we can use put options. Let's start from the left. The payoff decomposition is the following:

- The payoff starts from $-c_1$ and it is constant till K_1 . In order to have a straight line with slope $+1$ we consider a long position in a call option with strike K_1 :

$$\text{payoff}_T = -c_1 + (S_T + K_1)^+ + \dots$$

- Now, in order to change the slope from -1 to 1 we take two short positions in a call with strike K_2 :

$$\text{payoff}_T = -c_1 + (S_T - K_1)^+ - 2(S_T - K_2)^+ + \dots$$

and in order to have the discontinuity in K_2 we translate downward the payoff by taking a number of short positions in a digital option equal to the size of the step:

$$\begin{aligned} \text{payoff}_T = & -c_1 + (S_T - K_1)^+ - 2(S_T - K_2)^+ - \\ & -(K_2 - K_1 - c_1 - (K_3 - K_2 - c_2)) \mathbb{1}_{S_T \geq K_2} + \dots \end{aligned}$$

- Now at K_3 we change the slope from -1 to 0 by taking a long position in a call with strike K_3 :

$$\begin{aligned} \text{payoff}_T = & -c_1 + (S_T - K_1)^+ - 2(S_T - K_2)^+ - \\ & -(K_2 - K_1 - c_1 - (K_3 - K_2 - c_2)) \mathbb{1}_{S_T \geq K_2} + (S_T - K_3)^+. \end{aligned} \quad (3.118)$$

The corresponding price of this payoff is:

$$\begin{aligned} \text{price}_t(\text{payoff}_T) = & -c_1 e^{-r(T-t)} + S_t \Phi(d_1^{K_1}) - e^{-r(T-t)} K_1 \Phi(d_2^{K_1}) - \\ & - 2(S_t \Phi(d_1^{K_2}) - e^{-r(T-t)} K_2 \Phi(d_2^{K_2})) - \\ & - (2K_2 - K_1 - K_3 - c_1 + c_2) e^{-r(T-t)} \Phi(d_2^{K_2}) + \\ & + S_t \Phi(d_1^{K_3}) - e^{-r(T-t)} K_3 \Phi(d_2^{K_3}), \end{aligned} \quad (3.119)$$

where

$$d_1^{K_i} = \frac{\ln \frac{S}{K_i} + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma \sqrt{T-t}}, \quad i = 1, 2, 3. \quad (3.120)$$

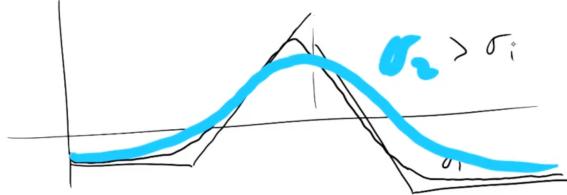
2. The delta is given by the sum of the building blocks' deltas:

$$\Delta_t = 0 + \Phi(d_1^{K_1}) - 2\Phi(d_1^{K_2}) - \Delta^{DIG, K_2} + \Phi(d_1^{K_3}) \quad (3.121)$$

and the gamma is

$$\Gamma_t = 0 + \frac{1}{S_t \sigma \sqrt{T-t}} \frac{e^{-\frac{d_1^{K_1}}{2}}}{\sqrt{2\pi}} - 2 \frac{1}{S \sigma \sqrt{T-t}} \frac{e^{-\frac{d_1^{K_2}}{2}}}{\sqrt{2\pi}} - \Gamma^{DIG, K_2} + \frac{1}{S_t \sigma \sqrt{T-t}} \frac{e^{-\frac{d_1^{K_3}}{2}}}{\sqrt{2\pi}}. \quad (3.122)$$

3. If the volatility is $\sigma_2 > \sigma_1$ the pricing function will be smoother.



3.9 Black & Scholes model with continuous dividends

In this section we consider the case when dividends are paid out continuously in time. As usual S_t denotes the price of the stock at time t , and we denote the cumulative dividends over the interval $[0, t]$ as

$$D(t) = \int_0^t \tilde{\delta}_s ds \quad (3.123)$$

where $\tilde{\delta}$ is the rate of dividends. Put in differential form, this means that over the infinitesimal interval $(t, t + dt]$ the holder of the stock receives the amount $dD(t) = D(t + dt) - D(t)$.

The starting point is the B&S formula:

$$\frac{dS}{S} = \mu dt + \sigma dW^{\mathbb{P}} \quad (3.124)$$

where \mathbb{P} is the historical probability measure. We still don't know which is the risk neutral probability measure because here the price of the asset changes differently according to the fact that some dividends are delivered or not. So, in order to find the risk neutral measure \mathbb{Q} we have to start from the beginning. We define our portfolio as

$$V_t = h_t \cdot S_t, \quad S_t = \begin{pmatrix} S_t \\ e^{rt} \end{pmatrix}. \quad (3.125)$$

The old portfolio (before dividends) has value $h(t - \Delta t)S(t)$. Suppose that at time t there is a delivery of some dividends, such that we get $h(t - \Delta t)(D(t) - D(t - \Delta t))$. The budget equation is given by:

$$h(t - \Delta t)S(t) + h(t - \Delta t)(D(t) - D(t - \Delta t)) = h(t)S(t) + c(t)\Delta t. \quad (3.126)$$

If we add and subtract $S(t - \Delta t)\Delta h(t)$ we end up with:

$$dV_t = h(t)dS_t + h(t)dD_t - c(t)dt$$

$$= h(t)(dS_t + dD_t) - c(t)dt. \quad (3.127)$$

In the presence of dividends, the self financing condition involves a *gain process*, because in addition to the price we get something more, i.e. the dividends (which can be seen as an incoming salary). Since $dD(t) = D(t+dt) - D(t)$ is a forward increment we can use Itô's formula. Now, under the measure \mathbb{Q} , we want the discounted asset with the additional presence of the value of the future streams of dividends

$$S_t e^{-rt} + \int_0^t e^{-rs} dD_s \quad (3.128)$$

to be a martingale. This process is called *normalized gain process*.

Now we look for the pricing PDE. Starting from the dynamics of our portfolio

$$\begin{cases} dV = hdS + hdD + 0 \\ V_T = \text{payoff}_T = F(S_T, T), \end{cases} \quad (3.129)$$

where the 0 means no consumption, we have

$$h_t dS_t + h_t dD_t = \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial S} dS + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 F}{\partial S^2} dt. \quad (3.130)$$

Recall that

$$h = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad S_t = \begin{pmatrix} S \\ e^{rt} \end{pmatrix}, \quad D = \begin{pmatrix} \int_0^t \tilde{\delta}_s ds \\ 0 \end{pmatrix}, \quad (3.131)$$

where the second element of D is zero because there are no dividends in the riskless asset. Substituting in (3.130) we get:

$$\begin{aligned} \alpha S(\mu dt + \sigma dW^{\mathbb{P}}) + \beta r e^{rt} dt + \alpha \tilde{\delta}_t dt &= \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial S} S(\mu dt + \sigma dW^{\mathbb{P}}) + \\ &\quad + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 F}{\partial S^2} dt. \end{aligned} \quad (3.132)$$

Comparing the pink boxes, we see that, even in presence of dividends, α is still given by

$$\alpha = \frac{\partial F}{\partial S} = \Delta_t. \quad (3.133)$$

Substituting (3.133) in (3.132), the two pink terms cancel out and we get the following PDE:

$$\frac{\partial F}{\partial t} dt + rS \frac{\partial F}{\partial S} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 F}{\partial S^2} - \tilde{\delta} \frac{\partial F}{\partial S} - rF = 0. \quad (3.134)$$

Financially speaking, it is not reasonable that $\tilde{\delta}$ is constant, because the rate of dividends should depend on the value of the asset. The typical assumption is that $\tilde{\delta}$ is given by a constant q times the value of the asset at time t :

$$\tilde{\delta} = qS_t. \quad (3.135)$$

In this case we speak about *proportional dividends*, and the corresponding PDE reads

$$\begin{cases} \frac{\partial F}{\partial t} dt + (r - q)S \frac{\partial F}{\partial S} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 F}{\partial S^2} - rF = 0 \\ F(S_T, T) = \text{payoff}_T. \end{cases} \quad (3.136)$$

Now we apply the Feynman-Kac methodology in order to say that if F is a solution of this PDE then there exist a probability space and a probability measure \mathbb{Q}^q under which we can define the B&S equation

$$\frac{dS}{S} = (r - q)dt + \sigma dW_T^{\mathbb{Q}^q}. \quad (3.137)$$

Such solution of the PDE can be written as

$$F(S_t, t) = e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}^q} [\text{payoff}_T]. \quad (3.138)$$

Recall that S_t is the solution of the geometric Brownian motion equation (3.137). In presence of dividends we have:

$$S_t = S_0 e^{(r-q-\frac{1}{2}\sigma^2)T+\sigma W_T^{\mathbb{Q}^q}} \quad (3.139)$$

$$= (S_0 e^{-qT}) e^{(r-\frac{1}{2}\sigma^2)T+\sigma W_T^{\mathbb{Q}^q}}. \quad (3.140)$$

Notice also that, since we can write $(S_0 e^{-qT})$ as a new initial value, we have that

$$F_q(S_t, t) = F_0(S_t e^{-q(T-t)}). \quad (3.141)$$

Now we are ready to write the Black & Scholes formula in presence of dividends:

$$\text{price}^{\text{Call}}(t, T, S_t, K, \sigma, r, q) = S_t e^{-q(T-t)} \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2) \quad (3.142)$$

with

$$d_1 = \frac{\ln \frac{S_t e^{-q(T-t)}}{K} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} = \frac{\ln \frac{S_t}{K} + (r - q + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \quad (3.143)$$

$$d_2 = d_1 - \sigma\sqrt{T-t}. \quad (3.144)$$

Using the put-call parity we immediately get the price of a put option.

3.10 Multi-dimensional market

Lesson 15 of
10/04/2020

Let us assume that we have n a priori given risky assets (“stocks”) with price processes $S_1(t), \dots, S_n(t)$ and a riskless asset with price process $S_0(t)$, which evolves according to a certain interest rate:

$$\frac{dS_0(t)}{S_0(t)} = r_t dt. \quad (3.145)$$

The entire asset price vector is denoted by $S(t)$, and in matrix notation we will write it as a column vector:

$$S(t) = (S_0(t), S_1(t), \dots, S_n(t))^T. \quad (3.146)$$

According to the B&S model, the individual risky asset dynamics follows the equation:

$$\frac{dS_i(t)}{S_i(t)} = \mu_i dt + \sum_{j=1}^d \sigma_{ij} dW_j^{\mathbb{P}}(t), \quad (3.147)$$

where $W^{\mathbb{P}}(t) \in \mathbb{R}^d$ is a standard Brownian motion, i.e. $W_i \perp \!\!\! \perp W_j \ \forall i \neq j$. Here we assume that r_t, μ_i and σ_{ij} are adapted processes to the corresponding filtration (that we assume generated by the Brownian motion itself). Notice that the number n

of assets is not necessarily equal to the number of Brownian motions d driving the assets, so $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^n$.

Let's give an intuition of what we want to do. Consider, for example, a two assets model ($n = 2$), that is a model with two risky assets and one riskless asset. Assume that μ, σ and r are constant. The dynamics of risky assets is given by

$$\frac{dS_1(t)}{S_1(t)} = \mu_1 dt + \sum_{j=1}^d \sigma_{1j} dW_j^{\mathbb{P}}(t) \quad (3.148)$$

$$\frac{dS_2(t)}{S_2(t)} = \mu_2 dt + \sum_{j=1}^d \sigma_{2j} dW_j^{\mathbb{P}}(t). \quad (3.149)$$

If $d = 1$ both assets are driven by the same Brownian motion, i.e. they are perfectly correlated (the source of randomness is the same). However, if the drifts are different, there can be a potential presence of an arbitrage opportunity (we can discount with respect to the volatility and the corresponding rate of growth may be different). If $d > 2$ we can always reduce the number of Brownian motions by normalizing the noise, in fact a linear combination of three Brownian motions can always be written as a combination of two or one Brownian motion(s).

Example 3.10.1: How to reduce the number of Brownian motions

If $d = 2$, eq. (3.148) become:

$$\frac{dS_1(t)}{S_1(t)} = \mu_1 dt + \sigma_{11} dW_1^{\mathbb{P}}(t) + \sigma_{12} dW_2^{\mathbb{P}}(t) \quad (3.150)$$

$$= \mu_1 dt + \sqrt{\sigma_{11}^2 + \sigma_{12}^2} d\widetilde{W}(t), \quad (3.151)$$

where

$$\widetilde{W} = \frac{1}{\sqrt{\sigma_{11}^2 + \sigma_{12}^2}} (\sigma_{11} W_1^{\mathbb{P}}(t) + \sigma_{12} W_2^{\mathbb{P}}(t)). \quad (3.152)$$

The quadratic variation of \widetilde{W} is

$$\langle \widetilde{W} \rangle_t = t \quad (3.153)$$

and this is the proof that \widetilde{W} is a true Brownian motion. So, one Brownian motion for process is enough.

So, we can re-write eq. (3.148) as:

$$\frac{dS_1(t)}{S_1(t)} = \mu_1 dt + \sigma_1 dW_1^{\mathbb{P}}(t) \quad (3.154)$$

$$\frac{dS_2(t)}{S_2(t)} = \mu_2 dt + \sigma_2 dW_2^{\mathbb{P}}(t). \quad (3.155)$$

We know that $dW_1^{\mathbb{P}}$ and $dW_2^{\mathbb{P}}$ are independent. This means that the two assets are independent, which is not a very realistic assumption – statistically speaking. So, we have to introduce some kind of correlation.

Let's consider a triangular construction involving, asset after asset, a new Brownian motion:

$$\frac{dS_1(t)}{S_1(t)} = \mu_1 dt + \sigma_1 dW_1^{\mathbb{P}}(t) \quad (3.156)$$

$$\frac{dS_2(t)}{S_2(t)} = \mu_2 dt + \sigma_2 \underbrace{(\rho dW_1^{\mathbb{P}}(t) + \sqrt{1 - \rho^2} dW_2^{\mathbb{P}})}_{\langle \cdot \rangle_t = t}. \quad (3.157)$$

where ρ represents the correlation between the two Brownian motions (i.e. between the two assets). This approach has two advantages: first, there is the possibility to decompose the matrix σ using the Cholesky decomposition⁴; second, it allows to introduce additional assets without affecting the evolution of the previous ones.

3.10.1 Absence of arbitrage

Recall that the absence of arbitrage is related to the existence of a probability measure \mathbb{Q} under which all the assets are martingales:

$$S_t^i e^{-\int_0^t r_s ds} = \mathbb{Q} - \text{martingale}.$$

Recall also that we found that in order to go from the Brownian motion under the historical probability measure \mathbb{P} to the one under the risk neutral probability measure \mathbb{Q} we just have to make a shift involving a finite variation process $\lambda \in \mathbb{R}^d$:

$$dW_t^{\mathbb{P}} \longrightarrow dW_t^{\mathbb{Q}} + \lambda(t)dt.$$

The change of measure is related to the Girsanov theorem through the Radon-Nikodym derivative:

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_t = \exp \left(\int_0^t \lambda_s^T \cdot dW_s^{\mathbb{P}} - \frac{1}{2} \int_0^t \|\lambda_s\|^2 ds \right). \quad (3.158)$$

The fact that

$$\mu + \sigma\lambda = (r, \dots, r)^T \in \mathbb{R}^n \quad (3.159)$$

means that, after the change of measure, the drifts μ_i are transformed into drifts which are equal to r for all the components. μ , σ , and r are given a priori and we want to solve for λ . Therefore, for each t , the n -dimensional vector $\mu(t) - r(t)$ must be in the image of the diffusion matrix $\sigma(t)$. This leads to the following result.

Proposition 3.10.1. *A necessary condition for the absence of arbitrage is that*

$$\mu(t) - r(t) \in \text{Im}\{\sigma\} \quad (3.160)$$

with probability one for each t .

We then have the following central result.

Proposition 3.10.2. *The model is generically arbitrage free (arbitrage free for every (sufficiently integrable) choice of μ) if and only if, for each $t \leq T$ and \mathbb{P} -a.s., the mapping*

$$\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^n \quad (3.161)$$

is surjective, i.e. if and only if the volatility matrix $\sigma(t)$ has rank n .

From this result it follows that in order to have absence of arbitrage we must necessarily have $d \geq n$, i.e. we must have at least as many independent Wiener processes as the number of risky assets (as we said before).

⁴ The Cholesky decomposition of a Hermitian positive-definite matrix A is a decomposition of the form $A = LL^*$ where L is a lower triangular matrix with real and positive diagonal entries, and L^* denotes the conjugate transpose of L . Every Hermitian positive-definite matrix (and thus also every real-valued symmetric positive-definite matrix) has a unique Cholesky decomposition.

3.10.2 Completeness

We now go on to obtain conditions for the model to be complete. In order to avoid pathological cases we assume that the model is generically arbitrage free. From the Second Fundamental Theorem 3.4.3 we know that the model is complete if and only if the martingale measure is unique, so it is tempting to draw the conclusion that we have completeness if and only if equation (3.159) has a unique solution, i.e. if and only if the condition

$$\ker(\sigma) = \{0\} \quad (3.162)$$

is satisfied for all t and with probability one. This is, however, not quite true. The reason is that, in case of a general filtered probability space, there is no guarantee that all equivalent measure transformations are of the Girsanov type above. In a general situation, where there are other sources of randomness beside the Wiener process W , the Girsanov transformation above will only change the measure for the Wiener process, but it will not affect the other processes. So, in the presence of jumps, we cannot replicate all the contingent claims.

In order to obtain sharp results we are therefore forced to make the assumption that all randomness in our model is generated by the Wiener process W .

Theorem 3.10.1. *Assume that the model is generically arbitrage free and that the filtration \mathcal{F} is defined by*

$$\mathcal{F}_t = \mathcal{F}_t^W, \quad (3.163)$$

i.e. there are no jumps. Then, disregarding integrability problems, the market is complete if and only if $k = n$ and the volatility matrix $\sigma(t)$ is invertible, i.e. $\ker(\sigma) = \{0\}$ for each $t \leq T$.

Proof. The martingale measure is unique if and only if the solution λ of the “martingale measure equation”

$$\sigma\lambda = r - \mu$$

is unique, and this occurs if and only if $\sigma(t)$ is injective, which implies $d \leq n$. But since we have assumed generic absence of arbitrage, we know that $n \geq d$ and that $\sigma(t)$ is surjective. Thus $d = n$ and $\sigma(t)$ is invertible.

Alternatively, from the fact that $\text{Im}\{\sigma\} = \mathbb{R}^n$ and $\text{Im}\{\sigma^T\} = \mathbb{R}^d$ we have that

$$\ker(\sigma) = (\text{Im}\{\sigma^T\})^\perp = \{0\}.$$

Thus $\sigma(t)$ is invertible and $d = n$. □

Therefore if we want completeness the number of Brownian motions must be equal to the number of assets.

3.10.3 The pricing problem with two risky assets in the Black & Scholes model

The dynamics of the assets under the neutral probability measure \mathbb{Q} is given by:

$$\frac{dS_1(t)}{S_1(t)} = r dt + \sigma_{11} dW_1^{\mathbb{Q}}(t) + \sigma_{12} dW_2^{\mathbb{Q}}(t) \quad (3.164)$$

$$\frac{dS_2(t)}{S_2(t)} = r dt + \sigma_{21} dW_1^{\mathbb{Q}}(t) + \sigma_{22} dW_2^{\mathbb{Q}}(t), \quad (3.165)$$

where in the first asset we keep two Brownian motions in order to stick with the Bjork's book notation.

We want to price a contract which payoff $\Phi(S_1(t), S_2(t))$ depends on both assets (S_1, S_2) . According the risk neutral methodology, the price is given by the discounted value up to maturity of the conditional expected value of the payoff under the probability measure \mathbb{Q} :

$$price_t = e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}}[\Phi(S_1(t), S_2(t))] \quad (3.166)$$

$$= F(t, S_1, S_2), \quad (3.167)$$

where in we used the Feynman-Kac formula, as the payoff depends only on the terminal value of the underlying. F solves the PDE:

$$\begin{cases} \frac{\partial F}{\partial t} + r \left(x_1 \frac{\partial F}{\partial x_1} + x_2 \frac{\partial F}{\partial x_2} \right) + \frac{1}{2} \left(x_1^2 \frac{\partial^2 F}{\partial x_1^2} c_{11} + 2x_1 x_2 \frac{\partial^2 F}{\partial x_1 \partial x_2} c_{12} + x_2^2 \frac{\partial^2 F}{\partial x_2^2} c_{22} \right) - rF = 0 \\ F(t, x_1, x_2) = \Phi(x_1(t), x_2(t)), \end{cases} \quad (3.168)$$

where from the Itô's formula we have the coefficients:

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \sigma \sigma^T = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}. \quad (3.169)$$

Example 3.10.2: Exchange option

Consider an *exchange option*, which gives the holder the right, but not the obligation, to exchange one S_2 share for one S_1 share at time T . So, the payoff is

$$\Phi(S_1(T), S_2(T)) = (S_1(T) - S_2(T))^+, \quad (3.170)$$

which is similar to the call option payoff if we see S_2 as the strike. Very naively, we could think to apply the B&S formula to S_1 with strike S_2 :

$$price_t^{EXC} = S_1(t) \mathcal{N}(d_1) + e^{-r(T-t)} S_2(T) \mathcal{N}(d_2). \quad (3.171)$$

This is wrong, because, at time t , $S_2(T)$ is not a measurable quantity. Then we could think to adapt the formula by using $S_2(t)$, but this is an arbitrary choice. Moreover, if we use this adaptation on d_1 or d_2 , we have (for example for d_1)

$$d_1 = \frac{\ln \frac{S_1(t)}{S_2(t)} + (r + \frac{1}{2}(\sigma_{11}^2 + \sigma_{12}^2))(T-t)}{\sqrt{(\sigma_{11}^2 + \sigma_{12}^2)(T-t)}}, \quad (3.172)$$

so the price of this option will depend only on the volatility of the first asset (there is no σ_{21} or σ_{22}). This is not reasonable because the second asset plays a role in the price and if the two assets are strongly correlated we cannot expect the same price as if the two assets were poorly correlated or strongly anti-correlated. We need to start from scratch.

Let us ask ourselves: which is the role of the correlation between S_1 and S_2 ? Is it possible to reduce the dimension of the problem (of the PDE)? Let's start by rewriting the payoff as

$$(S_1(T) - S_2(T))^+ = S_2(T) \left(\frac{S_1(T)}{S_2(T)} - 1 \right)^+ = F(t, x_1, x_2) \equiv x_2 G(T, z), \quad (3.173)$$

where we denoted $\frac{x_1}{x_2} \equiv z$. In this case we obtain the payoff of a true call, $\left(\frac{S_1(T)}{S_2(T)} - 1 \right)^+$, where $K = 1$. If the payoff has this property, maybe also the pricing function has it. Now we look for the PDE satisfied by G . We have:

$$\frac{\partial F}{\partial t} = x_2 \frac{\partial G}{\partial t},$$

$$\begin{aligned}
\frac{\partial F}{\partial x_1} &= x_2 \frac{\partial G}{\partial z} \frac{1}{x_2} = \frac{\partial G}{\partial z}, \\
\frac{\partial F}{\partial x_2} &= G + x_2 \frac{\partial G}{\partial z} \left(-\frac{x_1}{x_2^2} \right) = G - \frac{x_1}{x_2} \frac{\partial G}{\partial z}, \\
\frac{\partial^2 F}{\partial x_1^2} &= \frac{1}{x_2} \frac{\partial^2 G}{\partial z^2}, \\
\frac{\partial^2 F}{\partial x_2^2} &= \left(\cancel{-\frac{x_1}{x_2^2}} \frac{\partial G}{\partial z} - \left(\cancel{-\frac{x_1}{x_2^2}} \frac{\partial G}{\partial z} + \left(\frac{x_1}{x_2} \frac{\partial^2 G}{\partial z^2} \left(-\frac{x_1}{x_2^2} \right) \right) \right) \right) = \frac{x_1^2}{x_2^3} \frac{\partial^2 G}{\partial z^2}, \\
\frac{\partial^2 G}{\partial x_1 \partial x_2} &= -\frac{x_1}{x_2^2} \frac{\partial^2 G}{\partial z^2}.
\end{aligned}$$

Substituting in eq. (3.168) we get:

$$\begin{aligned}
x_2 \frac{\partial G}{\partial t} + r \left(x_1 \frac{\partial G}{\partial z} + x_2 \left(G - \frac{x_1}{x_2} \frac{\partial G}{\partial z} \right) \right) + \\
+ \frac{1}{2} \left(\frac{x_1^2}{x_2} c_{11} \frac{\partial^2 G}{\partial z^2} + 2 \frac{x_1^2}{x_2} c_{12} + \frac{x_1^2}{x_2} c_{22} \frac{\partial^2 G}{\partial z^2} \right) - rx_2 G = 0,
\end{aligned}$$

where the highlighted terms cancel out. We are left with:

$$x_2 \frac{\partial G}{\partial t} + \frac{1}{2} \frac{x_1^2}{x_2} \frac{\partial^2 G}{\partial z^2} = 0.$$

Finally, by dividing by x_2 , we get:

$$\begin{cases} \frac{\partial G}{\partial t} + \frac{1}{2} z^2 \frac{\partial^2 G}{\partial z^2} = 0 \\ G(T, z) = (z - 1)^+. \end{cases} \quad (3.174)$$

This is the B&S PDE satisfied by the price of a call option with underlying z , interest rate $r = 0$, strike $K = 1$ and volatility $\sigma_z^2 = c_{11} + c_{22} - 2c_{12}$. So, we can write G using the B&S formula:

$$G(t, z) = z(t)\mathcal{N}(d_1) - e^{rt}\mathcal{N}(d_2) \quad (3.175)$$

and then $F = G/x_2$:

$$F(t, x_1, x_2) = x_1(t)\mathcal{N}(d_1) - x_2\mathcal{N}(d_2). \quad (3.176)$$

The crucial difference between this solution and the naive one is that now the definition of d_1 involves the whole volatility matrix:

$$\begin{aligned}
d_1 &= \frac{\ln z(t) - \frac{1}{2}\sigma_z^2(T-t)}{\sigma_z \sqrt{T-t}} \\
&= \frac{\ln \frac{S_1(t)}{S_2(t)} - \frac{1}{2}(c_{11} + c_{22} - 2c_{12})(T-t)}{\sqrt{(c_{11} + c_{22} - 2c_{12})(T-t)}}.
\end{aligned} \quad (3.177)$$

3.11 Change of numéraire

The change of numeraire is an alternative technique. We know that under the original probability measure \mathbb{P} one asset evolves according to the equation

$$\begin{cases} \frac{dS(t)}{S(t)} = \rho dt + \sigma dW^{\mathbb{P}}(t) \\ \frac{dS_0(t)}{S_0(t)} = rdt. \end{cases} \quad (3.178)$$

Under the risk neutral probability measure \mathbb{Q} the risky asset changes, while the riskless asset doesn't:

$$\begin{cases} \frac{dS}{S} = rdt + \sigma dW^{\mathbb{Q}}(t) \\ \frac{dS_0(t)}{S_0(t)} = rdt. \end{cases} \quad (3.179)$$

Moreover, we know that under \mathbb{Q} the discounted asset $\hat{S}_t = e^{-rt} S_t$ is a \mathbb{Q} -martingale and the discounted numéraire is $\hat{S}_0 \equiv 1$. This shift from \mathbb{P} to \mathbb{Q} starts from $(S(t), S_0(t))$ and leads to $(S(t)/S_0(t), 1)$. This is only one possibility.

In general, for any numéraire, provided that it is a tradable asset with positive price, there exists a probability measure \mathbb{Q}^{NUM} such that any discounted asset with respect to it becomes a \mathbb{Q}^{NUM} -martingale. This means that:

- $\frac{\text{assets}}{\text{numéraire}} = \mathbb{Q}^{\text{NUM}}$ -martingales;
- $\frac{\text{portfolios}}{\text{numéraire}} = \mathbb{Q}^{\text{NUM}}$ -marginalgales;
- $\frac{\text{option prices}}{\text{numéraire}} = \mathbb{Q}^{\text{NUM}}$ -marginalgales;

From the last point, in general we have:

$$\frac{\text{price}_t(\text{payoff}_T)}{\text{NUM}_T} = \mathbb{E}_t^{\mathbb{Q}^{\text{NUM}}} \left[\frac{\text{payoff}_T}{\text{NUM}_T} \right], \quad (3.180)$$

so we can recover the pricing function:

$$\text{price}_t(\text{payoff}_T) = \text{NUM}_T \mathbb{E}_t^{\mathbb{Q}^{\text{NUM}}} \left[\frac{\text{payoff}_T}{\text{NUM}_T} \right]. \quad (3.181)$$

If the numéraire is the riskless asset we can extract the NUM inside the expected value, recovering the usual formula. If it is different – in particular if it is stochastic – we have to keep it inside the expected value.

3.11.1 Application to the exchange option

We already know that the payoff of the exchange option is $(S_1(T) - S_2(T))^+$. Under the probability measure \mathbb{P} the dynamics is given by:

$$\frac{dS_1(t)}{S_1(t)} = \mu_1 dt + \sigma_{11} dW_1^{\mathbb{P}}(t) + \sigma_{12} dW_2^{\mathbb{P}}(t) \quad (3.182)$$

$$\frac{dS_2(t)}{S_2(t)} = \mu_2 dt + \sigma_{21} dW_1^{\mathbb{P}}(t) + \sigma_{22} dW_2^{\mathbb{P}}(t). \quad (3.183)$$

Let's use S_2 as numéraire:

$$(S_1, S_2) \longrightarrow (S_1/S_2, 1). \quad (3.184)$$

We have to look for the probability measure \mathbb{Q}^{S_2} under which S_1/S_2 becomes a martingale. If we define

$$\frac{S_1}{S_2} := Z_t \quad (3.185)$$

the dynamics of Z_t must involve only a drift. Using the Itô's formula we have to compute:

$$dZ_t = d\left(\frac{S_1}{S_2}\right) = d\left(S_1 \frac{1}{S_2}\right). \quad (3.186)$$

Let's start from

$$\begin{aligned} d\left(\frac{1}{S_2}\right) &= -\frac{1}{S_2^2}dS_2 + \frac{1}{S_2^3}(dS_2)^2 \\ &= -\frac{1}{S_2}(\mu_2 + \sigma_{21}dW_1^{\mathbb{P}} + \sigma_{22}dW_2^{\mathbb{P}}) + \frac{1}{S_2} \frac{S_2^2}{S_2^2}(\sigma_{21}^2 + \sigma_{22}^2)dt. \end{aligned}$$

Then

$$\frac{d\left(\frac{1}{S_2}\right)}{\frac{1}{S_2}} = (-\mu_2 + \sigma_{21} + \sigma_{22}^2)dt + \sigma_{21}^2 dW_1^{\mathbb{P}} + \sigma_{22}^2 dW_2^{\mathbb{P}}.$$

Recall that the dynamics of S_2^α is still a Brownian motion, in fact we can write $S_2(t) = S_2(0)e^N$ and if we consider a power the structure remains the same. Now let's plug what we found into (3.186) and apply the integration by parts:

$$\begin{aligned} dZ_t &= \frac{1}{S_2}dS_1 + S_1 d\left(\frac{1}{S_2}\right) + d\left\langle S_1, \frac{1}{S_2} \right\rangle \\ &= (\dots)dt + \underbrace{\frac{S_1}{S_2}}_{Z_t} \left[(\sigma_{11} - \sigma_{21})dW_1^{\mathbb{P}} + (\sigma_{12} - \sigma_{22})dW_2^{\mathbb{P}} \right]. \end{aligned}$$

Let's focus only in the diffusive part (recall that, under the measure \mathbb{Q}^{S_2} , Z_t is a martingale). We can write:

$$\frac{dZ_t}{Z_t} = (\sigma_{11} - \sigma_{21})dW_1^{\mathbb{Q}^{S_2}} + (\sigma_{12} - \sigma_{22})dW_2^{\mathbb{Q}^{S_2}}. \quad (3.187)$$

The key point is that:

$$\frac{price_t(\text{payoff}_T)}{S_2(t)} = \mathbb{E}_t^{\mathbb{Q}^{S_2}} \left[\frac{(S_1(T) - S_2(T))^+}{S_2(T)} \right] \quad (3.188)$$

$$= \mathbb{E}_t^{\mathbb{Q}^{S_2}}[(Z(T) - 1)^+]. \quad (3.189)$$

This is exactly the problem of pricing a call option on Z with strike $K = 1$ and $r = 0$. In order to write the pricing function using the B&S formula, we need the value of the volatility corresponding to Z :

$$\sigma_Z = \sqrt{(\sigma_{11} - \sigma_{21})^2 + (\sigma_{12} - \sigma_{22})^2} = \sqrt{c_{11} + c_{22} - 2c_{12}}, \quad (3.190)$$

where $c = \sigma\sigma^T$. Finally, by applying the B&S formula, we get:

$$\frac{price_t(\text{payoff}_T)}{S_2(t)} = Z_t \Phi(d_1) - \Phi(d_2), \quad (3.191)$$

which – after some simplifications – can be written as

$$price_t(\text{payoff}_T) = S_1(t)\Phi(d_1) - S_2(t)\Phi(d_2) \quad (3.192)$$

with

$$d_1 = \frac{\ln Z(t) - \frac{1}{2}\sigma_z^2(T-t)}{\sigma_z \sqrt{(T-t)}}$$

$$= \frac{\ln \frac{S_1(t)}{S_2(t)} - \frac{1}{2}(c_{11} + c_{22} - 2c_{12})(T-t)}{\sqrt{(c_{11} + c_{22} - 2c_{12})(T-t)}}. \quad (3.193)$$

This is exactly what we found using the PDE, but in this case we did less computations.

The switch from \mathbb{P} to the martingale measure \mathbb{Q}^{S_2} is expressed through the Radon-Nikodym derivative:

$$\frac{d\mathbb{Q}^{S_2}}{d\mathbb{P}} \Big|_t = e^{\int_0^t \lambda^T \cdot dW - \frac{1}{2} \int_0^t \|\lambda\|^2 ds} \quad (3.194)$$

where λ is the shift:

$$dW_1^{\mathbb{P}}(t) \longrightarrow dW_1^{\mathbb{Q}^{S_2}}(t) + \lambda_1(t)dt \quad (3.195)$$

$$dW_2^{\mathbb{P}}(t) \longrightarrow dW_2^{\mathbb{Q}^{S_2}}(t) + \lambda_2(t)dt. \quad (3.196)$$

Remark 3.11.1. In the B&S formula (3.192) there is no discounting. This is not surprising, because the exchange option is like a bet of one asset with respect to another one, so there is no interest rate (in the standard call option we are betting an asset with respect to a deterministic/constant strike, so there must be an interest rate).

3.11.2 Application in the case of stochastic interest rates: Forward Measures

The starting point is, as always, the B&S PDE:

$$\begin{cases} \frac{dS}{S} = r_t dt + \sigma dW^{\mathbb{Q}}(t) \\ \frac{dS_0(t)}{S_0(t)} = r_t dt \end{cases} \quad (3.197)$$

where we suppose that r_t is an adapted stochastic process. In this case the riskless asset is no more deterministic, but locally deterministic. In fact, in principle, we can still solve the locally deterministic differential equation, which has solution

$$S_0(t) = e^{\int_0^t r_s ds}.$$

With the usual risk neutral approach, if we use the savings account (the riskless asset) as a numéraire,

$$(S, S_0) \longrightarrow (S/S_0, 1), \quad (3.198)$$

we have that S/S_0 is a \mathbb{Q} -martingale and

$$\frac{\text{price}_t(\text{payoff}_T)}{S_0(t)} = \mathbb{E}_t^{\mathbb{Q}} \left[\frac{\text{payoff}_T}{S_0(T)} \right]. \quad (3.199)$$

So, we have:

$$\text{price}_t(\text{payoff}_T) = e^{\int_0^t r_s ds} \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_0^T r_s ds} \text{payoff}_T \right] = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \text{payoff}_T \right]. \quad (3.200)$$

The problem is that the highlighted term is not \mathcal{F}_t -measurable, because it involves future realizations of the interest rate r . Moreover, the payoff typically involves the interest rate. So, the highlighted term and the payoff are not independent and – in general – the computation of this expected value is a very difficult problem.

We can solve this problem by changing the numéraire. In particular, we use the price of a zero-coupon bond as numéraire. Recall that the payoff of the zero-coupon bond is 1 and the price at time t is given by the risk neutral valuation.

Notation. We use $B(t, T)$ to indicate the price of one unit of money given at time T (i.e. the zero-coupon bond). A natural boundary condition is that $B(T, T) = 1$.

By the risk neutral valuation we have that:

$$B(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \cdot 1 \right] = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \right], \quad (3.201)$$

which is \mathcal{F}_t -measurable. Choosing the zero-coupon bond as numéraire we can consider the change of variables

$$(S(t), B(t, T)) \longrightarrow (S(t)/B(t, T), 1) \quad (3.202)$$

and we have that

$$\frac{S(t)}{B(t, T)} = \mathbb{Q}^T - \text{martingale}. \quad (3.203)$$

This particular risk neutral probability measure associated to the zero-coupon bond at maturity time T is called *forward risk neutral measure*. Now, we can apply the universal formula:

$$\frac{\text{price}_t(\text{payoff}_T)}{B(t, T)} = \mathbb{E}_t^{\mathbb{Q}^T} \left[\frac{\text{payoff}_T}{B(T, T)} \right], \quad (3.204)$$

which leads to

$$\begin{aligned} \text{price}_t(\text{payoff}_T) &= B(t, T) \mathbb{E}_t^{\mathbb{Q}^T} \left[\frac{\text{payoff}_T}{B(T, T)} \right] \\ &= \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \right] \mathbb{E}_t^{\mathbb{Q}^T} [\text{payoff}_T], \end{aligned} \quad (3.205)$$

where we use the fact that $B(T, T) = 1$. We end up with two conditional expected values:

- the first (in green), under the risk neutral probability measure \mathbb{Q} , is related to the interest rate market;
- the second, under the forward probability measure \mathbb{Q}^T , is related to the stock market with $r_t = 0 \forall t$.

Using this procedure we are able to disentangle the computation of the conditional expected value of the discount factor from the one of the payoff. By doing so, the randomness of the interest rate is all included in the green term. The price we pay is that we need to consider a new probability measure \mathbb{Q}^T – which is different from the risk neutral probability measure \mathbb{Q} – and now we have to understand which is the dynamics of the underlying under it.

Comparing (3.200) and (3.205) we have

$$\mathbb{E}_t^{\mathbb{Q}^T} [\text{payoff}_T] = \mathbb{E}_t^{\mathbb{Q}} \left[\text{payoff}_T \frac{e^{-\int_t^T r_s ds}}{B(t, T)} \right]. \quad (3.206)$$

From this equality we deduce that the highlighted term plays the role of the Radon-Nikodym derivative for the change of measure $\mathbb{Q} \rightarrow \mathbb{Q}^T$:

$$\frac{d\mathbb{Q}^T}{d\mathbb{Q}} \Big|_t = \frac{e^{-\int_t^T r_s ds}}{B(t, T)}. \quad (3.207)$$

Lesson 17 of
16/04/2020

Now, we have all we need to find the generic B&S formula with stochastic interest rate. For a call option

$$\text{payoff}_T = (S(T) - K)^+, \quad (3.208)$$

so the price is

$$\begin{aligned} \text{Call}_t &= \mathbb{E}_t^{\mathbb{Q}} \left[(S(T) - K)^+ e^{-\int_t^T r_s ds} \right] \\ &= \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} S(T) \mathbf{1}_{S_T \geq K} \right] - K \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \mathbf{1}_{S_T \geq K} \right] \\ &\stackrel{(a)}{=} \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} S(t) e^{\int_t^T r_s ds - \frac{\sigma^2}{2}(T-t) + \sigma(W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}})} \mathbf{1}_{S_T \geq K} \right] - K B(t, T) \mathbb{E}_t^{\mathbb{Q}} [\mathbf{1}_{S_T \geq K}] \\ &= S(t) \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\frac{\sigma^2}{2}(T-t) + \sigma(W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}})} \mathbf{1}_{S_T \geq K} \right] - K B(t, T) \mathbb{Q}^T (\mathbf{1}_{S_T \geq K}) \\ &\stackrel{(b)}{=} S(t) \mathbb{E}_t^{\mathbb{Q}^s} [S_T \geq K] - K B(t, T) \mathbb{Q}^T (S_T \geq K) \\ &= S(t) \mathbb{Q}^s (S_T \geq K) - K B(t, T) \mathbb{Q}^T (S_T \geq K), \end{aligned} \quad (3.209)$$

where in (a) we use the explicit expression for $S(T)$ and in order to split the second expected value we use the zero coupon bond as numéraire. In (b) we use the change of measure $\mathbb{Q} \rightarrow \mathbb{Q}^s$ associated to s 's numéraires, with corresponding Radon-Nikodym derivative

$$\frac{d\mathbb{Q}^s}{d\mathbb{Q}} \Big|_t = \frac{S(T)}{\mathbb{E}^{\mathbb{Q}}[S(T)]}. \quad (3.210)$$

Eq. (3.209) is the analogous of the B&S formula in the general case with stochastic interest rate (in the standard B&S formula we would have the gaussian CDFs). However, in principle we don't know the distributions of $S(t)$ under \mathbb{Q}^s and \mathbb{Q}^T .

Let's consider the particular case in which the volatility of the process $S(t)/B(t, T)$ is deterministic. In order to deal with the volatility we need to write the dynamics of the process in terms of a geometric Brownian motion-like form:

$$\frac{d \left(\frac{S(t)}{B(t, T)} \right)}{\frac{S(t)}{B(t, T)}} = (\dots) dt + \sigma dW. \quad (3.211)$$

As usual, if we introduce the auxiliary process

$$Z(t) := \frac{S(t)}{B(t, T)} \quad (3.212)$$

we have that:

$$\frac{dZ(t)}{Z(t)} = (\dots) dt + \sigma_Z(t) dW^{\mathbb{Q}}, \quad (3.213)$$

where σ_Z is deterministic. Notice that the denominator of $Z(t)$ is a possible numéraire. In particular, if we set the dynamics of $Z(t)$ under probability measure associated to B , i.e. the forward measure, then $Z(t)$ becomes a martingale. So, we have that

$$\frac{dZ(t)}{Z(t)} = 0 dt + \sigma_Z(t) dW^{\mathbb{Q}^T}. \quad (3.214)$$

The solution at time T is:

$$Z(T) = Z_0 \exp \left\{ -\frac{1}{2} \int_0^T \|\sigma_Z(t)\|^2 dt + \int_0^T \sigma_Z(t) \cdot dW^{\mathbb{Q}^T} \right\} \quad (3.215)$$

$$\stackrel{(a)}{=} \frac{S(0)}{B(0, T)} \exp \left\{ \mathcal{N} \left(-\frac{1}{2} \int_0^T \|\sigma_Z(t)\|^2 dt; \int_0^T \|\sigma_Z(t)\|^2 dt \right) \right\}, \quad (3.216)$$

where in (a) we used the fact that if σ_Z is deterministic then $\int_0^T \sigma_Z \cdot dW^{\mathbb{Q}^T}$ is gaussian (martingale) with expected value equal to zero. With that being said, the first term represents only a shift in the expected value. The variance of the Gaussian comes from the *Itô's isometry*⁵.

Now, we can compute the price of the call (3.209):

$$\text{Call}_T = S(t) \mathbb{Q}^s(S(T) > K) - KB(t, T) \mathbb{Q}^T(S(T) > K). \quad (3.217)$$

By using the fact that $B(T, T) = 1$, we can write:

$$\mathbb{Q}^T(S(T) > K) = \mathbb{Q}^T\left(\frac{S(T) > K}{1}\right) = \mathbb{Q}^T\left(\frac{S(T) > K}{B(T, T)}\right) = \mathbb{Q}^T(Z(T) > K) = \Phi(d_2)$$

with

$$d_2 = \frac{\ln \frac{S(t)}{B(t, T)K} - \frac{1}{2} \int_t^T \|\sigma_Z\|^2 ds}{\sqrt{\int_t^T \|\sigma_Z\|^2 ds}}. \quad (3.218)$$

Notice that if we consider a constant interest rate we have $B(t, T) = e^{-r(T-t)}$ and we recover the B&S formula for d_2 .

Then, we can write:

$$\mathbb{Q}^s(S(T) \geq K) = \mathbb{Q}^s\left(\frac{1}{S(T)} \leq \frac{1}{K}\right) = \mathbb{Q}^s\left(\frac{B(T, T)}{S(T)} \leq \frac{1}{K}\right) = \mathbb{Q}^s\left(\frac{1}{Z(T)} \leq \frac{1}{K}\right), \quad (3.219)$$

where

$$\frac{1}{Z(T)} = \frac{B(T, T)}{S(T)} = \mathbb{Q}^s - \text{martingale} \quad (3.220)$$

so that

$$\frac{d\frac{1}{Z(T)}}{\frac{1}{Z(t)}} = 0 dt + \sigma_{1/Z} dW^{\mathbb{Q}^s}(t) = -\sigma_Z dW^{\mathbb{Q}^s}(t). \quad (3.221)$$

This equation is a Brownian motion differential equation, which has solution:

$$\frac{1}{Z(T)} = \frac{1}{Z(0)} \exp\left\{-\frac{1}{2} \int_0^T \|\sigma_Z(t)\|^2 dt - \int_0^T \sigma_Z(t) \cdot dW^{\mathbb{Q}^T}\right\} \quad (3.222)$$

$$= \frac{B(0, T)}{S(0)} \exp\left\{\mathcal{N}\left(-\frac{1}{2} \int_0^T \|\sigma_Z(t)\|^2 dt; \int_0^T \|\sigma_Z(t)\|^2 dt\right)\right\}. \quad (3.223)$$

Now we can come back to eq. (3.219):

$$\mathbb{Q}^s(S(T) \geq K) = \mathbb{Q}^s\left(\frac{1}{Z(T)} \leq \frac{1}{K}\right) = \Phi(d_1) \quad (3.224)$$

⁵ \wedge Itô's isometry states that if $H \in \mathcal{H}$, then

$$\mathbb{E}\left[\left(\int_0^T H_s dW_s\right)^2\right] = \mathbb{E}\left[\left(\int_0^T H_s^2 ds\right)^2\right]. \quad (\diamond)$$

Then:

$$\text{Var}\left(\int_0^T \sigma_Z(s) \cdot dW^{\mathbb{Q}^T}\right) = \mathbb{E}\left[\left(\int_0^T \sigma_Z(s) \cdot dW^{\mathbb{Q}^T}\right)^2\right] \stackrel{(\diamond)}{=} \mathbb{E}\left[\int_0^T \|\sigma_Z(s)\|^2 ds\right] = \int_0^T \|\sigma_Z(s)\|^2 ds.$$

In other words, the relationship between deterministic and stochastic integrals is not linear, but quadratic.

where

$$d_1 = d_2 + \sqrt{\int_0^T \|\sigma_Z(t)\|^2 dt}. \quad (3.225)$$

Finally, we conclude that, if the volatility of the process $S(t)/B(t, T)$ is constant, the price of a call option is:

$$\text{Call}_t = S(t)\Phi(d_1) - KB(t, T)\Phi(d_2). \quad (3.226)$$

Eq. (3.226) has the same shape as the B&S formula and the only difference is in the expression of d_1 and d_2 .

However, we have still to specify the dynamics of the interest rate.

Example 3.11.1: Vašíček model

The Vašíček model (1979) is the first model that described the evolution of interest rates. It considers the *instantaneous (spot) interest rate* r_t corresponding to the infinitesimal interval $[t, t + dt]$ and the main assumption is that:

$$dr_t = (a - br_t)dt + \sigma_r dW^Q(t) \quad (3.227)$$

where W^Q is a Brownian motion (under the risk neutral probability measure \mathbb{Q}) that models the continuous inflow of randomness into the system. The typical parameters can be quickly characterized as follows:

- $b > 0$ represents the “long term mean level”, i.e. all future trajectories of r will evolve around a mean level b in the long run;
- $a > 0$ characterizes the “speed of reversion”, i.e. the velocity at which such trajectories will regroup around b in time;
- $\sigma_r > 0$ represents the “instantaneous volatility”, which measures instant by instant the amplitude of randomness entering the system. Higher σ_r implies more randomness.

The starting interest rate r_0 is not observable, since it is instantaneous (the short interest rate is an abstraction). This mean that in this model the interest rate is not tradable.

Let's see why this is the dynamics. If we suppose that $\sigma_r = 0$ (deterministic dynamics), then eq. (3.227) becomes:

$$dr_t = (a - br_t)dt \quad \Rightarrow \quad \dot{r}_t = b \left(\frac{a}{b} - r_t \right) \quad (3.228)$$

The solution r_t of this differential equation will have an exponential behavior with convergence to the asymptotic value r_∞ . The presence of Brownian noise induces some fluctuations around the deterministic trajectory, as shown in Figure 16.

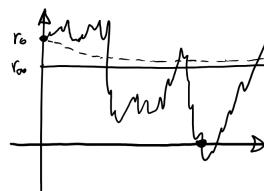


Figure 3.16: Dynamics of the interest rate r_t with initial condition r_0 and asymptotic value r_∞ .

In order to find the solution of eq. (3.227), we consider the following differential:

$$\begin{aligned} d(r_t e^{bt}) &= e^{bt} dr_t + r_t d(e^{bt}) \\ &= e^{bt}(a - br_t) dt + e^{bt} \sigma_r dW^{\mathbb{Q}} + br_t e^{bt} dt \\ &= ae^{bt} dt + \sigma_r e^{bt} dW^{\mathbb{Q}}(t) \end{aligned} \quad (3.229)$$

Integrating from 0 to t , we get:

$$r_t e^{bt} - r_0 = \frac{a}{b}(e^{bt} - 1) + \sigma_r \int_0^t e^{bs} dW^{\mathbb{Q}}(s) \quad (3.230)$$

from which:

$$r_t = r_0 e^{-bt} + \frac{a}{b}(1 - e^{-bt}) + \sigma_r \int_0^t e^{-b(t-s)} dW^{\mathbb{Q}}(s) \quad (3.231)$$

It is very simple to check that $e^{-b(t-s)} \in \mathcal{H}$, so the integral is a martingale, in particular a gaussian variable $\mathcal{N}\left(0; \sigma_r^2 \int_0^t e^{-2b(t-s)} ds\right) = \mathcal{N}\left(0; \frac{\sigma_r^2}{2b} (1 - e^{-2bt})\right)$. The deterministic part results in a shift in the mean, so we end up with:

$$r_t \sim \mathcal{N}\left(r_0 + \frac{a}{b}(1 - e^{-bt}); \frac{\sigma_r^2}{2b} (1 - e^{-2bt})\right) \quad (3.232)$$

It turns out that in the Vašíček model the interest rate is a gaussian variable. In particular, $\mathbb{Q}(r_t < 0) > 0$. In the past this fact represented a problem, because the short interest rate was supposed to be strictly positive. This convention was dropped after the 2008 crisis.

The simpler product to price in this framework is the zero-coupon bond. We would like to check if the volatility of the process $S(t)/B(t, T)$ is deterministic or not. In order to do that, we try to limit ourselves to the minimal needed information about $B(t, T)$.

If we fix $s \geq t$ as starting point, we have:

$$r_s = r_t e^{-b(s-t)} + \frac{a}{b}(1 - e^{-b(s-t)}) + \sigma_r \int_t^s e^{-b(s-u)} dW^{\mathbb{Q}}(u). \quad (3.233)$$

The price of the zero-coupon bond is given by:

$$B(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \right]. \quad (3.234)$$

First, we compute the integral:

$$\begin{aligned} - \int_t^T r_s ds &= -r_t \int_t^T e^{-b(s-t)} ds - \frac{a}{b} \int_t^T (1 - e^{-b(s-t)}) ds - \\ &\quad - \sigma_r \int_t^T \left(\int_t^s e^{-b(s-u)} dW_u^{\mathbb{Q}} \right) du. \end{aligned} \quad (3.235)$$

The first two integrals are easy to compute, while the third is feasible thanks to the stochastic Fubini theorem. Since there is no dependence on the interest rate, the third integral will result in a Gaussian variable $\mathcal{N}(0; \Sigma^2)$ with zero mean and variance given by the Itô's isometry. If we solve the first integral and we consider the second integral as a shift in the mean, we end up with:

$$- \int_t^T r_s ds = -\frac{r_t}{b}(e^{-b(T-t)} - 1) - \mathcal{N}(m; \Sigma^2) \quad (3.236)$$

Plugging in eq. (3.234) we get:

$$B(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\frac{r_t}{b}(e^{-b(T-t)} - 1)} e^{-\mathcal{N}(m; \Sigma^2)} \right] \quad (3.237)$$

By Itô, the dynamic of the zero-coupon bond is:

$$\begin{aligned} dB(t, T) &= \frac{\partial B(t, T)}{\partial t} dt + \frac{\partial B(t, T)}{\partial r_t} dr_t + \frac{1}{2} \frac{\partial^2 B(t, T)}{\partial r_t^2} d\langle r \rangle_t \\ &\stackrel{(3.227)}{=} (\dots) dt + B(t, T) \frac{e^{-b(T-t)} - 1}{b} ((a - br_t) dt + \sigma_r dW^{\mathbb{Q}}) \\ &\stackrel{(a)}{=} (\dots) dt + B(t, T) \frac{e^{-b(T-t)} - 1}{b} \sigma_r dW^{\mathbb{Q}} \end{aligned} \quad (3.238)$$

and then:

$$\frac{dB(t, T)}{B(t, T)} = (\dots) dt + \frac{e^{-b(T-t)} - 1}{b} \sigma_r dW^{\mathbb{Q}} \quad (3.239)$$

So, the volatility of $B(t, T)$ in the Vašíček model is given by:

$$\sigma_B = \frac{e^{-b(T-t)} - 1}{b} \quad (3.240)$$

Now, if we consider the ratio

$$\frac{d\frac{S(t)}{B(t, T)}}{\frac{S(t)}{B(t, T)}} = \left(\sigma - \sigma_r \frac{e^{-b(T-t)} - 1}{b} \right) dW^{\mathbb{Q}} + (\dots) dt \quad (3.241)$$

we have that the highlighted term is deterministic, so it is possible to find the B&S formula for the price of a call in the Vašíček model. If we denote

$$\sigma - \sigma_r \frac{e^{-b(T-t)} - 1}{b} \equiv \sigma_z \quad (3.242)$$

then we have

$$\text{Call}_t = S(t)\Phi(d_1) - KB(t, T)\Phi(d_2) \quad (3.243)$$

with

$$d_1 = \frac{\ln \frac{S(t)}{B(t, T)K} - \frac{1}{2} \int_t^T \|\sigma_Z\|^2 ds}{\sqrt{\int_t^T \|\sigma_Z\|^2 ds}} \quad (3.244)$$

Notice that, as expected, the zero-coupon bond becomes deterministic at the maturity (it gives one unit of money), in fact for $t = T$ eq. (3.239) becomes

$$\frac{dB(t, T)}{B(t, T)} = (\dots) dt + \frac{e^{-b(T-T)} - 1}{b} \sigma_r dW^{\mathbb{Q}} = (\dots) dt.$$

In conclusion:

- in this particular model we are able to find the price for the zero-coupon bond (but we are not really interested in it), ending up with $B(t, T)$ having a log-normal distribution.

- We focused on the dependence of the stochastic quantity r_t , because in order to find out which is the dynamics of the zero-coupon bond we need the coefficient of the Brownian motion and according to Itô's formula this coefficient is only related to dr_t .
- We found the volatility of the zero-coupon bond and then we checked that the volatility of the process $S(t)/B(t, T)$ is deterministic. Then, if the volatility is deterministic, we can write down an explicit B&S formula with CDFs $\Phi(d_1), \Phi(d_2)$. Is it surprising? No, because in general we know that:

$$S(t) = S(0) \exp \left\{ \int_0^T r_s ds - \frac{1}{2} \sigma^2 t + \sigma dW^{\mathbb{Q}}(t) \right\}$$

In the Vašíček model r_t is gaussian, so $\int_0^T r_s ds$ is gaussian too, together with $dW^{\mathbb{Q}}(t)$. This means that $S(t)$ is log-normal distributed and that we can adopt the B&S model.

3.12 Black formula

The B&S formula gives the price of a call option on an asset. The purpose of this section is to derive the Black formula for the price of call options on a *futures* or *forward contract*.

Lesson 18 of
17/04/2020

3.12.1 Forward contracts

Recall the main features of the forward contract:

- it gives us the obligation to buy the asset at a fixed price K which is fixed at the signature date;
- the payoff is $(S_T - K)$;
- the strike K is such that at the signature time the value of the contract is zero (it is cost-less to enter in a forward contract). Furthermore, we can introduce the notion of *forward price* $K(t, T)$, which is the strike price decided at time t such that the value of the “new contract” starting at time t is zero. In other words, it represents an “updated baricenter”;
- using the absence of arbitrage, we found that if the interest rate is constant then

$$K(0, T) = S(0)e^{rT} \quad \Rightarrow \quad K(t, T) = S(t)e^{r(T-t)}. \quad (3.245)$$

If the interest rate is stochastic we have to consider a change of numéraire.

- if the strike $K(0, T) = K$ is the fair one, the price is given by:

$$\text{price}_t(\text{payoff}_{[0, T]}) = \begin{cases} 0 & t = 0 \\ \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} (S(T) - K(0, T)) \right] & 0 < t < T \\ S(T) - K(0, T) & t = T \end{cases} \quad (3.246)$$

Let's compute the expected value in (3.246) for $0 < t < T$:

$$\begin{aligned} (3.246) &= e^{-\int_t^T r_s ds} \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_0^t r_s ds} S(T) \right] - K(0, T) B(t, T) \\ &\stackrel{(a)}{=} e^{-\int_0^t r_s ds} e^{-\int_0^t r_s ds} S(t) - K(0, T) B(t, T) \\ &= S(t) - K(0, T) B(t, T) \end{aligned}$$

where in (a) we used the fact that the discounted asset $S(t)e^{-\int_0^t r_s ds}$ is a true \mathbb{Q} -martingale.

If the interest rate is flat, $r_t = r \forall t$, we have:

$$price_t(S(T) - K(0, T)) = S(t) - K(0, T) e^{-r(T-t)} \geq 0 \quad (3.247)$$

which can be positive or negative. Introducing the expression of the forward strike at time $t = 0$,

$$K(0, T) = S(0) e^{rT},$$

we can rewrite (3.247) as

$$price_t(S(T) - K(0, T)) = S(t) - K(0, T) e^{-r(T-t)} \quad (3.248)$$

$$\begin{aligned} &= S(t) - S(0) e^{rT} e^{-r(T-t)} \\ &= S(t) - S(0) e^{rt} \end{aligned} \quad (3.249)$$

Eq. (3.248) is a *forward looking* expression, in the sense that we look for the correct price to exchange with the underlying at time T , we discount it and we compare this with the current realization of the underlying $S(t)$. Eq. (3.249) is a *backward looking* expression, in the sense that the value of the forward contract is positive or negative according to the comparison between the current value of the underlying $S(t)$ and the capitalization of the initial value $S(0)$.

3.12.2 Futures contracts

Futures contracts are a particular type of forward contracts in which the counterparts “do not know/trust each other”. If the two counterparts agree to trade an asset in the future for a certain price, there are obvious risks. One of the investors may regret the deal and try to back out. Alternatively, the investor simply may not have the financial resources to honor the agreement. One of the key roles of the exchange is to organize trading so that contract defaults are avoided. This is where futures contracts and margin accounts come in.

Instead of signing the contract at time $t = 0$ and hope that at time T our counterpart will deliver the corresponding payoff (which can be positive or negative), we require it to deposit funds in a *margin account*. The amount that must be deposited at the time the contract is entered into is known as the *initial margin* and, at the end of each trading day, the margin account is adjusted to reflect the investor's gain or loss $S(t) \geq K$. This practice is referred to as *daily settlement* or *marking to market*.

The investor is entitled to withdraw any balance in the margin account in excess of the initial margin. To ensure that the balance in the margin account never becomes negative a *maintenance margin*, which is somewhat lower than the initial margin, is set. If the balance in the margin account falls below the maintenance margin, the investor receives a *margin call* and is expected to top up the margin account to the initial margin level by the end of the next day. The extra funds deposited are known as *variation margin*. If the investor does not provide the variation margin, the broker

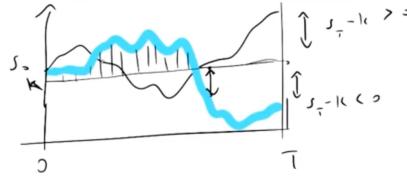


Figure 3.17: Futures contracts.

closes out the position⁶.

Here we will consider a simplified marking to market framework in which we make the following assumptions:

- the market to market is on the whole value of the contract, in such a way that if the payoff of the futures at time T is $S(T) - F(0, T)$, then:

$$price_t(\text{futures}) \equiv 0 \quad \forall t$$

- for all t there is a quoted strike price $F(t, T)$ called *strike price* which makes zero the value of the contract. This means that in the time interval $(s, t]$ the holder of a futures receives the dividend $F(t, T) - F(s, T)$.
For example, if there is a shock in the market so that $F(t, T) - F(s, T) < 0$, the long position has to pay this difference (margin call).

- The convergence condition

$$F(T, T) = S(T) \tag{3.250}$$

holds.

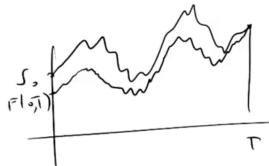


Figure 3.18: Convergence condition of the underlying

In summary:

- $price_t = 0 \quad \forall t \leq T$ (this plays the role of S_t);
- The dividend is equal to the futures price: $D(t) = F(t, T)$;
- $F(T, T) = S(T)$ (boundary condition).

From Section 3.9 we know that in presence of dividends the discounted asset is not a martingale, but the normalized gain process

$$S(t)e^{-\int_0^t r_s ds} + \int_0^t e^{-\int_0^s r_u du} dD \tag{3.251}$$

is a \mathbb{Q} -martingale. But in futures contracts

$$S(t) \leftrightarrow price_t \equiv 0, \quad dD \leftrightarrow dF$$

⁶ ^See [2], section 2.4.

so we are left with:

$$\int_0^t e^{-\int_0^s r_u \, du} dF_s \stackrel{(a)}{=} \int_0^t h_s \, dW_s^{\mathbb{Q}} \quad (3.252)$$

where in (a) we used the representation theorem for continuous (Brownian) martingales, which allows us to write the integral as a Brownian integral introducing an appropriate adapted process h_s . Writing (3.252) in its infinitesimal form we have:

$$e^{-\int_0^s r_u \, du} dF_t = h_s \, dW_s^{\mathbb{Q}} \quad \Rightarrow \quad dF(t, T) = \left(e^{\int_0^s r_u \, du} h_t \right) \, dW_s^{\mathbb{Q}}. \quad (3.253)$$

In other words, we have that $F(t, T)$ is a \mathbb{Q} -martingale. So, by martingality, the futures' price at time t is given by:

$$F(t, T) = \mathbb{E}_t^{\mathbb{Q}}[F(T, T)] = \mathbb{E}_t^{\mathbb{Q}}[S(T)] = \mathbb{E}^{\mathbb{Q}}[S(T)|\mathcal{F}_t] \quad (3.254)$$

Recall that for the forward contract we found that:

$$K(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[\frac{e^{-\int_t^T r_s \, ds} S(T)}{B(t, T)} \right] = \mathbb{E}_t^{\mathbb{Q}^T}[S(T)] \quad (3.255)$$

So forward price and futures price are both the expected value of the underlying at time T but they are computed under different probability measures (forward risk neutral measure and risk neutral measure).

Forward price and futures are the same if the two probability measures are the same. For example, this is true when the interest rate is deterministic.

3.12.3 Black formula

The Black formula gives the price of a call on a futures (on S). Here we assume that the underlying evolves according to the B&S dynamics with constant interest rate, $r_t \equiv r$. In this framework, if we denote with X the strike price of the call and we choose $T_1 > T$ as maturity of the call option, the payoff is

$$\text{payoff}_T = (F(T, T_1) - X)^+ + \underbrace{\text{payoff}_T(\text{long position on the futures})}_{=0} \quad (3.256)$$

In other words, we have a contract on a contract, which is called *compounded option*. Using eq. (3.254) we can write:

$$\begin{aligned} (F(T, T_1) - X)^+ &= \left(S(T) e^{r(T_1-T)} - X \right)^+ \\ &= e^{r(T_1-T)} \left(S(T) - X e^{-r(T_1-T)} \right)^+ \end{aligned} \quad (3.257)$$

Recall that according to the B&S formula the price of a call is given by:

$$\text{price}_t(\text{Call}) = S(t)\Phi(d_1) - K e^{-r(T-t)}\Phi(d_2)$$

In this case

$$K = X e^{-r(T_1-T)}, \quad S(t) = F(t, T_1) e^{-r(T_1-t)}$$

so we have:

$$\begin{aligned} \text{price}_t(\text{Call(futures)}) &= e^{r(T_1-T)} \left(S(t)\Phi(d_1) - e^{-r(T-t)} X e^{-r(T_1-T)} \Phi(d_2) \right) \\ &= e^{r(T_1-T)} S(t)\Phi(d_1) - e^{-r(T-t)} X e\Phi(d_2) \end{aligned}$$

$$= e^{-r(T-t)} \left(\underbrace{S(t)e^{-r(T_1-t)} \Phi(d_1) - X\Phi(d_2)}_{F(t,T_1)} \right)$$

We end up with the so called *Black formula*:

$$price_t(\text{Call(futures)}) = e^{-r(T-t)} (F(t, T_1)\Phi(d_1) - X\Phi(d_2)) \quad (3.258)$$

where in this case:

$$d_1 = \frac{\ln \frac{F(t, T_1)}{X} + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t}. \quad (3.259)$$

3.13 Quantity-adjusted options

Quantity-adjusted options, also called *quanto* options, are cash settled, cross currency derivatives where the underlying asset is denominated in a currency different from the one in which the option is settled. In other words, quanto options are derivatives where the payoff is defined by variables associated with one currency but is paid in another currency.

Lesson 19 of
22/04/2020

For example, the transfer between 1€ and 1\$ is given through an *exchange rate*

$$X(t)^{\text{EURUSD}} = \frac{1\text{€}}{1\$} \quad (3.260)$$

which is usually given using the *FORDOM* (foreign-domestic currency) notation. If we consider a third additional currency, for example the Japanese Yen ¥, in order to be consistent it must be that if X_1, X_2 are two exchange rates, then $X_3 = X_1 X_2$. So, the product of two exchange rates is an exchange rate.

In general, the interest rates in different countries are different, so – at least in principle – there can be a potential arbitrage opportunity. However, this is only an apparent problem since, as soon as we may want to exploit this arbitrage, we would have to consider the currency conversion, which works through a stochastic quantity, the interest rate.

As starting point, we assume that the evolution of the exchange rate between the two currencies is given by a B&S dynamics under the subjective probability measure \mathbb{P} :

$$\frac{dX(t)}{X(t)} = \mu_X dt + \sigma_X dW^{\mathbb{P}}(t) \quad (3.261)$$

This is a good candidate, in fact the products of two Brownian motions is itself a Brownian motion. The evolution of the corresponding riskless asset in the domestic currency is given by:

$$\frac{dB_d}{B_d} = r_d dt \quad (3.262)$$

Similarly, the riskless asset in the foreign currency evolves according to the equation:

$$\frac{dB_f}{B_f} = r_f dt \quad (3.263)$$

We already know the solutions of (3.262) and (3.263):

$$B_d(t) = e^{r_d t}, \quad B_f(t) = e^{r_f t} \quad (3.264)$$

Here we assume that r_d, r_f are constant and that if we buy foreign currency, this will evolve according r_f ⁷.

⁷ This is not always the case. For example, if we buy dollars in the euro-zone the evolution of this EUR-dollar is different from the one of USD-dollar. Anyway, this is a reasonable assumption.

We want to price options on the exchange rate at maturity X_T . Notice that these are not properly quanto options, because we need to introduce an underlying denominated in a different currency. For example, consider the call option that allows to buy 1\\$ at price $K\text{\euro}$, which has payoff $(X_T - K)^+$. According to the risk neutral methodology the price is given by:

$$\text{price}_t = e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}}[(X_T - K)^+]$$

But we are not able to define the risk neutral measure \mathbb{Q} if there is no underlying. In fact, in order to find \mathbb{Q}^d we have to impose that whatever discounted asset in our country is a martingale, i.e. it evolves with a drift r_d .

Which are the domestic assets? Let's consider a stylized economy in which we have the foreign and domestic saving accounts and we consider the transfers between the two. In this case there are two assets:

- B_d , in fact its drift is r_d ;
- $B_f X = \tilde{B}_f$, i.e. the foreign riskless asset converted into the domestic currency.

By Itô, we have that:

$$\begin{aligned} d\tilde{B}_f &= d(B_f X) = (dB_f)X + B_f dX \\ &= \tilde{B}_f r_f dt + \tilde{B}_f (\mu_X dt + \sigma_X dW^{\mathbb{P}}) \\ &= \tilde{B}_f (\mu_X + r_f) dt + \tilde{B}_f \sigma_X dW^{\mathbb{P}} \end{aligned} \quad (3.265)$$

This is the general dynamics of a foreign riskless asset converted in the domestic currency. Now we want that

$$\mu_X + r_f = r_d$$

so we perform the same trick we used when we applied the martingale approach to the B&S model, that is we absorb the excessive drift into the Brownian motion:

$$(3.265) = \tilde{B}_f \left(r_d dt + \sigma_X \left(dW^{\mathbb{P}} + \frac{\mu_X + r_f - r_d}{\sigma_X} dt \right) \right) \quad (3.266)$$

If we define the change of probability measure given by the Radon-Nikodym derivative⁸:

$$\begin{aligned} \frac{d\mathbb{Q}^d}{d\mathbb{P}} \Big|_t &= \mathcal{E} \left(\frac{\mu_X + r_f - r_d}{\sigma_X} \right)_t \\ &= \exp \left\{ \int_0^t \left(\frac{\mu_X + r_f - r_d}{\sigma_X} \right) dW^{\mathbb{P}} - \int_0^t \left(\frac{\mu_X + r_f - r_d}{\sigma_X} \right)^2 ds \right\} \\ &\sim \exp \left\{ \alpha W(t) - \frac{1}{2} \sigma^2 t \right\} = \text{true martingale} \end{aligned} \quad (3.267)$$

we can apply the Girsanov theorem:

$$dW^{\mathbb{P}} + \frac{\mu_X + r_f - r_d}{\sigma_X} dt = dW^{\mathbb{Q}^d}$$

Then:

$$d\tilde{B}_f = \tilde{B}_f (r_d dt + \sigma_X dW^{\mathbb{Q}^d}) \quad (3.268)$$

⁸ $\wedge \mathcal{E}$ is the so called Doléans-Dade exponential.

So, the dynamics of X_t under the domestic risk neutral pricing probability measure \mathbb{Q}^d is given by:

$$\frac{dX_t}{X} = (r_d - r_f)dt + \sigma_X dW^{\mathbb{Q}^d} \quad (3.269)$$

In a sense, the short foreign interest rate r_f plays the role of a dividend.

Remark 3.13.1. Notice that the exchange rate X/B^d , i.e. the discounted X , is not a \mathbb{Q}^d -martingale. This is confirmed by the fact that there is an additional drift r_f .

Now we are ready to price an exchange option.

Proposition 3.13.1. *The price at time t of whatever payoff written on the exchange rate at maturity T is given by:*

- the discounted value under the domestic interest rate of the conditional risk neutral expectation of the payoff:

$$price_t = e^{-r_d(T-t)} \mathbb{E}_{t,X}^{\mathbb{Q}^d} [\text{payoff}(X_T)] \quad (3.270)$$

where the dynamics of X_t is given by (3.269);

- the solution:

$$price_t = F(t, X_t) \quad (3.271)$$

of the PDE:

$$\begin{cases} \frac{\partial F}{\partial t} + (r_d - r_f)X \frac{\partial F}{\partial X} + \frac{1}{2}X^2 \sigma_X^2 \frac{\partial^2 F}{\partial X^2} - r_d F = 0 \\ F(X, T) = \text{payoff}(X) \end{cases} \quad (3.272)$$

- the solution

$$F(t, X) = F_0(t, X e^{-r_f(T-t)}) \quad (3.273)$$

where F_0 is the solution corresponding to no dividends, which obeys to the equation:

$$\frac{dX}{X} = r_d dt + \sigma_X dW^{\mathbb{Q}} \quad (3.274)$$

Example 3.13.1: Call option on the exchange rate

The price of a call option on X is given by:

$$price_t(\text{Call on } X) = X_t e^{-r_f(T-t)} \Phi(d_1) - e^{-r_d(T-t)} K \Phi(d_2) \quad (3.275)$$

where

$$d_1 = \frac{1}{\sigma_X \sqrt{T-t}} \left(\ln \frac{X_t e^{-r_f(T-t)}}{K} + \left(r_d + \frac{1}{2} \sigma_X^2 \right) (T-t) \right) \quad (3.276)$$

$$d_2 = \frac{1}{\sigma_X \sqrt{T-t}} \left(\ln \frac{X_t e^{-r_f(T-t)}}{K} + \left(r_d - \frac{1}{2} \sigma_X^2 \right) (T-t) \right) \quad (3.277)$$

Yet, we are not ready to price quanto options. We need to do an upgrade, introducing in the market some risky assets. The starting point is that the dynamics under \mathbb{P} is given by

$$\frac{dX}{X} = \mu_X dt + \sigma_X \cdot dW^{\mathbb{P}} \quad (3.278)$$

for the exchange rate, by

$$\frac{dS_d}{S_d} = \mu_d dt + \sigma_d \cdot dW^{\mathbb{P}} \quad (3.279)$$

for the domestic asset and by

$$\frac{dS_f}{S_f} = \mu_f dt + \sigma_f \cdot dW^{\mathbb{P}} \quad (3.280)$$

for the foreign asset. We used the scalar product (\cdot) because σ is a row vector and dW is a column vector.

The volatility matrix of our market is given by

$$\sigma = \begin{pmatrix} \sigma_X \\ \sigma_d \\ \sigma_f \end{pmatrix} = \begin{pmatrix} \sigma_{X_1} & \sigma_{X_2} & \sigma_{X_3} \\ \sigma_{d_1} & \sigma_{d_2} & \sigma_{d_3} \\ \sigma_{f_1} & \sigma_{f_2} & \sigma_{f_3} \end{pmatrix} \quad (3.281)$$

We also add a domestic and a foreign riskless assets which evolve according to the equations:

$$\frac{dB_d}{B_d} = r_d dt \quad (3.282)$$

$$\frac{dB_f}{B_f} = r_f dt \quad (3.283)$$

Now we can consider different problems involving different contract. For example:

- pricing in euros a call on S_f in dollars with K in dollars, which has payoff $(S_f(T) - K)^+ X_T^{\text{USDEUR}}$ (not a real quanto option);
- pricing in euros a call on S_f in dollars with K in euros, which has payoff $(S_f X_T - K)^+$ (real quanto option).

Now we have to characterize the probability measure \mathbb{Q}^d under which we price this payoffs. Again, we impose that under \mathbb{Q}^d all the domestic assets discounted by the domestic interest rate are \mathbb{Q}^d -martingales, i.e. the domestic assets evolve with drift r_d . Our domestic assets are $B_d, S_d, \tilde{S}_f = S_f X_t$ and $\tilde{B}_f = B_f X_t$, so this quantities have – once discounted by B_d – must be \mathbb{Q}^d -martingales.

Let's start from S_d and rewrite (3.279) in such a way that the drift is r_d :

$$\frac{dS_d}{S_d} = r_d dt + \sigma_d \left(dW^{\mathbb{P}} + \frac{\mu_d - r_d}{\sigma_d} dt \right) \quad (3.284)$$

This is only a formal expression because, since σ_d is a row vector, $1/\sigma_d$ is not well defined.

Then we have:

$$\begin{aligned} \frac{d\tilde{S}_f}{\tilde{S}_f} &= (\mu_f + \mu_X + \sigma_f \sigma_X^T) dt + (\sigma_f + \sigma_d) dW^{\mathbb{P}} \\ &= r_d dt + (\sigma_f + \sigma_X) \left(dW^{\mathbb{P}} + \frac{\mu_f + \mu_X + \sigma_f \sigma_X^T - r_d}{\sigma_f \sigma_X} dt \right) \end{aligned} \quad (3.285)$$

which again is just a formal expression. Finally, we can do the same thing for the foreign riskless asset denominated in the domestic currency:

$$\frac{d\tilde{B}_f}{\tilde{B}_f} = (\mu_f + \mu_X) dt + \sigma_X dW^{\mathbb{P}}$$

$$= r_d dt + \sigma_X \left(dW^{\mathbb{P}} + \frac{\mu_X + r_f - r_d}{\sigma_X} dt \right) \quad (3.286)$$

Provided that the Girsanov Theorem can be applied to a 3-dimensional geometric Brownian motion $W^{\mathbb{P}}$ we have to find the corresponding Radon-Nikodym derivative

$$\frac{dQ^d}{d\mathbb{P}} \Big|_t = \text{homework} \quad (3.287)$$

Now, for example, consider the problem of pricing in euros a call on S_f in dollars with K in dollars, which has payoff $(S_f(T) - K)^+ X_T^{\text{USDEUR}}$. We can solve this problem by pricing directly in the foreign country, that is under \mathbb{Q}^f , using the B&S model and then convert the result using the exchange rate:

$$\text{price} = X_t(S_f(t)\Phi(d_1) - e^{-r_f(T-t)}K\Phi(d_2)) \quad (3.288)$$

where

$$d_1 = \frac{1}{\|\sigma_f\|\sqrt{T-t}} \left(\ln \frac{S_f(t)}{K} + \left(r_f + \frac{1}{2}\|\sigma_f\|^2 \right) (T-t) \right) \quad (3.289)$$

Again, we see that this is not a real quanto problem but just a matter of conversion. Let's consider the real quanto option problem, i.e. pricing in euros a call on S_f in dollars with K in euros, which has payoff $(S_f X_T - K)^+$. This is a call written on \tilde{S}_f so we can apply the \mathbb{Q}^d -pricing methodology on it. So, the price of this quanto option is:

$$\text{price}_t((S_f X_T - K)^+) = \tilde{S}_f(t)\Phi(d_1) - e^{-r_d(T-t)}K\Phi(d_2) \quad (3.290)$$

with

$$d_1 = \frac{1}{\|\sigma_X + \sigma_f\|\sqrt{T-t}} \left(\ln \frac{\tilde{S}_f(t)}{K} + \left(r_d + \frac{1}{2}\|\sigma_X + \sigma_f\|^2 \right) (T-t) \right) \quad (3.291)$$

Remark 3.13.2. $\mathbb{Q}^d \neq \mathbb{Q}^f$ (Siegel paradox). This happens because under \mathbb{Q}^d we impose that everything in the domestic market – after being discounted by the domestic interest rate – is a martingale while under \mathbb{Q}^f we impose that all the foreign assets, including the domestic ones converted in the foreign currency, will be martingales – once discounted by the foreign riskless asset. This leads to different dynamics. In particular, the exchange rate under \mathbb{Q}^d evolves according to the equation:

$$\frac{dX}{X} = (r_d - r_f) dt + \sigma_X dW^{\mathbb{Q}^d} \quad (3.292)$$

Consider the euro-zone and the dollar-zone. If the exchange rate between euros and dollars is X then the one from dollars to euros must be $1/X$. But the dynamics under $1/X$ is given by

$$\frac{d\frac{1}{X}}{\frac{1}{X}} = (r_d - r_f \|\sigma_X\|^2) dt + \sigma_X dW^{\mathbb{Q}^d} \quad (3.293)$$

so there is an additional term, essentially the Itô term, which is responsible for the fact that the two measures are not the same. This is due to the randomness of the exchange rate.

Example 3.13.2: Quanto exchange call

Price the exchange quanto option, which has payoff:

$$(X_T S_f(T) - S_d(T))^+ = (\tilde{S}_f(T) - S_d(T))^+$$

Homework

3.14 Barrier options

Lesson 20 of
23/04/2020

A *barrier option* is a type of derivative where the payoff depends on whether or not the underlying asset has reached or exceeded a predetermined price. In order to deal with the past history of the underlying we need some stochastic analysis tools.

3.14.1 Mathematical background

Definition 3.14.1 (Mirror reflection). *The mirror reflection of a Brownian motion $B(t)$ at the level m and in the interval $[\tau, T]$ is defined as:*

$$\tilde{B}(t) = \begin{cases} B(t) & t < \tau \\ 2m - B(t) & \tau \leq t \leq T \end{cases} \quad (3.294)$$

The idea is that if we have an object in a point b and if in a point $m < b$ we put a mirror, then the image of b with respect to m will be at $2m - b$.

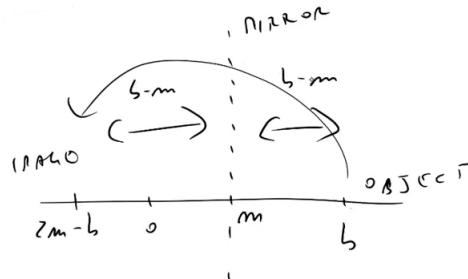


Figure 3.19: Mirror reflection.

One fundamental property is that, if the Brownian motion crosses the barrier at m , then

$$\mathbb{P}(B(T) > x) = \mathbb{P}(\tilde{B}(T) < 2m - x) \quad (3.295)$$

In order to have an intuition about that, look at Figure 3.20. For $t > \tau$ we can consider the original Brownian motion or its mirror reflection, so at time T the probability that $B(T) > x$ is equal to the probability that $\tilde{B}(T) < 2m - x$.

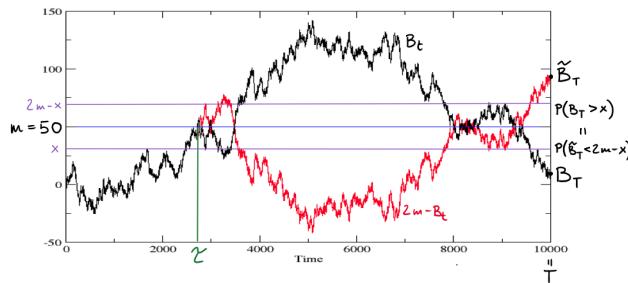


Figure 3.20: For $t > \tau$ we can consider the original Brownian motion or its mirror reflection.

Definition 3.14.2. *The running maximum of the Brownian motion $B(t)$ is defined as*

$$M(t) = \sup_{0 \leq s \leq t} B(s) \quad (3.296)$$

Definition 3.14.3. *The first hitting time is defined as*

$$T_a = \inf\{t : B(t) = a\}, \quad a > 0 \quad (3.297)$$

Since

$$\limsup_t B(t) = +\infty,$$

we have that

$$\mathbb{P}(T_a < \infty) = 1 \quad (3.298)$$

so the hitting time is finite but not bounded.

Another important property that comes from the definition of BM is that the increment $B(t+T) - B(T)$ is still a BM and it is independent of the filtration \mathcal{F}_T . This property holds also for stochastic times, for example the first hitting time:

$$B(t + T_a) - B(T_a) = \text{Brownian motion } \perp\!\!\!\perp \mathcal{F}_s \quad \forall s \leq T_a \quad (3.299)$$

Provided that the Brownian motion will reach a at some time, we have that

$$B(T_a) = a, \quad (3.300)$$

so the Brownian motion at the random time T_a becomes a (deterministic) constant.

Theorem 3.14.1 (Reflection principle). *Let $B(t)$ be a standard Brownian motion and let $a > 0$. Then*

$$\mathbb{P}(M(t) \geq a) = 2\mathbb{P}(B(t) \geq a) = \frac{2}{\sqrt{2\pi t}} \int_a^{+\infty} e^{-\frac{x^2}{2t}} dx. \quad (3.301)$$

Proof.

$$\begin{aligned} \mathbb{P}(B(t) \geq a) &= \mathbb{P}(B(t) \geq a, M(t) \geq a) + \underbrace{\mathbb{P}(B(t) \geq a, M(t) < a)}_{=0} \\ &= \mathbb{P}(B(t) \geq a | M(t) \geq a) \mathbb{P}(M(t) \geq a) \\ &\stackrel{(3.300)}{=} \mathbb{P}(B(T_a + (t - T_a)) - B(T_a) \geq 0 | M(t) \geq a) \mathbb{P}(M(t) \geq a) \\ &= \frac{1}{2} \mathbb{P}(M(t) \geq a) \end{aligned}$$

where we used the fact that the highlighted term is the probability of a Brownian motion, i.e. a Gaussian variable, to be greater than zero, i.e. $1/2$. \square

The importance of the reflection principle comes from the fact that it translates the problem of computing a $\mathbb{P}(M_t \geq a)$, which involves the whole past of the BM, in terms of the probability of a Brownian random variable, which is very simple, being Gaussian.

Now we need the probability of the joint distribution of the maximum of the BM and the BM itself.

Proposition 3.14.1. *For $a > 0$ and $y \geq 0$:*

$$\mathbb{P}(M(t) \geq a, B(t) \leq a - y) = \mathbb{P}(B(t) > a + y) \quad (3.302)$$

Alternatively, $\forall m > 0$ such that $m \geq b$:

$$\mathbb{P}(M(t) > m, B(t) < b) = \mathbb{P}(B(t) > 2m - b). \quad (3.303)$$

Proof.

$$\begin{aligned}\mathbb{P}(B(t) > a + y) &= \mathbb{P}(B(t) > a + y, M(t) \geq a) + \underbrace{\mathbb{P}(B(t) > a + y, M(t) < a)}_{=0} \\ &\stackrel{(a)}{=} \mathbb{P}(B(T_a + (t - T_a)) - a > -y | M(t) \geq a) \mathbb{P}(M(t) \geq a) \\ &= \mathbb{P}(B(T_a + (t - T_a)) < a - y, M(t) \geq a)\end{aligned}$$

where in (a) we used the symmetry of the BM. \square

Corollary 3.14.1.1. *For any $a > 0$ and $y \geq 0$*

$$\mathbb{P}(M(t) \leq a, B(t) \leq a - y) = \Phi\left(\frac{a - y}{\sqrt{t}}\right) - \Phi\left(\frac{-a - y}{\sqrt{t}}\right). \quad (3.304)$$

Proof. From the fact that

$$\mathbb{P}(B(t) < x) = \mathbb{P}(B(t) < x, M(t) < y) + \mathbb{P}(B(t) < x, M(t) \geq y)$$

we have that

$$\mathbb{P}(M(t) \leq a, B(t) \leq a - y) = \mathbb{P}(B(t) \leq a - y) - \mathbb{P}(B(t) \leq a - y, M(t) > a)$$

\square

By construction, the *joint probability density* is given by

$$\begin{aligned}\mathbb{P}(M(t) \in dm, B(t) \in db) &= -\frac{\partial^2}{\partial m \partial b} \mathbb{P}(B(t) > 2m - b) \\ &= \frac{2(2m - b)}{\sqrt{2\pi t^3}} e^{-\frac{(2m-b)^2}{2t}} \mathbf{1}_{m \geq 0} \mathbf{1}_{m \geq b},\end{aligned} \quad (3.305)$$

as we can see by computing

$$\begin{aligned}-\frac{\partial^2}{\partial m \partial b} \frac{1}{\sqrt{2\pi t}} \int_{2m-b}^{+\infty} e^{-\frac{x^2}{2t}} dx &= -\frac{\partial}{\partial m} \left(\frac{1}{\sqrt{2\pi t}} e^{-\frac{(2m-b)^2}{2t}} \right) \\ &= \frac{2(2m-b)}{\sqrt{2\pi t^3}} e^{-\frac{(2m-b)^2}{2t}}.\end{aligned}$$

From the reflection principle we have that

$$\begin{aligned}\mathbb{P}(T_a \leq t) &= \mathbb{P}(M(t) \geq a) = 2\mathbb{P}(B(t) \geq a) \\ &= \frac{2}{\sqrt{2\pi t}} \int_a^{+\infty} e^{-\frac{x^2}{2t}} dx \\ &= 2 \left(1 - \Phi\left(\frac{a}{\sqrt{t}}\right) \right).\end{aligned} \quad (3.306)$$

So, the density of the hitting time is given by

$$\begin{aligned}f_{T_a}(t) &= \frac{\partial \mathbb{P}(T_a \leq t)}{\partial t} \\ &= -\frac{t^{-3/2}}{2} \frac{2}{\sqrt{2\pi}} \int_a^{+\infty} e^{-\frac{x^2}{2t}} dx + \frac{2}{\sqrt{2\pi t}} \int_a^{+\infty} \frac{x^2}{2t^2} e^{-\frac{x^2}{2t}} dx \\ &\stackrel{(a)}{=} -\frac{t^{-3/2}}{2} \frac{2}{\sqrt{2\pi}} \int_a^{+\infty} e^{-\frac{x^2}{2t}} dx + \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}} + \frac{1}{\sqrt{2\pi t^3}} \int_a^{+\infty} e^{-\frac{x^2}{2t}} dx\end{aligned}$$

$$= \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}} \mathbf{1}_{t>0} \quad (3.307)$$

where in (a) we computed the last integral:

$$\begin{aligned} \frac{2}{\sqrt{2\pi t}} \int_a^{+\infty} \frac{x^2}{2t^2} e^{-\frac{x^2}{2t}} dx &= \frac{2}{\sqrt{2\pi t}} \int_a^{+\infty} \left(-\frac{x}{t} \right) e^{-\frac{x^2}{2t}} \left(-\frac{x}{2t} \right) dx \\ &= \frac{2}{\sqrt{2\pi t}} \left(e^{-\frac{x^2}{2t}} \left(-\frac{x}{2t} \right) \Big|_a^{+\infty} - \int_a^{+\infty} e^{-\frac{x^2}{2t}} \left(-\frac{1}{2t} \right) dx \right) \\ &= \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}} + \frac{1}{\sqrt{2\pi t^3}} \int_a^{+\infty} e^{-\frac{x^2}{2t}} dx. \end{aligned}$$

Now we can compute the moment generating function of T_a :

$$\mathbb{E}[e^{\lambda T_a}] = \int_0^\infty e^{-\lambda t} \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}} dt = \text{homework} = e^{-\sqrt{2\lambda}a} \quad (3.308)$$

where $\lambda > 0$. An alternative way of computing it comes from the fact that $e^{\alpha B(t) - \frac{1}{2}\alpha^2 B(t)}$ is an exponential martingale, so

$$\mathbb{E} \left[e^{\alpha B(t) - \frac{1}{2}\alpha^2 t} \right] = 1$$

For the case in which $t = T_a$ we get:

$$\begin{aligned} 1 &= \mathbb{E} \left[e^{\alpha B(T_a) - \frac{1}{2}\alpha^2 T_a} \right] \stackrel{(a)}{=} \overbrace{\mathbb{E} \left[e^{\alpha B(T_a) - \frac{1}{2}\alpha^2 T_a} \right]}^{\cancel{\text{---}}} \\ &= \overbrace{\mathbb{E} \left[e^{\alpha a - \frac{1}{2}\alpha^2 T_a} \right]}^{\cancel{\text{---}}} - e^{\alpha a} \overbrace{\mathbb{E} \left[e^{-\frac{1}{2}\alpha^2 T_a} \right]}^{\cancel{\text{---}}} \end{aligned}$$

However, in principle, the equality (a) is not true. In fact, in general, if $X(t)$ is a martingale then $X(\tau)$ is a martingale only if τ is bounded. But T_a is finite but not bounded! So, we cannot use the naive version of the martingale property. The intuition is still good, but we need to find a proper way to do the calculation.

Let's start by applying the martingale property to a bounded stopping time $t \wedge T_a = \min\{t, T_a\}$. If

$$e^{\alpha B(t) - \frac{1}{2}\alpha^2 t} = \text{martingale} \quad (3.309)$$

then, by the optional sampling theorem:

$$e^{\alpha B(t \wedge T_a) - \frac{1}{2}\alpha^2(t \wedge T_a)} = \text{martingale} \quad (3.310)$$

But now we cannot substitute $B(t \wedge T_a)$ with a , in fact

$$B(t \wedge T_a) \leq a \quad (3.311)$$

By martingality, we have that:

$$\mathbb{E} \left[e^{\alpha B(t \wedge T_a) - \frac{1}{2}\alpha^2(t \wedge T_a)} \right] = 1 \quad (3.312)$$

From (3.311) and the fact that $\frac{1}{2}\alpha^2(t \wedge T_a) > 0$, we get:

$$e^{\alpha B(t \wedge T_a) - \frac{1}{2}\alpha^2(t \wedge T_a)} \leq e^{\alpha a} < \infty \quad (3.313)$$

Since $e^{\alpha a}$ is uniformly bounded in t , we can apply the dominated convergence theorem taking the limit of both sides and interchange it with the expected value:

$$1 = \mathbb{E} \left[\lim_{t \rightarrow \infty} e^{\alpha B(t \wedge T_a) - \frac{1}{2}\alpha^2(t \wedge T_a)} \right] = \begin{cases} e^{\alpha B(T_a) - \frac{1}{2}\alpha^2 T_a} & T_a < \infty \\ 0 & T_a = \infty \end{cases}$$

$$= \mathbb{E} \left[\lim_{t \rightarrow \infty} e^{\alpha B(T_a) - \frac{1}{2} \alpha^2 T_a} \mathbf{1}_{T_a < \infty} \right] \quad (3.314)$$

This is true for all α . Then, if we take $\alpha = 0$ we get:

$$1 = \mathbb{E} [\mathbf{1}_{T_a < \infty}] = \mathbb{P}(T_a < \infty) = 1 \quad (3.315)$$

which again states that the stopping time is finite. So, we can write:

$$1 = e^{\alpha a} \mathbb{E} \left[e^{-\frac{1}{2} \alpha^2 T_a} \right] \quad (3.316)$$

which is the same expression that we found naively at the beginning. Rearranging eq. (3.316) we get:

$$\mathbb{E} \left[e^{-\frac{1}{2} \alpha^2 T_a} \right] = e^{-\alpha a} \quad (3.317)$$

Then, if we define $\frac{1}{2} \alpha^2 \equiv \lambda$, we obtain:

$$\mathbb{E} \left[e^{-\lambda T_a} \right] = e^{-\sqrt{2\lambda} a} \quad (3.318)$$

which is the same expression of (3.308).

What happens if $a < 0$? The idea is that the stopping time can be re-defined as

$$T_a = \inf \{t \geq 0 : \tilde{B}(t) \equiv -B(t) = -a \geq 0\} \quad (3.319)$$

which leads to

$$\mathbb{E} \left[e^{-\lambda T_a} \right] = e^{\sqrt{2\lambda} a} \quad (3.320)$$

In summary, for $\lambda > 0$ and $a \in \mathbb{R}$, we have:

$$\mathbb{E} \left[e^{-\lambda T_a} \right] = e^{-\sqrt{2\lambda} |a|} \quad (3.321)$$

Furthermore, it turns out that $\mathbb{E}[T_a] = +\infty$. In fact:

$$\lim_{\lambda \rightarrow 0} \frac{\partial}{\partial \lambda} \mathbb{E} \left[e^{-\lambda T_a} \right] = \lim_{\lambda \rightarrow 0} -\frac{|a|}{\sqrt{\lambda}} e^{-|a|\sqrt{2\lambda}} = +\infty \quad (3.322)$$

So far we considered stopping times and hitting times for the BM. This is not enough, because if we want to consider contracts written on the underlying we need to consider the geometric BM.

Now let's consider a (arithmetic, not geometric) Brownian motion with a linear drift μ :

$$\begin{cases} dX(t) = \mu dt + \sigma dW(t) \\ X(0) = \alpha \end{cases} \quad (3.323)$$

with μ, σ constants. We want to compute the hitting time: the idea is to change the probability measure in such a way that the drift is absorbed in a shift of the BM. Solving (3.323) we get:

$$X(t) = \alpha + \mu t + \sigma W(t) = \alpha + \sigma (\theta t + W(t)) \quad (3.324)$$

where $\theta = \mu/\sigma$. Now we change the BM

$$W(t) + \theta t \equiv \tilde{W}(t) \quad (3.325)$$

in such a way that under the new probability measure $\tilde{\mathbb{P}}$ associated to \tilde{W} there is no drift:

$$X(t) = \alpha + \sigma \tilde{W}(t) \quad (3.326)$$

The corresponding Radon-Nikodym derivative is:

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \Big|_t = \exp \left\{ -\theta W(t) - \frac{1}{2} \theta^2 t \right\} \quad (3.327)$$

Let's define the hitting time for the BM with drift as:

$$T_\beta = \inf \{t : X(t) = \beta\} \stackrel{(3.326)}{=} \inf \left\{ t : \tilde{W}(t) = \frac{\beta - \alpha}{\sigma} \right\} \quad (3.328)$$

Now we can compute the probability:

$$\mathbb{P} \left(\max_{s \leq t} X_s \geq \beta \right) = \mathbb{E}^\mathbb{P} [\mathbf{1}_{T_\beta \leq t}] \quad (3.329)$$

Recall that the expected value of a given process Y is given by:

$$\mathbb{E}^\mathbb{P}[Y] = \int Y \frac{d\mathbb{P}}{d\tilde{\mathbb{P}}} d\tilde{\mathbb{P}} = \mathbb{E}^{\tilde{\mathbb{P}}} \left[\frac{\mathbb{P}}{\tilde{\mathbb{P}}} \right]$$

so we have to consider the converse of the Radon-Nikodym derivative:

$$\begin{aligned} (3.329) &= \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{\theta W(t) + \frac{1}{2} \theta^2 t} \mathbf{1}_{T_\beta \leq t} \right] \\ &\stackrel{(3.325)}{=} \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{\theta \tilde{W}(t) - \frac{1}{2} \theta^2 t} \mathbf{1}_{T_\beta \leq t} \right] \\ &= \mathbb{E}^{\tilde{\mathbb{P}}} \left[\mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{\theta \tilde{W}(t) - \frac{1}{2} \theta^2 t} \mathbf{1}_{T_\beta \leq t} \right] \middle| \mathcal{F}_{t \wedge T_\beta} \right] \\ &= \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{\theta \frac{\beta - \alpha}{\sigma} - \frac{1}{2} \theta^2 t} \mathbf{1}_{T_\beta \leq t} \right] \\ &= \int_0^t e^{\theta \frac{\beta - \alpha}{\sigma} - \frac{1}{2} \theta^2 y} \tilde{\mathbb{P}}(T_\beta \in dy) \\ &\stackrel{(3.307)}{=} \int_0^t e^{\theta \frac{\beta - \alpha}{\sigma} - \frac{1}{2} \theta^2 y} \frac{\frac{\beta - \alpha}{\sigma}}{\sqrt{2\pi y^3}} e^{-\frac{1}{2y}(\frac{\beta - \alpha}{\sigma})} dy \\ &= \int_0^t \frac{\beta - \alpha}{\sigma} \frac{1}{\sqrt{2\pi y^3}} \exp \left\{ -\frac{\left(\frac{\beta - \alpha}{\sigma} - \theta y\right)^2}{2y} \right\} dy \end{aligned} \quad (3.330)$$

Summarizing, we found that for a BM $X(t) = \alpha + \mu t + \sigma W(t)$ with drift the probability of its maximum to be greater or equal to β , such that $T_\beta = \inf \{t : X(t) = \beta\}$, is given by:

$$\mathbb{P} \left(\max_{s \leq t} X_s \geq \beta \right) = \int_0^t \frac{\beta - \alpha}{\sigma} \frac{1}{\sqrt{2\pi y^3}} \exp \left\{ -\frac{\left(\frac{\beta - \alpha}{\sigma} - \theta y\right)^2}{2y} \right\} dy \quad (3.331)$$

where $\theta = \mu/\sigma$.

Example 3.14.1

Consider the exponential martingale

$$Y(t) = e^{W(t) - \frac{1}{2}t} \quad (3.332)$$

and the stopping time

$$\tau_n = \inf\{t : Y(t) = n\}, \quad n \in \mathbb{N} \quad (3.333)$$

We would like to understand which is the probability

$$\mathbb{P}(Y(t) = n \text{ for some } t) = ? \quad (3.334)$$

The problem is that $Y(t)$ is not uniformly integrable (UI)^a. In fact,

$$\lim_{t \rightarrow \infty} Y(t) = \lim_{t \rightarrow \infty} e^{t\left(\frac{W(t)}{t} - \frac{1}{2}\right)} = 0$$

and so

$$1 = \mathbb{E}[Y(t)] \neq \mathbb{E}[Y(\infty)] = 0$$

So, since there is no convergence, the martingale is not uniform integrable. The idea to solve this problem is to exploit the UI of a particular version of this exponential martingale. Let's consider the exponential martingale computed at time $t \wedge \tau_n$: $Y(t \wedge \tau_n)$. Being a martingale stopped at a bounded time, this is still a martingale. If we consider the limit, we get:

$$\lim_{t \rightarrow \infty} Y(t \wedge \tau_n) = \begin{cases} 0 & \text{if } \nexists t : Y(t) = n \\ n & \text{if } \exists t : Y(t) = n \end{cases} \quad (3.335)$$

We found that the martingale $Y(t \wedge \tau_n)$ is bounded both from above (by n) and below (by construction) and so it is UI and it converges:

$$1 = \mathbb{E}[Y(t)] = \mathbb{E}[Y(\infty \wedge \tau_n)] \stackrel{(3.335)}{=} n\mathbb{P}(Y(t) = n \text{ for some } t) + 0$$

Solving this equation we get:

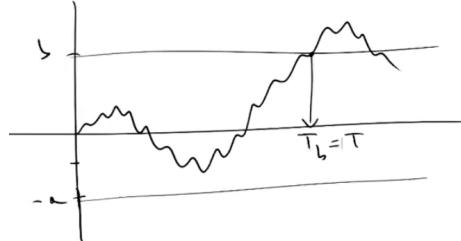
$$\mathbb{P}(Y(t) = n \text{ for some } t) = \frac{1}{n}. \quad (3.336)$$

^a The criterion for the uniform integrability is that $\lim_{M \rightarrow \infty} \sup_t \mathbb{E}[X(t) \mathbf{1}_{|X(t)| > M}] = 0$. This condition is quite difficult to verify. The usual procedure consists in bounding the martingale. If the martingale is bounded, then it is UI.

Example 3.14.2

Consider two real numbers $a, b > 0$, a BM $B(t)$ and the first hitting time

$$T = \inf\{t : B(t) = -a \text{ or } B(t) = b\} = T_{-a} \wedge T_b \quad (3.337)$$



We want to compute $\mathbb{E}[T]$. Since T is bounded, we have to construct something which is bounded. We know that, by construction, $|B(t)| \leq \max\{a, b\}$. Then,

from the optional stopping theorem, $B_{t \wedge T}$ is a martingale, so it is UI. By martingality we have that the expected value of $B_{t \wedge T}$ is equal to the initial value, which is zero:

$$0 = \mathbb{E}[B_{t \wedge T}]$$

So, if we consider the limit, we have:

$$\begin{aligned} 0 &= \lim_{t \rightarrow \infty} \mathbb{E}[B_{t \wedge T}] = \mathbb{E}\left[\lim_{t \rightarrow \infty} B_{t \wedge T}\right] = \mathbb{E}[B(T)] \\ &= -a\mathbb{P}(B(T) = -a) + b\mathbb{P}(B(T) = b) + (?)\mathbb{P}(B(\infty) \in (a, b)) \\ &= -a\mathbb{P}(B(T) = -a) + b\mathbb{P}(B(T) = b) \end{aligned} \quad (3.338)$$

where the last term stands for the case in which the BM will always remain inside the interval (a, b) and it is zero because the BM has infinite upper and lower limit. From (3.338) we conclude that the two events are exclusive:

$$\mathbb{P}(B(T) = b) = 1 - \mathbb{P}(B(T) = -a) \quad (3.339)$$

Moreover, we can rewrite (3.338) as:

$$0 = -a(1 - \mathbb{P}(B(T) = b)) + b\mathbb{P}(B(T) = b) \quad (3.340)$$

so that

$$\mathbb{P}(B(T) = b) = \frac{a}{b} \frac{1}{1 - \frac{a}{b}} = \frac{a}{a+b} \Rightarrow \mathbb{P}(B(T) = -a) = \frac{b}{a+b} \quad (3.341)$$

$\mathbb{P}(B(T) = b)$ is called *gain probability* and $\mathbb{P}(B(T) = -a)$ *ruin probability*.

Now we use another martingale, which is $B(t)^2 - t$. By the optional stopping theorem we have that

$$B(t \wedge T)^2 - (t \wedge T) = \text{martingale} \quad (3.342)$$

and by martingality we get:

$$\mathbb{E}[B(t \wedge T)^2] = \mathbb{E}[t \wedge T] \quad (3.343)$$

Now, recalling that $B(t \wedge T)^2 \leq \max\{a^2, b^2\}$, if we take the limit for t going to infinite we get:

$$\mathbb{E}[B(T)^2] = \mathbb{E}[T] \quad (3.344)$$

So we are able to express the expected value we want to compute in terms of a discrete random variable which takes values a^2, b^2 . Using eq. (3.341) we can write eq. (3.344) as:

$$a^2 \frac{b}{a+b} + b^2 \frac{a}{a+b} = \mathbb{E}[T] \quad (3.345)$$

Thus, we get:

$$\mathbb{E}[T] = ab. \quad (3.346)$$

3.14.2 Barrier options: a classic approach

Consider a B&S model in which the underlying under the risk neutral probability measure is given by

$$\begin{aligned} S(t) &= S(0) \exp \left\{ \left(r - \frac{1}{2}\sigma^2 \right) t + \sigma B^{\mathbb{Q}}(t) \right\} \\ &= S(0) \exp \left\{ \sigma \left(\left(\frac{r}{\sigma} - \frac{\sigma}{2} \right) t + B^{\mathbb{Q}}(t) \right) \right\} \\ &\stackrel{(a)}{=} S(0) \exp \left\{ \sigma \tilde{B}^{\mathbb{Q}}(t) \right\} \end{aligned} \quad (3.347)$$

where in (a) we defined

$$\theta \equiv \frac{r}{\sigma} - \frac{\sigma}{2} \quad \text{and} \quad \tilde{B}^{\mathbb{Q}}(t) \equiv B^{\mathbb{Q}}(t) + \theta t \quad (3.348)$$

Let's define the running maximum as

$$M^S(t) = \max_{s \leq t} S_s = S(0) \exp \left\{ \sigma M^{\tilde{B}}(t) \right\} \quad (3.349)$$

In order to price something which involves the past trajectory of the underlying, we need to consider the joint probability of the running maximum and the current value of the underlying:

$$\begin{aligned} \mathbb{Q}(M^S(t) \leq L, S(t) \geq K) &= \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{M^S(t) \leq L, S(t) \geq K}] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{M^{\tilde{B}}(t) \leq \frac{1}{\sigma} \ln \frac{L}{S(0)}, \tilde{B}(t) \geq \frac{1}{\sigma} \ln \frac{K}{S(0)}} \right] \end{aligned} \quad (3.350)$$

Now we move from the probability measure \mathbb{Q} to $\tilde{\mathbb{Q}}$, under which $\tilde{B}(t)$ is a BM. The corresponding Radon-Nikodym derivative is:

$$\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}} \Big|_t = e^{-\theta B(t) - \frac{1}{2}\theta^2 t} \quad \Rightarrow \quad \frac{d\mathbb{Q}}{d\tilde{\mathbb{Q}}} \Big|_t = e^{\theta B(t) + \frac{1}{2}\theta^2 t} = e^{\theta \tilde{B}(t) - \frac{1}{2}\theta^2 t} \quad (3.351)$$

This change of measure leads to:

$$\mathbb{Q}(M^S(t) \leq L, S(t) \geq K) = \mathbb{E}^{\tilde{\mathbb{Q}}} \left[e^{\theta \tilde{B}(t) - \frac{1}{2}\theta^2 t} \mathbb{1}_{M^{\tilde{B}}(t) \leq \frac{1}{\sigma} \ln \frac{L}{S(0)}, \tilde{B}(t) \geq \frac{1}{\sigma} \ln \frac{K}{S(0)}} \right] \quad (3.352)$$

Let's apply this result to an example.

Up-and-out call

The *Up-and-out call* (UOC) is a type of knock-out⁹ barrier option that ceases to exist when the price of the underlying rises above a specific price level, called barrier price. In other words, we get the payoff of the call only if the underlying never reaches a certain upper level L :

$$\text{payoff}_T(\text{UOC}) = (S(T) - K)^+ \mathbb{1}_{S(T) < L \forall t \leq T} \quad (3.353)$$

There are two possible situations:

1. $K < S(0) < L$;
2. $S(0) < K < L$.

Figure 3.21: Case in which $S(0) < K < L$.

In fact, the case $K > L$ is financially meaningless, because the payoff is zero.

Let's consider the case in which $S(0) < K < L$. The price of the UOC is given by:

$$\begin{aligned}
 \text{price}_0(\text{UOC}) &= e^{-rt} \mathbb{E}^{\mathbb{Q}}[(S(T) - K)^+ \mathbf{1}_{M^S(T) < L}] \\
 &= e^{-rt} \mathbb{E}^{\mathbb{Q}}[(S(0)e^{-\sigma \tilde{B}(t)} - K) \mathbf{1}_{M^{\tilde{B}}(t) \leq \frac{1}{\sigma} \ln \frac{L}{S(0)}, \tilde{B}(t) \geq \frac{1}{\sigma} \ln \frac{K}{S(0)}}] \\
 &= e^{-rt} \mathbb{E}^{\tilde{\mathbb{Q}}}[e^{\theta \tilde{B}(t) - \frac{1}{2} \theta^2 t} (S(0)e^{-\sigma \tilde{B}(t)} - K) \mathbf{1}_{M^{\tilde{B}}(t) \leq \frac{1}{\sigma} \ln \frac{L}{S(0)}, \tilde{B}(t) \geq \frac{1}{\sigma} \ln \frac{K}{S(0)}}] \\
 &\stackrel{(a)}{=} e^{-rt} \int_{\frac{1}{\sigma} \ln \frac{K}{S(0)}}^{\frac{1}{\sigma} \ln \frac{L}{S(0)}} db \int_b^{\frac{1}{\sigma} \ln \frac{L}{S(0)}} e^{\theta b - \frac{1}{2} \theta^2 T} (S(0)e^{-\sigma b} - K) (\text{JD}) dm \\
 &\stackrel{(b)}{=} e^{-rt} \int_{\tilde{b}}^{\tilde{m}} db \int_b^{\tilde{m}} e^{\theta b - \frac{1}{2} \theta^2 T} (S(0)e^{-\sigma b} - K) \frac{2(2m-b)}{\sqrt{2\pi t^3}} e^{-\frac{(2m-b)^2}{2t}} dm \\
 &\stackrel{(c)}{=} e^{-rt} \int_{\tilde{b}}^{\tilde{m}} e^{\theta b - \frac{1}{2} \theta^2 T} (S(0)e^{-\sigma b} - K) \left(e^{-\frac{(2\tilde{m}-b)^2}{2t}} - e^{-\frac{(2b-b)^2}{2t}} \right) db \\
 &= \text{complete the squares ecc.}
 \end{aligned} \tag{3.354}$$

where in (a) we recall that

$$\frac{1}{\sigma} \ln \frac{K}{S(0)} < b \leq m < \frac{1}{\sigma} \ln \frac{L}{S(0)}, \quad b \leq m < \frac{1}{\sigma} \ln \frac{L}{S(0)},$$

and JD stands for joint density, in (b) we define

$$\tilde{b} \equiv \frac{1}{\sigma} \ln \frac{K}{S(0)}, \quad \tilde{m} \equiv \frac{1}{\sigma} \ln \frac{L}{S(0)}$$

and in (c) we use the fact that $\frac{2(2m-b)}{\sqrt{2\pi t^3}} e^{-\frac{(2m-b)^2}{2t}}$ is already a derivative.

3.14.3 Barrier options: a modern approach

Let's consider a Wiener process with drift μ and diffusion σ starting at point α , i.e.

Lesson 22 of
29/04/2020

$$\begin{cases} dX(t) = \mu dt + \sigma dW(t) \\ X(0) = \alpha \end{cases} \tag{3.355}$$

and a hitting time

$$T_\beta = \inf\{t \geq 0 : X(t) = \beta\} \tag{3.356}$$

Definition 3.14.4 (Absorbed process). *The X -process absorbed at β is defined by*

$$X(t \wedge T_\beta) = \begin{cases} X(t) & t < T_\beta \\ \beta & t \geq T_\beta \end{cases} \tag{3.357}$$

Figure 3.22: Absorbed process (a) for $\beta > \alpha$ (b) for $\beta < \alpha$.

We are primarily interested in the one-dimensional marginal distribution for $X(t \wedge T_\beta)$, i.e. the distribution at time t of the X -process, absorbed at the point β . The distribution of $X(t \wedge T_\beta)$ is of course a mixed distribution in the sense that it has a point mass at $x = \beta$ (the probability that the process is absorbed prior to time t) and a density:

$$f_{X(t \wedge T_\beta)} = \text{density} + \text{mass in } X = \beta \quad (3.358)$$

This density has its support on the interval (β, ∞) if $\alpha > \beta$, whereas the support is the interval $(-\infty, \beta)$ if $\alpha < \beta$.

Definition 3.14.5 (Density of a normal distribution). *Let $\varphi(x; \mu, \sigma)$ denote the density of a normal distribution with mean μ and variance σ^2 , i.e.*

$$\varphi(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\} \quad (3.359)$$

Proposition 3.14.2. *The density $f_\beta(x; t, \alpha)$ of the absorbed process $X(t \wedge T_\beta)$ is given by*

$$f_{X(t \wedge T_\beta)}(x; t, \alpha) = \varphi(x; \mu t + \alpha, \sigma \sqrt{t}) - \exp\left\{-\frac{2\mu(\alpha - \beta)}{\sigma^2}\right\} \varphi(x; \mu t - \alpha + 2\beta, \sigma \sqrt{t}) \quad (3.360)$$

The support of this density is the interval (β, ∞) if $\alpha > \beta$, and the interval $(-\infty, \beta)$ if $\alpha < \beta$.

Down-and-out contracts

Let's apply this result in a B&S market with

$$\begin{cases} \frac{dS(t)}{S(t)} = r dt + \sigma dW^Q(t) \\ \frac{dS(0)}{S(0)} = r dt \end{cases} \quad (3.361)$$

We are interested in pricing a particular contract. Fix a real number $L < S(0)$, which will act as the barrier, and consider the following contract, which we denote by Z_{LO} :

- If the stock price stays above the barrier L during the entire contract period, then the amount Z is paid to the holder of the contract.
- If the stock price, at some time before the delivery time T , hits the barrier L , then the contract ceases to exist, and nothing is paid to the holder of the contract.

⁹A knock-out option is an option with a built-in mechanism to expire worthless if a specified price level in the underlying asset is reached.

The contract Z_{LO} is called the “down-and-out” version of the contract Z above, and our main problem is to price Z_{LO} . More formally we can describe Z_{LO} as:

$$Z_{LO} = \Phi(S(T)) \mathbb{1}_{S(T) < L} = \begin{cases} \Phi(S_T) & \text{if } S(t) > L \ \forall t \in [0, T] \\ 0 & \text{otherwise} \end{cases} \quad (3.362)$$

Concerning the notation, L as a subscript indicates a “down”-type contract, whereas the letter O indicates that we are considering an “out” claim.

Theorem 3.14.2 (Pricing down-and-out contracts). *The pricing function, denoted by F_{LO} , of a down-and-out contract Z_{LO} is given by*

$$F_{LO}(t, S(t) = s, \Phi) = \left[F(t, s, \Phi_L) - \left(\frac{L}{s} \right)^{\frac{2(r - \frac{\sigma^2}{2})}{\sigma^2}} F \left(t, \frac{L^2}{s}, \Phi_L \right) \right] \mathbb{1}_{s > L} \quad (3.363)$$

where

$$\Phi_L(X) = \Phi(X) \mathbb{1}_{x > L}. \quad (3.364)$$

Notice that the indicator function $\mathbb{1}_{x > L}$ represents an European-style barrier, because it involves the terminal value of the underlying. So, this theorem tells us that the price of the barrier option can be written as a combination of prices of the usual European options.

Proof. Without loss of generality we may set $t = 0$. Assume then that $S(0) = s > L$, and recall that $S(T \wedge T_L)$ denotes the process S with (possible) absorption at L . Using the risk neutral valuation we have:

$$\begin{aligned} F_{LO}(0, S(0) = s, \Phi) &= e^{-rT} \mathbb{E}_{0,s}^{\mathbb{Q}}[Z_{LO}] = e^{-rT} \mathbb{E}_{0,s}^{\mathbb{Q}} \left[\Phi(S(T)) \mathbb{1}_{\inf_{0 \leq t \leq T} S(t) > L} \right] \\ &= e^{-rT} \mathbb{E}_{0,s}^{\mathbb{Q}} \left[\Phi_L(S(T \wedge T_L)) \mathbb{1}_{\inf_{0 \leq t \leq T} S(t) > L} \right] \\ &= e^{-rT} \mathbb{E}_{0,s}^{\mathbb{Q}} [\Phi_L(S(T \wedge T_L))] \\ &\stackrel{(a)}{=} e^{-rT} \int_L^\infty \Phi_L(x) h(x) dx \end{aligned}$$

where $h(x)$ is the density function for the stochastic variable $S(T \wedge T_L)$. From the theory we know that

$$\begin{aligned} S(T) &= se^{\left(r - \frac{\sigma^2}{2} \right) T + \sigma W^{\mathbb{Q}}(T)} \\ &= e^{\ln s + \tilde{r}T + \sigma W^{\mathbb{Q}}(T)} = e^{X(T)} \end{aligned}$$

where we used the notation $\tilde{r} = r - \frac{\sigma^2}{2}$ and where the process X is defined by

$$\begin{cases} dX(t) = \tilde{r} dt + \sigma dW(t) \\ X(0) = \ln s \end{cases}$$

Thus we have

$$S(T \wedge T_L) = e^{X(t \wedge T_{\ln L})}$$

so we may write

$$\mathbb{E}_{0,s}^{\mathbb{Q}} [\Phi_L(S(T \wedge T_L))] = \int_{\ln L}^\infty \Phi_L(e^x) f(x) dx$$

where f is the density of the stochastic variable $X(t \wedge T_{\ln L})$. This density is, however, given by Proposition 3.14.2 as

$$\begin{aligned} f(x) &= \varphi \left(x; \tilde{r}T + \ln s, \sigma\sqrt{T} \right) - \\ &\quad - \exp \left\{ -\frac{2\tilde{r}(\ln s - \ln L)}{\sigma^2} \right\} \varphi \left(x; \tilde{r}T - \ln s + 2\ln L, \sigma\sqrt{T} \right) \\ &= \varphi \left(x; \tilde{r}T + \ln s, \sigma\sqrt{T} \right) - \left(\frac{L}{s} \right)^{\frac{2\tilde{r}}{\sigma^2}} \varphi \left(x; \tilde{r}T + \ln \left(\frac{L^2}{s} \right), \sigma\sqrt{T} \right) \end{aligned}$$

Thus we have

$$\begin{aligned} \mathbb{E}_{0,s}^{\mathbb{Q}}[\Phi_L(S(T \wedge T_L))] &= \int_{\ln L}^{\infty} \Phi_L(e^x) f(x) dx \\ &= \int_{\ln L}^{\infty} \Phi_L(e^x) \varphi \left(x; \tilde{r}T + \ln s, \sigma\sqrt{T} \right) - \\ &\quad - \left(\frac{L}{s} \right)^{\frac{2\tilde{r}}{\sigma^2}} \int_{\ln L}^{\infty} \Phi_L(e^x) \varphi \left(x; \tilde{r}T + \ln \left(\frac{L^2}{s} \right), \sigma\sqrt{T} \right) dx \\ &= \int_{\mathbb{R}} \Phi_L(e^x) \varphi \left(x; \tilde{r}T + \ln s, \sigma\sqrt{T} \right) dx - \\ &\quad - \left(\frac{L}{s} \right)^{\frac{2\tilde{r}}{\sigma^2}} \int_{\mathbb{R}} \Phi_L(e^x) \varphi \left(x; \tilde{r}T + \ln \left(\frac{L^2}{s} \right), \sigma\sqrt{T} \right) dx \end{aligned}$$

Inspecting the last two lines we see that the density in the first integral is the density of $X(T)$ under the usual martingale measure \mathbb{Q} , given the starting value $S(0) = s$. The density in the second integral is, in the same way, the density (under \mathbb{Q}) of $X(T)$, given the starting point $S(0) = L^2/s$. Thus we have

$$\mathbb{E}_{0,s}^{\mathbb{Q}}[\Phi_L(S(T \wedge T_L))] = \mathbb{E}_{0,s}^{\mathbb{Q}}[\Phi_L(S(T))] - \left(\frac{L}{s} \right)^{\frac{2\tilde{r}}{\sigma^2}} \mathbb{E}_{0,\frac{L^2}{s}}^{\mathbb{Q}}[\Phi_L(S(T))]$$

which gives us the pricing function

$$\begin{aligned} F_{LO}(0, s, \Phi) &= e^{-rT} \mathbb{E}_{0,s}^{\mathbb{Q}}[\Phi_L(S(T))] - e^{-rT} \left(\frac{L}{s} \right)^{\frac{2\tilde{r}}{\sigma^2}} \mathbb{E}_{0,\frac{L^2}{s}}^{\mathbb{Q}}[\Phi_L(S(T))] \\ &= \left[F(t, s, \Phi_L) - \left(\frac{L}{s} \right)^{\frac{2\tilde{r}}{\sigma^2}} F \left(t, \frac{L^2}{s}, \Phi_L \right) \right] \mathbf{1}_{s>L} \end{aligned}$$

□

We again emphasize the point of this result: the problem of computing the price for a down-and-out claim reduces to the standard problem of computing the price of an ordinary (related) claim without a barrier.

We also note the fact that down-and-out pricing is a linear operation.

Corollary 3.14.2.1. *For any contract payoffs Φ and Ψ , and for any real numbers α and β , the following relation holds:*

$$F_{LO}(t, s, \alpha\Phi + \beta\Psi) = \alpha F_{LO}(t, s, \Phi) + \beta F_{LO}(t, s, \Psi). \quad (3.365)$$

Proof. The result follows immediately from Theorem 3.14.2 together with the linearity of the ordinary pricing functional F and the linearity of the chopping operation:

$$\begin{aligned} (\alpha\Phi(x) + \beta\Psi(x))_L &= (\alpha\Phi(x) + \beta\Psi(x))\mathbf{1}_{x>L} \\ &= \alpha\Phi(x)\mathbf{1}_{x>L} + \beta\Psi(x)\mathbf{1}_{x>L} \\ &= \alpha\Phi(x)_L + \beta\Psi(x)_L \end{aligned}$$

□

Up-and-out contracts

We now describe the up-and-out version of Z . This is the contract which at the time of delivery, T , will pay Z if the underlying price process during the entire contract period has stayed below the barrier L . If, at some time during the contract period, the price process exceeds L , then the contract is worthless. In formal terms this reads as follows.

$$Z^{LO} = \Phi(S(T)) \mathbb{1}_{S(t) < L \forall t \in [0, T]} = \begin{cases} \Phi(S(T)) & \text{if } S(t) < L \ \forall t \in [0, T] \\ 0 & \text{otherwise} \end{cases} \quad (3.366)$$

The pricing functional for Z^{LO} is denoted by $F^{LO}(t, s, \Phi)$.

L as a superscript indicates an “up”-type contract, whereas the superscript O indicates that the contract is an “out” contract. As in the previous section we will relate the up-and-out contract to an associated standard contract.

Theorem 3.14.3 (Pricing up-and-out contracts). *Consider a fixed T -claim $Z = \Phi(S(T))$. Then the pricing function, F^{LO} , of the corresponding up-and-out contract Z^{LO} is given by*

$$F^{LO}(t, s, \Phi) = \left[F(t, s, \Phi^L) - \left(\frac{L}{s} \right)^{\frac{2\tilde{r}}{\sigma^2}} F \left(t, \frac{L^2}{s}, \Phi^L \right) \right] \mathbb{1}_{s < L} \quad (3.367)$$

with

$$\Phi^L = \begin{cases} \Phi(x) & \text{for } x < L \\ 0 & \text{for } x \geq L \end{cases}, \quad \tilde{r} = r - \frac{\sigma^2}{2}. \quad (3.368)$$

Remark 3.14.1. From eqs. (3.363) and (3.367) we see that the price of a barrier option is given by a difference between the price of a contract and the price of a modified version of the same contract, which is positive. This means that a barrier option is cheaper than the corresponding vanilla option (for example a barrier call option will be cheaper than a vanilla call). On the other hand, entering a barrier option in order to hedge is dangerous, because if the underlying touches the barrier then there every possibility to hedge is lost.

Examples

Let us define the following standard contracts, which will be the basic building blocks in the sequel. Fix a delivery time T . For fixed parameters K and L define the following claims:

- $ST(x) = x$ is the payoff which gives (the price of) one unit of the underlying stock at delivery time T . Basically, it is a long position on the underlying $S(t)$;
- $BO(x) = 1$ is an ordinary zero coupon bond paying one at maturity T ;
- $H(x, L) = \mathbb{1}_{x > L}$ gives the owner one if the value of the underlying stock exceeds L at delivery time T , otherwise nothing is paid out (digital call);
- $C(x, K) = (x - K)^+$ is the ordinary European call with strike price K .

We now list the pricing functions for the standard contracts above.

- the value of ST at time t is equal to the value of the underlying stock at the same time:

$$ST(t, S(t) = s) = \text{price}_t(ST) = e^{-rt}(T-t) \mathbb{E}_{t,s}^{\mathbb{Q}}[S(T)] = S(t) = s \quad (3.369)$$

- the value of BO at time t is

$$BO(t, s) = \text{price}_t(BO) = e^{-r(T-t)} \quad (3.370)$$

- the value of H is calculated by using the risk neutral methodology:

$$\begin{aligned} H(t, s, L) &= \text{price}_t(H(S(T), L)) = e^{-r(T-t)} \Phi(d_2) \\ &= e^{-r(T-t)} \Phi\left(\frac{\tilde{r}(T-t) + \ln\left(\frac{s}{L}\right)}{\sigma\sqrt{T-t}}\right) \end{aligned} \quad (3.371)$$

- The value of C is given by the B&S formula.

Notice that

$$BO_L(x) = 1 \cdot \mathbb{1}_{x>L} = H(x, L).$$

Lesson 23 of
30/04/2020

This fact has the following consequence.

Proposition 3.14.3 (Defaultable bond). *The down-and-out bond with barrier L is priced by the formula*

$$BO_{LO}(t, s) = \left[H(t, s, L) - \left(\frac{2\tilde{r}}{\sigma^2} \right) H\left(t, \frac{L^2}{s}, L\right) \right] \mathbb{1}_{s>L}. \quad (3.372)$$

This contract will thus pay out 1 dollar at time T only if the stock price is above the level L during the entire contract period.

Now we consider something which is not quoted in the market but that is a building block to price something which indeed is quoted.

Proposition 3.14.4 (Down-and-out on stock). *The down-and-out contract on the underlying stock is given by*

$$\begin{aligned} ST_{LO}(t, s) &= \left[LH(t, s, L) - L \left(\frac{2\tilde{r}}{\sigma^2} \right) H\left(t, \frac{L^2}{s}, L\right) + \right. \\ &\quad \left. + C(t, s, L) - \left(\frac{2\tilde{r}}{\sigma^2} \right) C\left(t, \frac{L^2}{s}, L\right) \right] \mathbb{1}_{s>L}. \end{aligned} \quad (3.373)$$

Proof. From Theorem 3.14.2 we have that

$$ST_{LO} = \left[F(t, s, ST_L) - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2} \right)} F\left(t, \frac{L^2}{s}, ST_L\right) \right] \mathbb{1}_{s>L} \quad (\diamond)$$

Recall that

$$ST_L(x) = x \mathbb{1}_{x>L}$$

We can write it as

$$ST_L(x) = (x - L + L) \mathbb{1}_{x>L} = (x - L)^+ + L \mathbb{1}_{x>L}$$

which is a payoff of a call option plus a digital option, which gives a up shift at $x = L$:

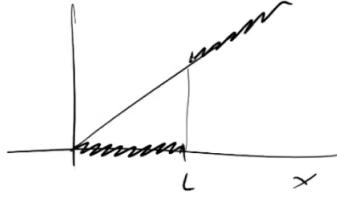
$$ST_L(x) = LH(x, L) + C(x, L)$$

Substituting this into (\diamond) and using the linearity (Corollary 3.14.2.1) we get

$$ST_{LO} = \left[F(t, s, LH(S(T), L) + C(S(T), L)) - \right.$$

$$\begin{aligned}
& - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{\tau}}{\sigma^2}\right)} F \left(t, \frac{L^2}{s}, LH(S(T), L) + C(S(T), L) \right) \Big] \mathbb{1}_{s>L} \\
= & \left[LF(t, s, LH(S(T), L)) - L \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{\tau}}{\sigma^2}\right)} F \left(t, \frac{L^2}{s}, H(S(T), L) \right) + \right. \\
& \quad \left. + F(t, s, C(S(T), L)) - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{\tau}}{\sigma^2}\right)} F \left(t, \frac{L^2}{s}, C(S(T), L) \right) \right] \mathbb{1}_{s>L}.
\end{aligned}$$

□

Figure 3.23: Graphical representation of $ST_L(x)$.

We now turn to a more interesting example, which is quoted in the market: the **down-and-out European call** (DOC) with strike price K . The payoff of this contract is given by

$$\text{payoff}_T(\text{DOC}) = (S(T) - K)^+ \mathbb{1}_{S(t)>L, \forall t \in [0, T]} \quad (3.374)$$

In order to price this contract we have to consider the modified payoff

$$(S(T) - K)^+ \mathbb{1}_{S(T)>L} = (S(T) - K) \mathbb{1}_{S(T)>K} \mathbb{1}_{S(T)>L} \quad (3.375)$$

According to the possible values of L and K we have different situations:

- if $K > L$ then the indicator function $\mathbb{1}_{S(t)>L}$ is redundant and the modification of the payoff coincides with the original payoff. However, even if the modification of the payoff coincides with the original payoff, we cannot say that the barrier option price coincides with the vanilla call option price, as we will see.
- if $K > L$ then we can exercise the contract but due to the presence of the indicator function of L the terminal payoff can be different. In this case the indicator function $\mathbb{1}_{S(t)>K}$ is redundant and the modified payoff is given by

$$(S(T) - K) \mathbb{1}_{S(T)>L}$$

This can be seen as the payoff of a combination of a call option and a digital option. In fact, we can write

$$\begin{aligned}
(S(T) - K) \mathbb{1}_{S(T)>L} &= (S(T) - L + L - K) \mathbb{1}_{S(T)>L} \\
&= (S(T) - L)^+ + (L - K) \mathbb{1}_{S(T)>L}
\end{aligned}$$

The result for the pricing function in these two cases is as follows.

Proposition 3.14.5 (Down-and-out call pricing). *The down-and-out European call option is priced as follows. For $L < K$:*

$$C_{LO}(t, s, K) = \left[C(t, s, K) - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{\tau}}{\sigma^2}\right)} C \left(t, \frac{L^2}{s}, K \right) \right] \mathbb{1}_{s>L}. \quad (3.376)$$

For $L > K$:

$$C_{LO}(t, s, K) = \left\{ C(t, s, K) + (L - K)H(t, s, L) - \left(\frac{L}{s} \right)^{\left(\frac{2\bar{r}}{\sigma^2}\right)} \left[C\left(t, \frac{L^2}{s}, K\right) + (L - K)H\left(t, \frac{L^2}{s}, K\right) \right] \right\} \mathbb{1}_{s>L} \quad (3.377)$$

In Section 2.3.1 we used this linearity to prove the standard put–call parity relation for standard European options, and we can now derive the put–call parity result for down-and-out options, often called **DOC-DOP parity**. Recall that the payoff of a put option is given by

$$\begin{aligned} P(x, K) &= (K - x)^+ = K - x - (x - K)^+ \\ &= K \cdot BO(x) - ST(x) + C(x, K) \end{aligned}$$

Using Corollary 3.14.2.1 we immediately have the following result.

Proposition 3.14.6 (Down-and-out put pricing). *The down-and-out put price P_{LO} , and call price C_{LO} , are related by the formula*

$$P_{LO}(t, s, K) = K \cdot B_{LO}(t, s) - ST_{LO}(t, s) + C_{LO}(t, s, K). \quad (3.378)$$

Note that when $L = 0$ we have the usual put–call parity.

We now want to compute the price of a European **up-and-out put option** (UOP) with barrier L and strike price K . As usual, we have to consider a modification of the payoff:

$$P^L(x, K) = (K - x) + \mathbb{1}_{x < L} = (K - x)\mathbb{1}_{x < K}\mathbb{1}_{x < L} \quad (3.379)$$

There are two possibilities:

- if $L > K$ then the indicator function $\mathbb{1}_{x < L}$ is redundant and the modification of the payoff is the same of the vanilla put:

$$P^L(x, K) = (K - x)^+$$

Again, this does not mean that the price of this barrier option is the same of the vanilla put option.

- if $L < K$ then the indicator function $\mathbb{1}_{x < K}$ is redundant and the modification of the payoff is given by

$$\begin{aligned} P^L(x, K) &= (K - x)\mathbb{1}_{x < L} = (L - x)\mathbb{1}_{x < L} + (K - L)\mathbb{1}_{x < L} \\ &= (L - x)^+ + (K - L) - (K - L)\mathbb{1}_{x < L} \end{aligned}$$

which is the combination of the payoff of a put option and two digital options.

Now we know how to price every building block of the up-and-out, so we can give the pricing function.

Proposition 3.14.7 (Up-and-out put pricing). *The price of an up-and-out European put option is given by the following formulas. For $L > K$:*

$$P^{LO}(t, s, K) = \left[P(t, s, K) - \left(\frac{L}{s} \right)^{\left(\frac{2\bar{r}}{\sigma^2}\right)} P\left(t, \frac{L^2}{s}, K\right) \right] \mathbb{1}_{s < L}. \quad (3.380)$$

For $L > K$:

$$\begin{aligned} P^{LO}(t, s, K) = & \left\{ P(t, s, L) + (K - L)H(t, s, L) - \right. \\ & - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2}\right)} \left[P \left(t, \frac{L^2}{s}, L \right) - (K - L)H \left(t, \frac{L^2}{s}, L \right) \right] + \\ & \left. \left[1 - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2}\right)} \right] (K - L)e^{-r(T-t)} \right\} \mathbb{1}_{s>L}. \end{aligned} \quad (3.381)$$

In order to price a **up-and-out call option** (UOC) we can use the put-call parity. Alternatively, if we use the “direct approach”, we have to consider the payoff modification

$$\begin{aligned} (S - K)^+ \mathbb{1}_{s < L} &= (S - K)^+ \mathbb{1}_{s < L} \\ &= (S - K) \mathbb{1}_{s > K} (1 - \mathbb{1}_{s > L}) \\ &= (S - K)^+ - (S - K) \mathbb{1}_{s > K} \mathbb{1}_{s > L} \end{aligned}$$

and then consider the different cases. [homework](#)

Knock-in contracts

Knock-in contracts are contracts which will start to exist if and only if the price of the underlying stock hits a prespecified barrier level at some time during the contract period. We thus fix a standard T -claim of the form $Z = \Phi(S(T))$, and we also fix a barrier L .

We start by studying the **down-and-in** version of Z , which is defined as follows:

$$Z_{LI} = \begin{cases} 0 & \text{if } S(t) > L \ \forall t \in [0, T] \\ \Phi(S(T)) & \text{otherwise} \end{cases} \quad (3.382)$$

L as a subscript indicates a “down” contract, whereas the subscript I denotes a “in” contract. Pricing a down-and-in contract turns out to be fairly easy, since we can in fact price it in terms of the corresponding down-and-out contract.

Lemma 3.14.4 (In-out parity). $F_{LI}(t, s, \Phi) = F(t, s, \Phi) - F_{LO}(t, s, \Phi)$, $\forall s$.

Proof. If, at time t , we have a portfolio consisting of a down-and-out version of Z as well as a down-and-in version of Z (with the same barrier L) then

$$Z_{LI} + Z_{LO} = \Phi,$$

so we wil receive exactly Z at time T . We thus have

$$F_{LI}(t, s, \Phi) = F(t, s, \Phi) - F_{LO}(t, s, \Phi).$$

□

We can now formulate the basic result.

Proposition 3.14.8 (Down-and-in pricing). *Consider a fixed T -contract $Z = \Phi(S(T))$. Then the price of the corresponding down-and-in contract Z_{LI} is given by*

$$F_{LI}(t, s, \Phi) = \left[F(t, s, \Phi^L) - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2}\right)} F \left(t, \frac{L^2}{s}, \Phi^L \right) \right] \mathbb{1}_{s < L}. \quad (3.383)$$

Proof. From the equality $\Phi_L + \Phi^L = \Phi$ we have

$$\begin{aligned} F_{LI}(t, s, \Phi) &= F(t, s, \Phi) + F_{LO}(t, s, \Phi) \\ &\stackrel{3.376}{=} \cancel{F(t, s, \Phi_L)} + F(t, s, \Phi^L) - \left[\cancel{F(t, s, \Phi_L)} - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2}\right)} F(t, s, \Phi_L) \right] \\ &= \left[F(t, s, \Phi^L) - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2}\right)} F\left(t, \frac{L^2}{s}, \Phi^L\right) \right]. \end{aligned}$$

□

The treatment of up-and-in contracts is of course parallel to down-and-in contracts.

Proposition 3.14.9 (Up-and-in pricing). *Consider a fixed T -contract $Z = \Phi(S(T))$. Then the price of the corresponding up-and-in contract Z^{LI} is given by*

$$F^{LI}(t, s, \Phi) = \left[F(t, s, \Phi_L) - \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2}\right)} F\left(t, \frac{L^2}{s}, \Phi_L\right) \right] \mathbb{1}_{s < L}. \quad (3.384)$$

Proof. Same argument as before. □

As an example, we give the pricing formula for a **down-and-in European call** with strike price K .

Proposition 3.14.10 (Down-and-in call pricing). *For $s > L$ the down-and-in European call option is priced as follows. For $L < K$*

$$C_{LI}(t, s, \Phi) = \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2}\right)} C\left(t, \frac{L^2}{s}, K\right). \quad (3.385)$$

For $L > K$:

$$C_{LI}(t, s, \Phi) = \left(\frac{L}{s} \right)^{\left(\frac{2\tilde{r}}{\sigma^2}\right)} \left[C\left(t, \frac{L^2}{s}, K\right) + (L - K)H\left(t, \frac{L^2}{s}, L\right) \right] - (L - K)H(t, s, L). \quad (3.386)$$

Lookbacks

Lookback options are contracts which at the delivery time T allow us to take advantage of the realized maximum or minimum of the underlying price process over the entire contract period. Typical examples are:

$$\begin{aligned} \Phi(T) &= (S(T) - \min_{t \leq T} S(t)) && \text{Lookback call} \\ \Phi(T) &= \max_{t \leq T} S(t) - S(t) && \text{Lookback put} \\ \Phi(T) &= \left(\max_{t \leq T} S(t) - K \right)^+ && \text{Forward lookback call} \\ \Phi(T) &= \left(K - \min_{t \leq T} S(t) \right)^+ && \text{Forward lookback put} \end{aligned}$$

In order to price this kind of contracts, we need the distribution of the maximum

$$M_S(t) = \max_{t \leq T} S(t) \quad (3.387)$$

and of the minimum

$$m_S(t) = \min_{t \leq T} S(t) \quad (3.388)$$

Recall that

$$S(t) = e^{\ln s + \tilde{r}t + \sigma W(t)} = e^{X(t)}$$

where

$$\begin{cases} dX(t) = \tilde{r} dt + \sigma dW(t) \\ X(0) = \ln s \end{cases}$$

Therefore

$$M_S(t) = e^{M_X(t)}$$

The distribution for $M_X(T)$ is known to us from Proposition 3.14.2. Thanks to this proposition we obtain the distribution function, F , for $M_X(T)$:

$$F_{\max_{s \leq t} X_s}(x) = \left[\varphi\left(\frac{x - \alpha - \mu t}{\sigma\sqrt{t}}\right) - \exp\left\{-\frac{2\mu(\alpha - \beta)}{\sigma^2}\right\} \varphi\left(-\frac{x - \alpha + \mu t}{\sigma\sqrt{t}}\right) \right] \mathbb{1}_{x \geq \ln s} \quad (3.389)$$

From this expression we can compute the density function $f = F'$ and then the expected value

$$\mathbb{E}^{\mathbb{Q}} \left[\max_{t \leq T} S(t) \right] = \mathbb{E}^{\mathbb{Q}} \left[e^{M_X(T)} \right] = \int_{\ln s}^{\infty} e^x f(x) dx \quad (3.390)$$

After a series of elementary, but extremely tedious, partial integrations we end up with the following result.

Proposition 3.14.11. *The price, at $t = 0$, of the lookback put is given by*

$$\text{price}_0(\text{lookback put}) = -s\varphi(-d) + se^{-rT} \varphi(-d + \sigma\sqrt{T}) \left(1 - \frac{\sigma^2}{2r} + \frac{s\sigma^2}{2r} N(d) \right) \quad (3.391)$$

where

$$d = \frac{rT + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}. \quad (3.392)$$

3.15 Bonds and interest rates

Definition 3.15.1 (Zero coupon bond). *A zero coupon bond with maturity date T , also called a T -bond, is a contract which guarantees the holder 1 unit of money to be paid on the date T . The price at time t of a bond with maturity date T is denoted by $p(t, T)$.*

Lesson 24 of
06/05/2020

In the real market not all maturities are traded, in fact there are some fixed maturities. According to the particular maturity we can invert the zero coupon bond in order to deduce the corresponding interest rate. So, the quantity which is traded is the zero coupon bond price, not the interest rate. However, here we assume that there are quoted bonds for any maturity.

Assumption 3.15.1. We assume that:

- there exists a market for T -bonds for every $T > 0$;
- the relation $p(t, t) = 1$ holds for all t ;

- for each fixed t , the bond price $p(t, T)$ is differentiable w.r.t. the time of maturity T .

Given this bond market, we may now define a number of interest rates, and the basic construction is as follows. Suppose that we are standing at time t , and let us fix two other points in time, S and T , with $t < S < T$. The immediate project is to write a contract at time t which allows us to make an investment of one unit for money at time S , and to have a deterministic rate of return, determined at the contract time t , over the interval $[S, T]$. This can easily be achieved as follows:

1. Go short: at time t we sell one S -bond. This will give us $p(t, S)$ dollars (for example).
2. Go long: we use this income to buy exactly $p(t, S)/p(t, T)$ T -bonds. Thus our net investment at time t equals zero.
3. At time S the S -bond matures, so we are obliged to pay out one dollar.
4. At time T the T -bonds mature at one dollar a piece, so we will receive the amount $p(t, S)/p(t, T)$ dollars.
5. The net effect of all this is that, based on a contract at t , an investment of one dollar at time S has yielded $p(t, S)/p(t, T)$ dollars at time T .
6. Thus, at time t , we have made a contract guaranteeing a riskless rate of interest over the future interval $[S, T]$. Such an interest rate is called *forward rate*.

	$t = S$	$t = T$	
1)	$+p(t, S)$	−1	0
2)	$-p(t, T) \frac{p(t, S)}{p(t, T)}$	0	$\frac{p(t, S)}{p(t, T)}$
	0	−1	$\frac{p(t, S)}{p(t, T)} \in \mathcal{F}_t$

We now go on computing the relevant interest rates implied by the construction above. The *simple forward rate* $F(t, S, T)$, is the solution to the equation

$$1 + F(t, S, T)(T - S) = \frac{p(t, S)}{p(t, T)} \quad (3.393)$$

that is

$$F(t, S, T) = \frac{1}{T - S} \left(\frac{p(t, S)}{p(t, T)} - 1 \right) \quad (3.394)$$

where $(T - S)$ is called *tenor*. If $t = S$ we are considering a particular forward interest rate for which the capitalization starts immediately, the so called *simple spot rate*:

$$F(S, S, T) = \frac{1}{T - S} \left(\frac{1}{p(S, T)} - 1 \right) \quad (3.395)$$

The simple spot rate is also called *yield to maturity*.

At the beginning of the course we defined the forward contract as a linear contract which gives the obligation to exchange one asset against a fixed strike. Here the notion is the same, in the sense that we exchange an interest rate with respect to the corresponding strike rate.

Let's introduce the notion of *forward rate agreement*, which is the analog of forward contracts for the interest rates. A forward rate agreement (FRA) is a contract, by

convention entered into at $t = 0$, where the parties (a lender and a borrower) fix the rate of interest R to be paid on an agreed upon date in the future period $[S, T]$.

$$\text{payoff}_T(\text{FRA}) = (T - S)(F(S, S, T) - R) \quad (3.396)$$

So, in this case, R plays the role of the strike of the forward. If we suppose that the payoff of the FRA is delivered at time S , the price of the FRA at time t is given by the risk neutral methodology

$$\text{price}_{t,FRA}(t, S, T, R) = "e^{-r(T-t)}" \mathbb{E}_t^{\mathbb{Q}}[\text{payoff}_T(\text{FRA})]$$

However, we don't know both the interest rate to put in the discounting factor and the probability measure \mathbb{Q} . Moreover, we know that in presence of a generic interest rate we have to keep the discounting factor inside the expected value

$$\text{price}_{t,FRA}(t, S, T, R) = \mathbb{E}_t^{\mathbb{Q}} \left["e^{-\int_t^T r_s ds}" \text{payoff}_T(\text{FRA}) \right]$$

In order to disentangle this product, we can consider a change of numéraire using the zero coupon bond expiring at maturity

$$\begin{aligned} \text{price}_{t,FRA}(t, S, T, R) &= p(t, T) \mathbb{E}_t^{\mathbb{Q}^T} [\text{payoff}_T(\text{FRA})] \\ &= p(t, T) \mathbb{E}_t^{\mathbb{Q}^T} [(T - S)(F(S, S, T) - R)] \end{aligned}$$

where \mathbb{Q}^T is the forward risk neutral probability measure. Now the discounting problem is no more present and we go on with the calculation:

$$\begin{aligned} \text{price}_{t,FRA}(t, S, T, R) &= p(t, T) \mathbb{E}_t^{\mathbb{Q}^T} [(T - S)(F(S, S, T) - R)] \\ &= p(t, T) \mathbb{E}_t^{\mathbb{Q}^T} \left[\cancel{(T-S)} \left(\frac{1}{\cancel{T-S}} \left(\frac{1}{p(S, T)} - 1 \right) - R \right) \right] \\ &= p(t, T) \mathbb{E}_t^{\mathbb{Q}^T} \left[\left(\frac{p(S, S)}{p(S, T)} - 1 \right) - (T - S)R \right] \\ &\stackrel{(a)}{=} p(t, T) \left(\left(\frac{p(t, S)}{p(t, T)} - 1 \right) - (T - S)R \right) \end{aligned}$$

where in (a) we use the fact that $\frac{p(S, S)}{p(S, T)}$ is a \mathbb{Q}^T martingale (it is divided by the numéraire $p(t, T)$).

Now we can find the value of R that makes zero the value of the FRA contract (i.e. that makes the FRA a fair contract):

$$\text{price}_{t,FRA} = 0 \Leftrightarrow R = \frac{1}{(T - S)} \left(\frac{p(t, S)}{p(t, T)} - 1 \right) = F(t, S, T) \quad (3.397)$$

So, the forward price (i.e. the strike price that at the inception of a forward contract makes the value of the contract zero) of the FRA is given by the forward interest rate. This leads to the following proposition.

Proposition 3.15.1. *The simple forward rate is given by*

$$F(t, S, T) = \mathbb{E}_t^{\mathbb{Q}^T} [F(S, S, T)] = \frac{1}{T - S} \left(\frac{p(t, S)}{p(t, T)} - 1 \right) \quad (3.398)$$

So, there is an exact relation between forward rates and bond prices.

3.15.1 The crisis of 2007

Before the crisis there were some quoted simple spot rates called XBOR, where X changes according to the country. For example, in the US there was the LIBOR rate, which is the standard simple spot rate:

$$\text{LIBOR} = F(S, S, T) = \frac{1}{T - S} \left(\frac{1}{p(S, T)} - 1 \right) \quad (3.399)$$

LIBOR stands for London Inter-Bank Offered Rate, and it is the interest rate at which different investment banks can lend or borrow money to each other. The LIBOR is managed by the Intercontinental Exchange (ICE), which in turn is managed and regulated by the British Bankers Association (BBA) and the ICE Benchmark Administration (IBA).

In the European market there were the EURIBOR, which is managed by the European Money Markets Institute (EMMI).

All these institutions were created in order to manage the behavior of the XBORs, in order to avoid manipulations of these rates, which were quoted in the market. The problem is that the LIBOR was an average interest rate calculated through submissions of interest rates by major banks across the world¹⁰. Since this methodology was based only on the declarations of the banks and not on traded contracts, some banks were able to manipulate this rate. In fact, there were a series of fraudulent actions connected to the LIBOR: the scandal arose when it was discovered that banks were falsely inflating or deflating their rates to profit from trades, or to give the impression that they were more creditworthy than they were.

If for example we consider a *forward LIBOR*, before the crisis it was possible to express it in terms of the zero coupon bond price, i.e. the usual forward rates

$$L(t, S, T) = \mathbb{E}_t^{\mathbb{Q}^T} (L(S, S, T)) = \frac{1}{T - S} \left(\frac{p(t, S)}{p(t, T)} - 1 \right) = F(t, S, T)$$

So, by construction, LIBOR rates were related to zero coupon bonds prices. The problem is that if we have two maturities, S and T , at time t our counterpart may have a certain credibility but at time S it can change, so the counterpart may not be able to honor its debt at time T . This issue becomes bigger as the tenor $T - S$ becomes larger.

After the crisis nobody was testing its counterpart, so it was not possible to express the LIBOR in terms of the zero coupon price. However, it is still possible to introduce the notion of forward LIBOR rate as

$$L(t, S, T) := \mathbb{E}_t^{\mathbb{Q}^T} [L(S, S, T)]$$

but now it is no more related to the price of the zero coupon bond. In fact, now we have that

$$L(t, S, T) > F(t, S, T)$$

So, now the LIBOR rate becomes very risky, because in addition to the variability of the market there is also the variability in the possible degradation of the credit scoring of the firm issuing the bond. Moreover, there is the so called *rollover effect*: if for example we borrow an amount of money today for a three months operation we have a certain LABOR rate, but after three months the condition of both counterparts

¹⁰ Every day at 11:30 AM the banks were asked to declare which was their typical offer rate. Once they collected all these information they cut the 25% of the lower value and of the higher and they considered the middle offer in the banks panel and the LIBOR was obtained by computing a ponderation of these values.

can be different, so we cannot deduce the corresponding LIBOR in six months by putting a rollover position in terms of a spot times a forward.

In other words, there is a *spread* between the forward LIBOR and the usual forward rate associated to a riskless operation, which typically is a function of the tenor $\Delta \equiv T - S$:

$$s(t, \Delta) = L(t, S, T) - F(t, S, T) \quad (3.400)$$

Of course, it holds that

$$s(t, 3 \text{ months}) < s(t, 6 \text{ months}) < \dots$$

So, the LIBOR is quoted for some fixed tenors (one/two weeks, one/two/three/six/nine/twelve months).

The LIBOR is too risky in order to be considered as the corresponding interest rate used to discount or capitalize cash flows, because there is a bias introduced by the credit risk. Then, what can we choose as “riskless” proxy? In general, it depends on the reputation of the institution that issues the obligation. However, in the interbank market there is no “riskless” proxy. What is typically used as proxy is the so called *overnight rate*, which is the interest rate for operations with maturity one day (the idea is that since the spread is smaller for smaller tenors, we consider the smallest maturity in order to have the less risky proxy). In the EU zone the overnight rate is called Euro Overnight Index Average (EONIA) while in the US zone it is called Federal Fund (FF). Another important quantity is the Overnight Index Swap (OIS), which is an interest rate swap involving the overnight rate (floating leg) being exchanged for a fixed interest rate (fixed leg).

We can also define different spreads, for example:

$$(\text{LIBOR swap rate}) - (\text{OIS}) = \text{LIBOR-OIS spread}$$

(this was supposed to be zero before the crisis) or

$$(\text{EURIBOR swap rate}) - (\text{EONIA swap rate}) = \text{EURIBOR-EONIA swap spread}$$

$$(\text{LIBOR swap rate 6 months}) - (\text{LIBOR swap rate 3 months}) = \text{basis swap spread}$$

The basis swap spread represents the fact that it is more risky to borrow money in 6 months than in 3 months.

However, our general assumption is that the forward rate can be considered as a proxy of the OIS and can be written in term of zero coupon bond prices:

$$F(t, S, T) = L^{\text{OIS}}(t, S, T) = \frac{1}{T - S} \left(\frac{p(t, S)}{p(t, T)} - 1 \right) \quad (3.401)$$

On the contrary, the forward LIBOR rate can only be expressed as an expected value under the forward risk measure:

$$L(t, S, T) = \mathbb{E}_t^{\mathbb{Q}^T} [L(S, S, T)] \neq F(t, S, T) \quad (3.402)$$

Remark 3.15.1. Since we know that $L(t, S, T) > F(t, S, T)$, we might think of adjusting the definition of LIBOR defining an artificial zero coupon bond $\bar{p}(S, T)$ – called *fictitious zero coupon bond* – such that

$$L(S, S, T) = \frac{1}{T - S} \left(\frac{1}{\bar{p}(S, T)} - 1 \right)$$

The problem is that, since $\bar{p}(S, T)$ is an artificial quantity, it is not quoted in the market and therefore it cannot be a numéraire under the forward probability measure \mathbb{Q}^T . This means that if we use the fictitious zero coupon bond we cannot compute forward contracts.

Remark 3.15.2. In 2021 the LIBOR rate will be replaced by a more reliable rate, which will be no more based on banks' declarations but on traded and certified contracts. It is called Secured Overnight Financing Rate (SOFR) and it basically is an overnight rate¹¹.

3.16 Relations between forward rate, zero coupon bond and short rate

Lesson 25 of
07/05/2020

Remark 3.16.1. For $S - T = \Delta \rightarrow 0$, the LIBOR rate tends to the *spot short rate*, $L(t, t, t + \Delta)$.

So, we can recover the notion of short rate in terms of LIBOR. We just need a change of notation.

Definition 3.16.1 (Continuously compounded forward rate). *The continuously compounded forward rate is defined by the equation*

$$e^{R(t,S,T)(T-S)} = \frac{p(t, S)}{p(t, T)} \quad (3.403)$$

from which we get

$$R(t, S, T) = \frac{1}{T - S} \ln \left(\frac{p(t, S)}{p(t, T)} \right). \quad (3.404)$$

Definition 3.16.2 (Continuously compounded spot rate). *The continuously compounded spot rate is defined as*

$$R(S, T) = \frac{1}{T - S} \ln p(S, T). \quad (3.405)$$

The continuously compounded spot rate plays the role of the yield in this different notation.

Definition 3.16.3 (Instantaneous forward rate). *The instantaneous forward rate is defined as*

$$f(t, T) = -\frac{\partial \ln p(t, T)}{\partial T} \quad (3.406)$$

Basically,

$$f(t, T) = \lim_{\Delta \rightarrow 0} (\ln p(t, T) + \ln p(t, T + \Delta)) \quad (3.407)$$

In order to well define the relation between the zero coupon bond and the instantaneous forward rate, we can express the price of the zero coupon bond as follows:

$$p(t, T) = \exp \left\{ - \int_t^T f(t, u) du \right\} \quad (3.408)$$

Of course, we can introduce also the instantaneous short interest rate.

Definition 3.16.4 (Instantaneous interest rate). *The instantaneous short interest rate is defined as*

$$r(t) = f(t, t). \quad (3.409)$$

¹¹ <https://www.investopedia.com/secured-overnight-financing-rate-sofr-4683954>

We note that spot rates are forward rates where the time of contracting coincides with the start of the interval over which the interest rate is effective, i.e. $t = S$. The instantaneous forward rate is the limit of the continuously compounded forward rate when $S \rightarrow T$. It can thus be interpreted as the riskless rate of interest, contracted at t , over the infinitesimal interval $[T, T + dT]$.

We now go on to define the money account process B :

$$B(t) = e^{\int_0^t r(s) ds} \quad (3.410)$$

where $r(s)$ is the instantaneous interest rate. This riskless asset can be seen as the self financing “rolling-over” trading strategy where for all t we invest in a zero coupon bond starting at time t with maturity $t + dt$. In other words, the savings account corresponds to infinite operations of capitalization from time t to time $t + dt$ where for each of them we introduce a spot rate $r(s)$ and we immediately reinvest the cash we get.

Now, since

$$p(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} \right] = e^{-\int_t^T f(t, u) du} \quad (3.411)$$

we are able to link a deterministic expression to a stochastic one. So, the zero coupon bond $p(t, T)$, the short rate $r(t)$ and the forward rate $f(t, T)$ are closely related. This “static” relationship of course leads to a “dynamic” relationship. Therefore, we want to understand which is the relation between $dp(t, T)$, $df(t, T)$ and $dr(t)$.

Let’s start from the following general dynamics:

$$dr(t) = a(t) dt + b(t) \cdot dW(t) \quad (3.412)$$

$$\frac{dp(t, T)}{p(t, T)} = m(t, T) dt + v(t, T) \cdot dW(t) \quad (3.413)$$

$$df(t, T) = \alpha(t, T) dt + \sigma(t, T) \cdot dW(t) \quad (3.414)$$

where $W(t)$ is the Brownian motion driving the three processes, which can be a vector (we can use the same Brownian motion because all the three processes are responsible for the fluctuations of the interest rate). The only assumptions we make are:

- $m, v, \alpha, \sigma \in C^1$ with respect to T ;
- all processes are assumed to be regular enough to allow us to differentiate under the integral sign as well as to interchange the order of integration.

Under these regularity assumptions we have the following result.

Theorem 3.16.1. *The following relationships hold:*

- *If we know the dynamics of $p(t, T)$, then:*

$$\alpha(t, T) = \frac{\partial v(t, T)}{\partial T} v(t, T) - \frac{\partial m(t, T)}{\partial T} \quad (3.415)$$

$$\sigma(t, T) = -\frac{\partial v(t, T)}{\partial T} \quad (3.416)$$

- *If we know the dynamics of $f(t, T)$, then:*

$$a(t) = \frac{\partial f(t, T)}{\partial T} + \alpha(t, T) \quad (3.417)$$

$$b(t) = \sigma(t, t) \quad (3.418)$$

- If we know the dynamics of $f(t, T)$, then:

$$\frac{dp(t, T)}{p(t, T)} = \left(r(t) + A(t, T) + \frac{1}{2} \|S(t, T)\|^2 \right) dt + S(t, T) \cdot dW(t) \quad (3.419)$$

where

$$A(t, T) = - \int_t^T \alpha(t, s) ds \quad (3.420)$$

$$S(t, T) = - \int_t^T \sigma(t, s) ds. \quad (3.421)$$

Proof. 1. From eqs. (3.413) and (3.406), by applying the Itô formula we have that

$$\begin{aligned} df(t, T) &= - \frac{\partial(d \ln p(t, T))}{\partial T} \\ &= - \frac{\partial}{\partial T} \left(\left(m(t, T) - \frac{1}{2} v^2(t, T) \right) dt + v(t, T) \cdot dW(t) \right) \\ &= \left(- \frac{\partial m(t, T)}{\partial T} + \frac{\partial v(t, T)}{\partial T} \cdot v(t, T) \right) dt - \frac{\partial v(t, T)}{\partial T} dW(t) \end{aligned}$$

2. By integrating the forward rate dynamics (3.414) we get

$$f(t, T) - f(0, T) = r(t) - f(0, t) = \int_0^t \alpha(s, T) ds + \int_0^t \sigma(s, T) dW(s) \quad (\star)$$

Now we can write

$$\begin{aligned} \alpha(s, t) &= \alpha(s, s) + \int_s^t \frac{\partial \alpha}{\partial T} du \\ \alpha(s, t) &= \alpha(s, s) + \int_s^t \frac{\partial \sigma}{\partial T} du \end{aligned}$$

and taking $t = T$ we can insert this into (\star) , getting

$$\begin{aligned} r(t) - f(0, t) &= \int_0^t \alpha(s, s) ds + \int_0^t \int_s^t \frac{\partial \alpha}{\partial T} du ds + \\ &\quad + \int_0^t \sigma(s, s) dW(s) + \int_0^t \int_s^t \frac{\partial \sigma}{\partial T} du ds \end{aligned} \quad (*)$$

Using stochastic Fubini in order to change the order of integration we obtain

$$\begin{aligned} r(t) &= f(0, t) + \int_0^t \alpha(s, s) ds + \int_0^t du \int_0^u \frac{\partial \alpha}{\partial T} ds + \\ &\quad + \int_0^t \sigma(s, s) dW(s) + \int_0^t du \int_0^u \frac{\partial \sigma}{\partial T} ds \end{aligned}$$

Now, using Itô we can take the differential:

$$\begin{aligned} dr(t) &= \left[\frac{\partial f(s, t)}{\partial T} dt + \alpha(t, t) dt + \left(\int_0^t \frac{\partial \alpha}{\partial T} ds \right) dt + \right. \\ &\quad \left. + \sigma(t, t) dW(t) + \left(\int_0^t \frac{\partial \sigma}{\partial T} dW(s) \right) dt \right] \end{aligned} \quad (\diamond)$$

From eq. (*) we have that

$$\frac{\partial f(t, t)}{\partial T} = \frac{\partial f(0, t)}{\partial T} + \int_0^t \frac{\partial \alpha(s, t)}{\partial T} ds + \int_0^t \frac{\partial \sigma(s, t)}{\partial T} dW(s)$$

which is exactly the highlighted term in eq. (◇). So, we obtain the result:

$$dr(t) = \left(\frac{\partial f(t, t)}{\partial T} + \alpha(t, t) \right) dt + \sigma(t, t) dW(t)$$

This is called *Musiela parametrization* and allows to switch from forward rates to short rates. This relation is not immediate, of course, because the relationship between zero coupon bond price and forward rate is deterministic, but the one between zero coupon bond price and short rate is stochastic.

3. The relationship between the zero coupon price and the forward rate is deterministic:

$$p(t, T) = e^{\int_t^T f(t, s) ds} \equiv e^{Y(t, T)}$$

From the Itô formula we then obtain the bond dynamics as

$$\frac{dp(t, T)}{p(t, T)} = dY(t, T) + \frac{1}{2} d\langle Y(\cdot, T) \rangle_t$$

where

$$dY(t, T) = d \left(- \int_t^T f(t, s) dt \right)$$

and the problem is that in the integral the t -variable occurs in two places: as the lower limit of integration and in the integrand $f(t, s)$. This is a situation that is not covered by the standard Itô formula, but it is easy to guess the answer. The t appearing as the lower limit of integration should give rise to the term $\frac{\partial}{\partial t} \left(\int_t^T f(t, s) ds \right) dt$ and since the stochastic differential is a linear operation, we should be allowed to move it inside the integral, thus providing us with the term $\int_t^T df(t, s) ds$. Therefore we arrive at

$$\begin{aligned} dY(t, T) &= d \left(- \int_t^T f(t, s) dt \right) \\ &= -\frac{\partial}{\partial t} \left(\int_t^T f(t, s) ds \right) dt - \int_t^T df(t, s) ds \\ &= f(t, t) dt - \int_t^T (\alpha(t, s) dt + \sigma(t, s) dW(t)) ds \\ &= r(t) dt - \left(\int_t^T \alpha(t, s) ds \right) dt + \left(\int_t^T \sigma(t, s) ds \right) dW(t) \end{aligned}$$

Now we have to compute the quadratic variation of Y :

$$d\langle Y(t, T) \rangle_t = \left(\int_t^T \sigma(t, s) ds \right)^2 dt$$

Then

$$\begin{aligned} \frac{dp(t, T)}{p(t, T)} &= r(t) dt + \left(- \int_t^T \alpha(t, s) ds + \frac{1}{2} \left(\int_t^T \sigma(t, s) ds \right)^2 \right) dt + \\ &\quad + \left(\int_t^T \sigma(t, s) ds \right) dW(t) \\ &= r(t) dt + \left(A(t, T) + \frac{1}{2} \|S(t, T)\|^2 \right) dt + S(t, T) dW(t) \end{aligned}$$

which is the result. \square

3.17 Linear fixed income products

Linear fixed income products are interest market's products for which the income is fixed by a certain rate of return, even if we know that the return can vary over time due to the randomness of the interest rate. The basic products that we know are the zero coupon bond and the FRA, which are the building blocks that allow to construct the first non trivial product that we consider: the *fixed coupon bond*. This is a bond which, at some intermediary points in time, will provide predetermined payments (coupons) to the holder of the bond. The formal description is as follows.

- Fix a number of dates, i.e. points in time, T_0, \dots, T_n . Here T_0 is interpreted as the emission date of the bond, whereas T_1, \dots, T_n are the coupon dates.
- At time T_i , $i = 1, \dots, n$, the owner of the bond receives the deterministic coupon c_i .
- At time T_n the owner receives the face value K .

We now go on computing the price of this bond. It is obvious that the coupon bond can be replicated by holding a portfolio of zero coupon bonds with maturities T_i , $i = 1, \dots, n$. More precisely, we will hold c_i zero coupon bonds with maturity T_i , $i = 1, \dots, n - 1$, and $K + c_n$ bonds with maturity T_n . So, the price $p(t)$ at a time $t < T_1$ of the coupon bond is given by

$$\text{price}_t^{\text{FIX}}(t) = K \cdot p(t, T_n) + \sum_{i=1}^n c_i \cdot p(t, T_i) \quad (3.422)$$

Very often the coupons are determined in terms of return, rather than in monetary (e.g. dollar) terms. The return for the i -th coupon is typically quoted as a simple rate acting on the face value K , over the period $[T_{i-1}, T_i]$. Thus, if, for example, the i -th coupon has a return equal to r_i , and the face value is K , this means that

$$c_i = r_i(T_i - T_{i-1})K.$$

For a standardized coupon bond, the time intervals will be equally spaced, i.e.

$$T_i = T_0 + i\delta,$$

and the coupon rates r_1, \dots, r_n will be equal to a common coupon rate r . The price of such a bond will, for $t \leq T_1$, be given by

$$\text{price}_t^{\text{FIX}}(t) = K \left(p(t, T_n) + r\delta \sum_{i=1}^n p(t, T_i) \right) \quad (3.423)$$

3.18 Floating rate bonds

There are various coupon bonds for which the value of the coupon is not fixed at the time the bond is issued, but rather reset for every coupon period. As an example, we will confine ourselves to discussing one of the simplest floating rate bonds, where the coupon rate r_i is set to the spot LIBOR rate $L(T_{i-1}, T_i)$. Thus

$$c_i = (T_i - T_{i-1})L(T_{i-1}, T_i)K,$$

and we note that $L(T_{i-1}, T_i)$ is determined already at time T_{i-1} , but that c_i is not delivered until time T_i . In other words, we need an information involving the forward LIBOR rate. As a consequence, we have two possibilities to price this contract:

- Before the crisis, if we set $T_i - T_{i-1} = \delta$ (coupon dates are equally spaced) and $K = 1$ we have

$$c_i = \delta L(T_{i-1}, T_i) = \delta \left(\frac{1 - p(T_{i-1}, T_i)}{\delta p(T_{i-1}, T_i)} \right) = \frac{1}{p(T_{i-1}, T_i)} - 1$$

This cash flow is delivered at time T_i but it is already measurable at time T_{i-1} with respect to the filtration $\mathcal{F}_{T_{i-1}}$. Now, if we consider

$$\text{price}_t(c_i) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} c_i \right]$$

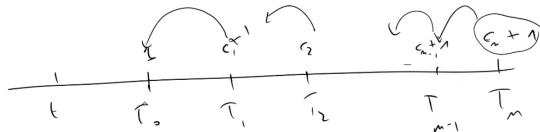
and we move to the forward measure \mathbb{Q}^T in order to eliminate the random discounting, we get

$$\begin{aligned} \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} c_i \right] &= p(t, T_i) \mathbb{E}_t^{\mathbb{Q}^{T_i}} [c_i] \\ &= p(t, T_i) \mathbb{E}_t^{\mathbb{Q}^{T_i}} \left[\frac{1}{p(T_{i-1}, T_i)} - 1 \right] \\ &= p(t, T_i) \mathbb{E}_t^{\mathbb{Q}^{T_i}} \left[\frac{p(T_i, T_i)}{p(T_{i-1}, T_i)} - 1 \right] \\ &\stackrel{(a)}{=} p(t, T_i) \left(\frac{p(t, T_i)}{p(t, T_i)} - 1 \right) \\ &= p(t, T_i) - p(t, T_i) \end{aligned}$$

where in (a) we use the martingality of the first term. Since this is the price of a single c_i , the total price will be

$$\text{price}_t^{\text{FLOAT}} = p(t, T_n) + \sum_{i=1}^n p(t, T_i) - p(t, T_i) = p(t, T_0) \quad (3.424)$$

So at the end we move one unit of the notional backward on time using as discount factor the corresponding c_i (because before the crisis we can discount and capitalize with respect to the LIBOR).



In particular, $\text{price}_{T_0}^{\text{FLOAT}} = 1$. This means that whatever happens after T_0 does not matter because the way in which we pay cash flows is exactly symmetric with respect to the way in which we discount money.

- After the crisis, keeping for simplicity $K = 1$ and $\delta = T_i - T_{i-1}$, the price is given by

$$\begin{aligned} \text{price}_t^{\text{FLOAT}} &= p(t, T_n) + \sum_{i=1}^n p(t, T_i) \mathbb{E}_t^{\mathbb{Q}^T} [\delta L(T_{i-1}, T_{i-1}, T_i)] \\ &= p(t, T_n) + \sum_{i=1}^n p(t, T_i) \delta L(t, T_{i-1}, T_i) \end{aligned} \quad (3.425)$$

where $L(t, T_{i-1}, T_i)$ is the forward LIBOR. In order to price floating coupon bonds we need a model which is able to describe the evolution of the forward

LIBOR. Before the crisis all the models were based on the description of the short interest rate were all LIBORs (3 months, 6 months and so on) evolved according to the same stochastic process. After the crisis the behavior of LIBORs for different tenors is different, due to the spread between rates, leading to multiple yield curves.

3.19 Yield and duration

Consider a zero coupon bond with market price $p(t, T)$. We now look for the bond's "internal rate of interest", i.e. the constant short rate of interest which will give the same value to this bond as the value given by the market. Denoting this value of the short rate by y , we thus want to solve the equation

$$p(t, T) = e^{-y(T-t)} \cdot 1 \quad (3.426)$$

We are thus led to the following definition.

Definition 3.19.1 (Continuously compounded zero coupon yield). *The continuously compounded zero coupon yield, $y(t, T)$, is given by*

$$y(t, T) = -\frac{\log p(t, T)}{T - t} \quad (3.427)$$

and, for a fixed t , the function $T \rightarrow y(t, T)$ is called the (zero coupon) yield curve.

The standard behavior of the yield curve is increasing, but in different historical periods we can see also decreasing or humped yield curves.

Now let us consider a fixed coupon bond where, for simplicity of notation, we include the face value in the coupon c_n . We denote its market value at t by $p(t)$. In the same spirit as above we now look for its internal rate of interest, i.e. the constant value of the short rate, which will give the market value of the coupon bond.

Definition 3.19.2 (Yield to maturity). *The yield to maturity, $y(t, T)$, of a fixed coupon bond at time t , with market price p , and payments c_i at T_i for $i = 1, \dots, n$, is defined as the value of y which solves the equation*

$$p(t) = \sum_{i=1}^n c_i e^{-y(T_i-t)}. \quad (3.428)$$

In other words, the yield to maturity corresponds to the barycenter of the interest rate such that the cash flow corresponding to the coupon is delivered according to a certain interest rate. An important concept in bond portfolio management is the duration. Without loss of generality we may assume that $t = 0$.

Definition 3.19.3 (Duration). *For the fixed coupon bond above, with price p at $t = 0$, and yield to maturity y , the duration, D , is defined as*

$$D = \frac{\sum_{i=1}^n T_i c_i e^{-y T_i}}{p}. \quad (3.429)$$

The duration is thus a weighted average of the coupon dates of the bond, where the discounted values of the coupon payments are used as weights, and it will in a sense provide you with the "mean time to coupon payment". It also acts as a measure of the sensitivity of the bond price w.r.t. changes in the yield:

$$\frac{dp}{dy} = -Dp$$

Thus we see that duration is essentially for bonds (w.r.t. yield) what delta is for derivatives (w.r.t. the underlying price)

$$D = -\% \text{ sensitivity to a parallel shift of the yield curve.} \quad (3.430)$$

Notice that in the trivial case of a zero coupon bond the duration coincides with the maturity.

3.20 Interest rate swaps

The interest rate swap is a scheme where we exchange a payment stream at a fixed rate of interest, known as the *swap rate*, for a payment stream at a floating rate (typically a LIBOR rate). Basically, a interest rate swap is a portfolio of FRAs.

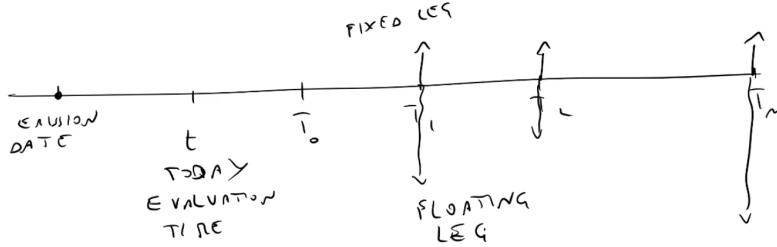


Figure 3.24: Interest rate swaps

There are many versions of interest rate swaps, and we will study the *forward swap settled in arrears*, which is defined as follows. We denote the principal by K , and the swap rate by R . By assumption we have a number of equally spaced dates T_0, \dots, T_n , and payment occurs at the dates T_1, \dots, T_n (not at T_0). If we swap a fixed rate for a floating rate (in this case the LIBOR spot rate), then, at time T_i , we will receive the amount

$$K\delta L(T_{i-1}, T_i)$$

At T_i we will pay the amount

$$K\delta R$$

so the net cash flow at T_i is thus given by

$$K\delta[L(T_{i-1}, T_i) - R]. \quad (3.431)$$

Basically, this is a floating bond minus a fixed bond. So, the corresponding price is given by

$$\text{price}_t^{\text{SWAP}} = \text{price}_t^{\text{FLOAT}} - \text{price}_t^{\text{FIXED}}.$$

Before the crisis this price is given by:

$$\text{price}_t^{\text{SWAP}} = Kp(t, T_0) - K \left(p(t, T_n) + \sum_i^n R\delta p(t, T_i) \right) \quad (3.432)$$

The swap rate R , which makes the whole contract fair ($\text{price}_0^{\text{SWAP}} = 0$), is different from the strike prices of the FRAs, which are fair only for the single component of the contract, and is given by

$$R = \frac{p(t, T_0) - p(0, T_n)}{\delta \sum_i^n p(0, T_i)}. \quad (3.433)$$

After the crisis, we have to adapt the formula by considering the aftermath in terms of floating and fixed bonds:

$$\text{price}_t^{\text{SWAP}} = K \sum_{i=1}^n p(t, T_i) \delta(L(t, T_{i-1}, T_i) - R) \quad (3.434)$$

and the swap rate is given by a convex combination of forward LIBOR rates:

$$R = \frac{\sum_{i=1}^n p(t, T_i)}{\delta \sum_{j=1}^n p(0, T_j)} L(t, T_{i-1}, T_i). \quad (3.435)$$

3.21 Non linear contracts: caps and floors

An *interest rate cap* is a financial insurance contract which protects us from having to pay more than a prespecified rate, the *cap rate*, even though we have a loan at a floating rate of interest. Technically speaking, a cap is the sum of a number of basic contracts, known as *caplets*, which are defined as follows:

- The interval $[0, T]$ is subdivided by the equidistant points $0 = T_0, T_1, \dots, T_n = T$. We use the notation δ for the length of an elementary interval, i.e. $\delta = T_i - T_{i-1}$.
- The cap is working on some principal amount of money, denoted by K , and the cap rate is denoted by R .
- The floating rate of interest underlying the cap is not the short rate r , but rather some market rate, and we will assume that over the interval $[T_{i-1}, T_i]$ it is the LIBOR spot rate $L(T_{i-1}, T_i)$.
- Caplet i is now defined as the contingent claim that has payoff at T_i :

$$K \delta(L(T_{i-1}, T_i) - R)^+ \quad (3.436)$$

There are also *floor contracts* which guarantee that the interest paid on a floating rate loan will never be below some predetermined floor rate. We now turn to the problem of pricing the caplet, and without loss of generality we may assume that $K = 1$.

Before the crisis

Before the crisis it is possible to write the LIBOR in terms of discount factor:

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

Substituting in eq. (3.436) (with $K = 1$) we get

$$\delta(L(T_{i-1}, T_i) - R)^+ = \delta \left(\frac{1 - p(T_{i-1}, T_i)}{\delta p(T_{i-1}, T_i)} - R \right)^+$$

Now the idea is to collect the factor $\frac{1 - \delta R}{p(T_{i-1}, T_i)}$ so that the payoff will correspond to a certain number of put options on the zero coupon bond:

$$\frac{1 - \delta R}{p(T_{i-1}, T_i)} \left(\frac{1}{1 + \delta R} - p(T_{i-1}, T_i) \right)^+ \quad (3.437)$$

This payoff is paid at time T_i but it involves quantities that are measurable with respect to the filtration at time T_{i-1} , $\mathcal{F}_{T_{i-1}}$. In other words, at time T_{i-1} this quantity is deterministic, so in order to find the price of the caplet we only have to discount the cash flow:

$$\frac{p(T_{i-1}, T_i)}{p(T_{i-1}, T_i)} \left(\frac{1}{1 + \delta R} - p(T_{i-1}, T_i) \right)^+ = (1 + \delta R)(\text{put on } p(T_{i-1}, T_i)) \quad (3.438)$$

So, in the end, the cap is a portfolio of put options on zero coupon bonds.

After the crisis

After the crisis we cannot consider the discount factor, so there will be a conditional expected value of the future realizations of the forward LIBOR, which in general are different for different tenors. Then, we would need a model to describe the LIBOR.

3.22 Short interest rates

We now consider different possible specifications for the short interest rate. We already introduced the Vašíček model, which describes the evolution of the short rate using a Brownian motion plus a drift and for which the distribution of the interest rate is gaussian. We also were able to find the price of the corresponding zero coupon bond (at least the diffusive part). Now we would like to consider other short interest rate models.

First, we need to extend the Feynman-Kac methodology to stochastic interest rates. If we define the infinitesimal generator

$$\mathcal{A} = \frac{\partial}{\partial x} K + \frac{1}{2} \frac{\partial^2}{\partial x^2} H^2$$

we obtain the following differential equation for the evolution of the price with stochastic interest rate $r(t, x)$:

$$\begin{cases} \frac{\partial F}{\partial t} + \mathcal{A} + r(t, x)F = 0 \\ F(T, x) = \Phi(x) \end{cases} \quad (3.439)$$

The solution is given by [homework](#)

$$F(t, x) = \mathbb{E}_{t,x} \left[\Phi(x_T) e^{- \int_t^T r(s, x) ds} \right] \quad (3.440)$$

where

$$dX = K dt + H dW \quad (3.441)$$

Let's consider, for example, the case $r(t, x) = x$. The dynamics under the risk neutral probability measure \mathbb{Q} is

$$dr(t) = \mu(t, r(t)) dt + \sigma(t, r(t)) dW^{\mathbb{Q}}(t) \quad (3.442)$$

In principle, we should start from the probability measure \mathbb{P} and then try to understand which is the change of measure that leads to \mathbb{Q} . However, this is a quite complicated problem because it is not clear what to impose to be a martingale, in fact the interest rate is not a traded asset. Practically, the fact that we do not know \mathbb{P} does not matter because we want to work under \mathbb{Q} , which is the probability measure that generate the prices in the market. From now on, we will consider every Brownian motion under the measure \mathbb{Q} .

Vašíček model

The dynamics of the short interest rate follows the equation

$$dr(t) = (b - ar(t)) dt + \sigma dW(t), \quad a > 0 \quad (3.443)$$

This dynamics is *affine*, that is the drift and the diffusion coefficients are linear combinations of the state variable (plus a constant). The short rate $r(t)$ is gaussian distributed and introducing an auxiliary process we are able to find it.

Cox-Ingersoll-Ross model

The CIR model was introduced in order to prevent the negativity of $r(t)$. It is defined by the equation

$$dr(t) = a(b - r(t)) dt + \sigma \sqrt{r(t)} dW(t) \quad (3.444)$$

and it is an affine model (the first term and the square root of the second term are affine). A process that evolves according this equation is called *squared Bessel process* and it is possible to prove that it has a strong unique solution and that if $ab > \sigma^2/2$ then $\mathbb{P}(r(t) > 0) = 1$. It also holds that CIR \sim Vašíček $\sim \chi^2$.

Dothan model

The dynamics is given by

$$dr(t) = ar(t) dt + \sigma r(t) dW(t) \quad (3.445)$$

The dynamics is not affine because the diffusive term is not the square of an affine term. The rate $r(t)$ is log-normal distributed. There are some problem in pricing the zero coupon bond because if we consider the expected value of the exponential of the integral of $r(t)$ (i.e. the zero coupon price) it can explode.

Black-Derman-Toy model

The dynamics is given by

$$dr(t) = \theta(t)r(t) dt + \sigma(t)r(t) dW(t) \quad (3.446)$$

where $\theta(t)$ and $\sigma(t)$ are deterministic functions that allow more flexibility to describe the process. This model is not affine.

Ho-Lee model

The dynamics is given by

$$dr(t) = \theta(t) dt + \sigma dW(t) \quad (3.447)$$

where $\theta(t)$ is a deterministic function and σ is a constant. This model is the simplest specification of a more general framework – the Heath–Jarrow–Morton framework – which describes the dynamics of forward interest rates. The rate $r(t)$ is gaussian. There can be some stationarity issues.

Hull-White model

The dynamics is given by

$$dr(t) = (\theta(t) - a(t)r(t)) dt + \sigma(t) dW(t) \quad (3.448)$$

The function $\theta(t)$ allows a fine calibration of the model to perfectly fit the initial yield curve.

Remark 3.22.1. We will prove that $r(t)$ is affine if and only if the moment generating function is such that $\mathbb{E}[e^{r(t)}] = e^{A+Bt}$, where A and B are two deterministic functions. This has as a consequence the fact that the price of the zero coupon bond can be written in terms of A and B :

$$\mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} \right] = e^{A(t,T)+B(t,T)}$$

The problem is: are we able to find the functions A and B for all the models above? The answer is positive, but we need to use the extension of the Feynman-Kac methodology.

3.23 Affine term structures

When we considered the short interest rate we said that the drift and the diffusion were affine functions of the short interest rate itself. Now, we have to deduce this fact from the properties of the pricing function of the zero coupon bond. So, the starting point is the price of the zero coupon bond:

$$\begin{aligned} p(t, T) &= \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} \right] \\ &= F(t, r(t), T) \end{aligned}$$

Lesson 27 of
13/05/2020

We introduced the Feynman-Kac methodology in order to solve a PDE in the case in which we do not have a candidate solution. In fact, thanks to the Feynman-Kac methodology we can transform the problem of finding the solution of a deterministic PDE into the problem of computing an expected value.

Now the situation is a bit different. We can write the price of the zero coupon bond as an expected value but in general we cannot compute it because we don't know the probability distribution of $r(t)$. So, we have to find a candidate solution of the PDE.

Suppose that

$$p(t, T) = e^{A(t,T)-B(t,T)r(t)} \quad (3.449)$$

where $A(t, T)$ and $B(t, T)$ are two deterministic functions and the conventional minus sign is coherent with the fact that if there is a rise up of the interest rate the price of the zero coupon bond will decrease. Then, the model associated to the short rate $r(t)$ is said to possess an *affine term structure* (ATS). Then, the question is: if the model has an ATS, which is the dynamics of the process? Let's consider the case

$$r(t) = x$$

and assume a generic SDE for r :

$$dr(t) = \mu(t, r(t)) dt + \sigma(t, r(t)) dW(t) \quad (3.450)$$

where measurability and integrability are enough to find a strong solution of this equation. Imposing the candidate solution

$$p(t, T) = e^{A(t, T) - B(t, T)r(t)}$$

to the Feynman-Kac equation

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

we get

$$\begin{aligned} \frac{\partial F}{\partial t} &= F \left(\frac{\partial A}{\partial t} - \frac{\partial B}{\partial t}r \right) \\ \frac{\partial^2 F}{\partial r^2} &= F \cdot B^2 \end{aligned}$$

and so

$$\begin{cases} F \left(\frac{\partial A}{\partial t} - \frac{\partial B}{\partial t}r \right) - B\mu F + \frac{1}{2}\sigma^2 B^2 F - rF = 0 \\ e^{A(T, T) - B(T, T)r(T)} = 1. \end{cases}$$

From the second equation we have that

$$A(T, T) = 0, \quad B(T, T) = 0 \quad (3.451)$$

Simplifying the first equation we get

$$\left(\frac{\partial A}{\partial t} - \frac{\partial B}{\partial t}r \right) - B\mu + \frac{1}{2}\sigma^2 B^2 - r = 0$$

Notice that there is a linear dependence on r and then a generic dependence on μ and σ^2 . If μ and σ^2 are both affine (i.e. linear plus a constant) functions of r , with possibly time dependent coefficients, then the differential equation becomes separable. Assume thus that μ and σ assume the following form

$$\mu(t, r) = \alpha(t)r + \beta(t), \quad \sigma(t, r) = \sqrt{\gamma(t)r + \delta(t)}.$$

Substituting we end up with

$$\dot{A} - \beta B + \frac{1}{2}\delta B^2 - \left(1 + \dot{B} + \alpha B - \frac{1}{2}\gamma B^2 \right) r = 0$$

This equation holds for all t, T and r , so let us consider it for a fixed choice of T and t . Since the equation holds for all values of r the coefficient of r must be equal to zero. Thus we have the equation

$$\begin{cases} \dot{B} + \alpha B - \frac{1}{2}\gamma B^2 = -1 \\ B(T, T) = 0 \end{cases} \quad (3.452)$$

which is called *Riccati quadratic ODE* and if B is a symmetric squared matrix it has a closed form solution. Since the r -term is zero we see that the other term must also vanish, giving us the equation

$$\begin{cases} \dot{A} = \beta B - \frac{1}{2}\delta B^2 \\ A(T, T) = 0 \end{cases} \quad (3.453)$$

Once we know the solution $B(t, T)$, solving the ODE for A is trivial.

Example 3.23.1: Vašíček model

In the Vašíček model

$$dr = (b - ar)dt + \sigma dW$$

so we have $\beta = b$, $\alpha = -a$, $\gamma = 0$ and $\delta = \sigma^2$. Eqs. (3.452) and (3.453) become

$$\begin{cases} \dot{B} - aB = -1 \\ B(T, T) = 0 \end{cases} \quad (3.454)$$

$$\begin{cases} \dot{A} = \beta B - \frac{1}{2}\delta B^2 \\ A(T, T) = 0 \end{cases} \quad (3.455)$$

Equation (3.454) is, for each fixed T , a simple linear ODE in the t -variable. It has solution

$$B(t, T) = \frac{1}{a} \left(1 - e^{-a(T-t)} \right)$$

which is exactly what we obtained before by computing the expected value. Integrating (3.455) we obtain

$$A(t, T) = \frac{\sigma^2}{2} \int_t^T B^2(s, T) ds - b \int_t^T B(s, t) ds$$

and, substituting the expression for B above, we obtain

$$A(t, T) = \frac{(B(t, T) - T + t)(ab - \frac{1}{2}\sigma^2)}{a^2} - \frac{\sigma^2 B^2(t, T)}{4a}.$$

The bond prices are given by

$$p(t, T) = e^{A(t, T) - B(t, T)r(t)}.$$

Example 3.23.2: Ho–Lee model

For the Ho–Lee model

$$dr(t) = \Theta(t) dt + \sigma dW(t)$$

so $\alpha = \gamma = 0$. The ATS equations become

$$\begin{cases} \dot{B} - aB = -1 \\ B(T, T) = 0 \end{cases} \quad \begin{cases} \dot{A} = \beta B - \frac{1}{2}\delta B^2 \\ A(T, T) = 0 \end{cases}$$

which can be solved as

$$\begin{aligned} B(t, T) &= T - t \\ A(t, T) &= \int_t^T \Theta(s)(s - T) ds + \frac{\sigma^2(T - t)^3}{6} \end{aligned}$$

Example 3.23.3: CIR model

In the CIR model

$$dr(t) = a(b - r(t))dt + \sigma\sqrt{r(t)}dW(t)$$

Notice that even if $\sqrt{r(t)}$ is not Lipschitz, the existence of a strong solution is guaranteed by the fact that it is Hölder continuous^a, so the Yamada condition^b can be applied. Since $\gamma = 1$, we get a true Riccati equation:

$$\begin{cases} \dot{B} - aB - \frac{1}{2}\sigma^2 B^2 = -1 \\ B(T, T) = 0 \end{cases} \quad \begin{cases} \dot{A} = abB \\ A(T, T) = 0 \end{cases}$$

where $\sigma = \sqrt{\gamma r + \delta}$. The solution of the Riccati equation is

$$B(t, T) = \frac{2(e^{\epsilon(T-t)} - 1)}{(\epsilon - a)(e^{\epsilon(T-t)} - 1) + 2\epsilon}$$

with

$$\epsilon = \sqrt{a^2 + 2\sigma^2}$$

^a A real or complex-valued function f on d -dimensional Euclidean space is Hölder continuous when there are non-negative real constants C and $\alpha > 0$ such that

$$|f(x) - f(y)| \leq C\|x - y\|^\alpha$$

for all x and y in the domain of f .

^b See <https://fam.tuwien.ac.at/~schmock/notes/Yamada-Watanabe.pdf>.

3.24 Stochastic volatility

We already know that the B&S model is not exact, in the sense that it is not able to reproduce the volatility smile. In order to go beyond, the idea is to assume the volatility to be random, driven by a stochastic process which in principle is (partially) correlated with the asset. Keeping the interest rate constant (we are not interested in its fluctuations in the short term) and considering a stochastic volatility we can write the B&S equation as

$$\frac{dS}{S} = r dt + \sigma(t) dW^Q(t) \quad (3.456)$$

It is possible to use the square root of the instantaneous variance instead of the volatility:

$$\frac{dS}{S} = r dt + \sqrt{V(t)} dW^Q(t) \quad (3.457)$$

One of the pioneering models for stochastic volatility is due to Heston (1993), who borrowed the CIR process for the interest rate and applied it to the volatility:

$$dV(t) = K(V(\infty) - V(t))dt + \eta\sqrt{V(t)}(\rho dW^Q(t) + \sqrt{1 - \rho^2} dW^\perp(t)) \quad (3.458)$$

where $W^Q \perp\!\!\!\perp W^\perp$. The parameter η is called *volatility of volatility parameter* and the parameter ρ is called *correlation parameter*. One interesting feature of this model is that the quadratic covariation between the asset returns and the instantaneous variance is

$$\left\langle \frac{dS}{S}, V \right\rangle = V\eta\rho$$

which is positive or negative according to ρ and can range from -1 to $+1$. Typically we have $\eta > 0$, $V > 0$ and $\rho < 0$. The fact that $\rho < 0$ is the explanation of the leverage effect of the volatility (when the asset value is growing the volatility is low and when the market crash the volatility is high). $\rho < 0$ is also responsible for the

negative slope of the smile effect, that is the skew.

The main advantage of this model is that it is analytically tractable, in the sense that we are able to do pricing even if we loose the log-normal property of the asset price. In other words, we are not able to compute prices as conditional expected value but through an analytical methodology which involves the Feynman-Kac formula and the Fourier transform. In order to introduce this methodology, we need some preliminary results on squared Bessel processes.

3.24.1 Review of Bessel squared processes

Let $V(t)$ be a time changed Bessel squared process (BESQ).

Proposition 3.24.1 (Feller condition). *If $2KV(\infty) \geq \eta^2$, then*

$$\text{Prob}(V(t) > 0 \forall t) = 1.$$

This condition is very important because it guarantees the strict positivity of the volatility process. Recall that a BESQ process is characterized by an infinitesimal generator

$$\mathcal{A} = 2x \frac{d^2}{dx^2} + \delta \frac{d}{dx} \quad (3.459)$$

We define the BESQ process parametrized by δ and starting from the value x such that it satisfies

$$\begin{cases} d\rho(t) = \delta dt + 2\sqrt{\rho(t)} dW(t) \\ \rho(0) = x \geq 0 \end{cases} \quad (3.460)$$

(where ρ has nothing to do with the correlation parameter). The parameter $\delta \geq 0$ is related to the dimension of the BESQ. Another way to parametrize the process is to define the *index* $\nu = \frac{\delta}{2} - 1$:

$$\text{BESQ}_x^\delta = \text{BESQ}_x^{(\nu)}.$$

Remark 3.24.1. If $\delta = n \in \mathbb{N}$ and (W_1, \dots, W_n) are n independent Brownian motions, then

$$\text{BESQ}_x^\delta = \|(W_1, \dots, W_n)\|$$

i.e. the BESQ is the radial distance from the origin. In fact, if we set $R(t) = \|W(t)\|$, then

$$R^2(t) = \sum_{i=1}^n W_i^2$$

and then

$$\begin{aligned} dR^2 &= n dt + \sum_{i=1}^n 2W_i dW_i \\ &= n dt + 2R(t) d\bar{W}(t) \end{aligned}$$

where $\bar{W}(t)$ is the scalar Brownian motion defined as

$$\bar{W}(t) = \frac{1}{R(t)} \sum_{i=1}^n W_i dW_i.$$

Now, let's consider the classification of the boundary:

- if $\delta \geq 2$ ($V \geq 0$) then BESQ_x^δ never reaches the zero (Feller condition);

Lesson 28 of
14/05/2020

- if $0 \leq \delta < 2$ ($-1 \leq V < 0$) then BESQ_x^δ may reach zero, so we have to specify the boundary conditions. For the CIR process

$$\begin{cases} dr(t) = k(\theta - r(t))dt + \sigma\sqrt{r(t)}dW(t) \\ r(0) = x > 0 \end{cases}$$

it is possible to prove that

$$r(t) = e^{-kt}\rho \left(\frac{\sigma^2}{4k}(e^{kt} - 1) \right)$$

where $\rho(t) \sim \text{BESQ}_x^\delta$ with $\delta = \frac{4k\theta}{\sigma^2}$. Moreover, the Feller condition is given by

$$r(t) > 0 \quad \forall t \quad \Leftrightarrow \quad 2k\theta \geq \sigma^2.$$

3.24.2 The pricing problem with stochastic volatility

Now we consider the problem of pricing a contract in this framework. The problem is that we loose the information about the distribution of the asset price, so even if the risk neutral methodology still holds true we cannot compute conditional expected values.

Of course, the underlying evolves according to the B&S equation with stochastic volatility:

$$\frac{dS}{S} = r dt + \sqrt{V(t)} dW(t) \tag{3.461}$$

Just for convenience, we shift from prices to returns by considering that

$$S = e^x.$$

Notice that $S(t)$ has an unknown distribution, so when if we want to the price of the payoff as

$$\begin{aligned} \text{price}_t(\text{payoff}_T) &= e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}}[\text{payoff}_T] \\ &= e^{-r(T-t)} \int_{\mathbb{R}} (e^{x_T} - K)^+ f(x_T | x_t) dx_T \end{aligned}$$

we cannot do it, because the conditional density function $f(x_T | x_t)$ is unknown. Since we know that there is a one to one correspondence between the density function and its Fourier transform, we can compute the density as the anti-transform of its Fourier transform:

$$\hat{f}(z) = \int_{\mathbb{R}} e^{itz} f(x | x_t) dx$$

In general z is real, but we can extend z to a complex domain \mathbb{C} and get the so called *generalized Fourier transform*, also known as the *Laplace transform*. From the FT we can recover the density as

$$f(x_T | x_t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-izx} \hat{f}(z) dt$$

We have that

$$\hat{f}(z) = \mathbb{E}_t^{\mathbb{Q}}[e^{izx_T}]$$

so we can write the price of the payoff as

$$\text{price}_t(\text{payoff}_T) = e^{-r(T-t)} \int_{\mathbb{R}} (e^{x_T} - K)^+ \left(\frac{1}{2\pi} \int_{\mathbb{R}} e^{-izx_T} \hat{f}(z) dz \right) dx_T$$



Figure 3.25: Covered call payoff.

$$\begin{aligned} &= \frac{e^{-r(T-t)}}{2\pi} \int_{\mathbb{R}} \hat{f}(z) dz \int_{\mathbb{R}} (e^{x_T} - K)^+ e^{-izx_T} dx_T \\ &= \frac{e^{-r(T-t)}}{2\pi} \int_{\mathbb{R}} \hat{f}(z) \widehat{\text{payoff}}(-z) dz \end{aligned}$$

where $\widehat{\text{payoff}}$ is the FT of the payoff. This integral is computed numerically. Notice that $\hat{f}(z) = \mathbb{E}_t^Q[e^{izx_T}]$ is product independent and model dependent, so once we found the characteristic function of the log-asset price in the Heston model we can use it to price whatever option (we just have to change the payoff). On the contrary, the FT of the payoff is product dependent and model independent.

So, the problem of pricing within the stochastic volatility framework consists in computing the integral of the product of two quantities that can be separately computed: the characteristic function of the model (which is computed only one time) and the Fourier transform of the payoff. To speed up the computation of the FT, in 1999 Carr and Madan introduced the *Fast Fourier Transform* in finance.

Let's consider some examples of payoff FT.

Covered call option

The payoff of a covered call is

$$h(x_T) = \min\{e^{x_T}, K\} \quad (3.462)$$

so its FT is

$$\begin{aligned} \hat{h}(z) &= \int_{\mathbb{R}} e^{ixz} \min\{e^x, K\} dx \\ &= \int_{-\infty}^{\ln K} e^{izx} e^x dx + \int_{\ln K}^{+\infty} e^{izx} K dx \\ &= \frac{1}{iz+1} e^{(iz+1)x} \Big|_{-\infty}^{\ln K} + \frac{1}{zi} e^{izx} \Big|_{\ln K}^{+\infty} \end{aligned} \quad (3.463)$$

In order to get convergence in the first integral, it must be

$$\operatorname{Re}\{iz+1\} > 0$$

so, if $z = \operatorname{Re}\{z\} + i \operatorname{Im}\{z\}$ then

$$\operatorname{Re}\{i(\operatorname{Re}\{z\} + i \operatorname{Im}\{z\}) + 1\} > 0 \Rightarrow \operatorname{Im}\{z\} < 1.$$

In order to get convergence in the second integral, it must be

$$\operatorname{Re}\{iz\} < 0 \Rightarrow \operatorname{Im}\{z\} > 0.$$

So, if

$$0 < \operatorname{Im}\{z\} < 1 \quad (3.464)$$

then we end up with

$$\hat{h}(z) = \frac{K^{iz+1}}{z^2 - iz}. \quad (3.465)$$

Then the integrability domain of convergence for the (generalized) FT of the covered

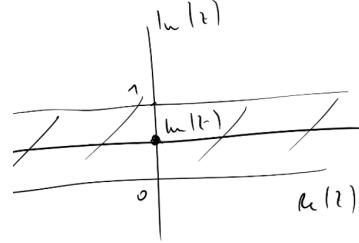


Figure 3.26: Strip of integrability.

call will be

$$\{z \in \mathbb{C} : -1 < \text{Im}\{z\} < 0\} \quad (3.466)$$

(the domain if the inverse of (3.464) because the argument of the payoff is $-z$).

Call option

The payoff of the call option is

$$g(x_T) = (e^{x_T} - K)^+ \quad (3.467)$$

and its FT is

$$\begin{aligned} \hat{g}(z) &= \int_{\mathbb{R}} e^{izx} (e^x - K)^+ dx \\ &= \frac{1}{iz + 1} e^{(iz+1)x} \Big|_{\ln K}^{+\infty} \end{aligned} \quad (3.468)$$

In order to get convergence it must be

$$\text{Re}\{iz + 1\} < 0 \quad \Rightarrow \quad \text{Im}\{z\} > 1 \quad (3.469)$$

and provided that this condition is satisfied, the FT is given by

$$\hat{g}(z) = -\frac{K^{iz+1}}{z^2 - iz}. \quad (3.470)$$

If we write z as $z = a + ib$ then we have

$$\mathbb{E}[e^{-izx_T}] = \mathbb{E}[e^{bx_T}] = \mathbb{E}[S^b]$$

where S is the underlying and $b \geq 1$. We know that, in principle, once discounted, S will be a martingale, i.e. integrable. However, this is not guaranteed for its moments. So we need to consider a strip of integrability such that these moments do not explode (for the B&S model all the moments converge).

Put option

The payoff of the put option is

$$g(x_T) = (K - e^{x_T})^+ \quad (3.471)$$

and its FT is

$$\begin{aligned} \hat{g}(z) &= \int_{\mathbb{R}} e^{izx} (K - e^x)^+ dx \\ &= \int_{-\infty}^{\ln K} e^{izx} (K - e^x) dx \\ &= \frac{K}{iz} e^{izx} \Big|_{-\infty}^{\ln K} - \frac{1}{iz + 1} e^{(iz+1)x} \Big|_{-\infty}^{\ln K} \end{aligned} \quad (3.472)$$

In order to get convergence it must be

$$\begin{aligned} \operatorname{Re}\{iz\} > 0 &\Rightarrow \operatorname{Im}\{z\} < 0 \\ \operatorname{Re}\{iz + 1\} > 0 &\Rightarrow \operatorname{Im}\{z\} < 1 \end{aligned}$$

so

$$\operatorname{Im}\{z\} < 0 \quad (3.473)$$

Provided that this condition is satisfied, the FT is given by

$$\hat{g}(z) = \frac{K^{1+iz}}{iz} - \frac{K^{iz+1}}{iz + 1} = -\frac{K^{iz+1}}{z^2 - iz}. \quad (3.474)$$

3.24.3 Characteristic functions

What can we say about the characteristic function of the model? If

$$x_T = \ln S_T$$

then the characteristic function is

$$\varphi(z) = \mathbb{E}_{t,x}^{\mathbb{Q}}[e^{izx_T}]$$

We know that in principle this guy can be also the solution of the corresponding Feynman-Kac PDE, so our guess is a function depending on the time to maturity $T - t$, on z , on the initial value x of the process and also on a state variable v related to the volatility Brownian motion:

$$\varphi(z) = G(T - t, z, x, v).$$

As usual, the underlying satisfies the SDE

Lesson 29 of
15/05/2020

$$\frac{dS}{S} = r dt + \sqrt{v(t)} dW(t)$$

and the volatility follows the CIR process dynamics

$$dv = K(\theta - v(t)) dt + \eta \sqrt{v(t)} (\rho dW(t) + \sqrt{1 - \rho^2} dW(t)^\perp) \quad (3.475)$$

which is an affine process. Let's shift to the dynamics of the log-price (the returns) $x(t) = \ln S(t)$. Using Itô, we get

$$dx(t) = \left(r - \frac{1}{2}v(t) \right) dt + \sqrt{v(t)} dW(t) \quad (3.476)$$

Since v is embedded in the dynamics of x , we have to consider it when we compute the expected value. Moreover, since both x and v are affine, we guess that the expected value has a joint affine structure:

$$\mathbb{E}_{t,x,v}^{\mathbb{Q}}[e^{izx_T}] = e^{A(t,T) + B(t,T)v(t) + C(t,T)x(t)} \equiv G(\tau, z, x, v), \quad \tau = T - t \quad (3.477)$$

In order to transform this conditional expected value into a PDE we use the Feynman-Kac formula:

$$\begin{aligned} 0 = & -\frac{\partial G}{\partial \tau} + \left(r - \frac{1}{2}v \right) \frac{\partial G}{\partial x} + K(\theta - v) \frac{\partial G}{\partial v} + \frac{1}{2} \underbrace{v}_{\langle x \rangle} \frac{\partial^2 G}{\partial x^2} + \\ & + \frac{1}{2} \underbrace{\eta^2 v}_{\langle v \rangle} \frac{\partial^2 G}{\partial v^2} + \underbrace{\eta \rho v}_{\langle x, v \rangle} \frac{\partial^2 G}{\partial x \partial v} \end{aligned} \quad (3.478)$$

with initial condition

$$G(0, z, x, v) = e^{izx}$$

Since the couple (x, v) is affine, we can guess that the solution can be written as a function of the time difference $\tau = T - t$:

$$G(\tau, z, x, v) = e^{A(\tau) + B(\tau)v(t) + C(\tau)x(t)}$$

so we save one parameter. This is our candidate solution, so now we have to impose the structure we want:

$$-\frac{\partial G}{\partial \tau} = -G(\dot{A} + \dot{B}v + \dot{C}x)$$

$$\frac{\partial G}{\partial x} = CG, \quad \frac{\partial^2 G}{\partial x^2} = C^2G, \quad \frac{\partial G}{\partial v} = BG$$

$$\frac{\partial^2 G}{\partial v^2} = B^2G, \quad \frac{\partial^2 G}{\partial x \partial v} = BCG.$$

Substituting in (3.478) we get:

$$\begin{aligned} 0 = & -G(\dot{A} + \dot{B}v + \dot{C}x) + \left(r - \frac{1}{2}v \right) CG + K(\theta - v)BG + \frac{1}{2}vC^2G + \\ & + \frac{1}{2}\eta^2vB^2G + \eta\rho vBCG \quad \forall(x, v) \end{aligned} \quad (3.479)$$

Separately identifying the coefficients of x and v , we get the following ODEs. Coefficients of x :

$$\begin{cases} -\dot{C} = 0 \\ C(0) = iz \end{cases} \Rightarrow C(\tau) = iz$$

Coefficients of v :

$$\begin{cases} -\dot{B} - \frac{1}{2}iz - KB + \frac{1}{2}(iz^2) + \frac{1}{2}\eta^2B^2 + \eta\rho B(iz) = 0 \\ B(0) = 0 \end{cases}$$

This is a true Riccati ODE with constant coefficients. Then, if we consider the constant terms we get

$$\begin{cases} -\dot{A} + rC + K\theta B = -\dot{A} + riz + K\theta B = 0 \\ A(0) = 0 \end{cases}$$

Now we have to solve the Riccati equation. In order to do that we can double the dimension and linearize or write solution in terms of a derivative of a function divided by the function. This last approach comes from a property of the Riccati equation, that lives in a space where variables can be parametrized by ratios. Let's introduce the change of variable

$$B(\tau) = -\frac{\dot{E}(\tau)}{\frac{\eta^2}{2}E(\tau)} \quad (3.480)$$

which transforms the quadratic first order ODE into a linear second order ODE:

$$\ddot{E} + (K - \rho\eta iz)\dot{E} - \frac{\eta^2}{4}(iz + z^2)E = 0$$

and let's look for solutions of the form $e^{\lambda\tau}$. Substituting we get

$$\lambda^2 + (K - \rho\eta iz)\lambda - \frac{\eta^2}{4}(iz - z^2) = 0$$

Solving this quadratic equation we get

$$\lambda_{1,2} = \frac{-(K - \rho\eta iz)^2 \pm \sqrt{\Delta}}{2}$$

where

$$\Delta = (K - \rho\eta iz)^2 + \eta^2(iz + z^2).$$

Now, if we introduce the two auxiliary functions

$$\psi^\pm = -(K - \rho\eta iz) \pm \sqrt{\Delta}$$

such that

$$\psi^+ - \psi^- = 2\sqrt{\Delta}, \quad \psi^+ - \psi^- = -\eta^2(iz + z^2),$$

the general solution can be written as

$$E(\tau) = \alpha_1 e^{\frac{\psi^+}{2}\tau} + \alpha_2 e^{\frac{\psi^-}{2}\tau} \quad (3.481)$$

where α_1, α_2 are two constants to determine according to the initial conditions

$$E(0) = \alpha_1 + \alpha_2$$

$$\dot{E}(t) = \frac{1}{2}\psi^+\alpha_1 e^{\frac{\psi^+}{2}\tau} + \frac{1}{2}\psi^-\alpha_2 e^{\frac{\psi^-}{2}\tau} \quad \Rightarrow \quad \dot{E}(0) = \frac{1}{2}\psi^+\alpha_1 + \frac{1}{2}\psi^-\alpha_2$$

This leads to

$$\alpha_1 = \frac{\psi^- E(0)}{2\sqrt{\Delta}}, \quad \alpha_2 = \frac{\psi^+ E(0)}{2\sqrt{\Delta}}$$

Finally, we can get an expression for B :

$$\begin{aligned} B(\tau) &= -\frac{\psi^+\psi^- \left(e^{\frac{1}{2}\psi^+\tau} - e^{\frac{1}{2}\psi^-\tau} \right)}{\eta^2 \left(\psi^- e^{\frac{1}{2}\psi^+\tau} - \psi^+ e^{\frac{1}{2}\psi^-\tau} \right)} \\ &= -(iz + z^2) \frac{e^{\frac{1}{2}\psi^+\tau} - e^{\frac{1}{2}\psi^-\tau}}{\psi^+ e^{\frac{1}{2}\psi^-\tau} - \psi^- e^{\frac{1}{2}\psi^+\tau}} \end{aligned} \quad (3.482)$$

Now we can find A by integrating:

$$A(\tau) = \int_0^\tau k\theta B(s) ds + riz\tau$$

$$\begin{aligned}
&= -\frac{2K\theta}{\eta^2} \int_0^\tau \frac{\dot{E}(s)}{E(s)} ds + riz\tau \\
&= -\frac{2K\theta}{\eta^2} \ln \left(\frac{E(\tau)}{E(0)} \right) + riz\tau \\
&= -\frac{k\theta}{\eta^2} \left[2 \ln \left(\frac{-\psi^- + \psi^+ e^{-\sqrt{\Delta}}}{2\sqrt{\Delta}} \right) + \psi^+ \tau \right]
\end{aligned} \tag{3.483}$$

Notice that this expression is completely explicit.

3.25 The Carr-Madan approach

Consider the specific case of a call option. The main idea of this approach is that a call option is a function not only of the spot but also of the strike. So, instead of considering the Laplace FT (LFT) and the anti-transform of the density function of the underlying (which is unknown) we can take the LFT of the price of the call and then take the inverse transform. The LFT is performed on the strike because we want the price of the call for many strikes (in view of the calibration of the model). So, if we parametrize the price of a call as $\text{Call}(x, l, T)$ where $x = \ln S$ and $l = \ln K$, then we have that

$$\text{Call}(x, l, T) = e^{-r(T-t)} \int_l^{+\infty} (e^y - e^l) f_x(y) dy \tag{3.484}$$

where $f_x(y)$ is the underlying density function. Notice that the call, as a function of l , is not L^1 , in fact

$$\lim_{l \rightarrow -\infty} \text{Call} = e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}}[e^{x_T}]$$

because for $l \rightarrow -\infty$ we have $K \rightarrow 0$ and the price of the call coincides with the underlying. This is the reason why in order to get the integrability of the call price we can't take the LFT but we need something more. So, we take the LFT with respect to l . If we consider a generic complex number $z = \xi + i\eta$, the LFT is given by

$$\widehat{\text{Call}}(z) = \widehat{\text{Call}}(\xi + i\eta) = \int_{\mathbb{R}} e^{i(\xi+i\eta)l} \text{Call}(x, l, t) dl$$

Putting the discount factor on the other side we have that

$$e^{r(T-t)} \widehat{\text{Call}}(z) = \int_{\mathbb{R}} e^{i(\xi+i\eta)l} \left(\int_l^{+\infty} (e^y - e^l) f_x(y) dy \right) dl$$

Let's use Fubini:

$$\begin{aligned}
&= \int_{\mathbb{R}} dy \int_{-\infty}^y e^{i(\xi+i\eta)l} (e^y - e^l) f_x(y) dl \\
&= \int_{\mathbb{R}} f_x(y) dy \int_{-\infty}^y e^{i(\xi+i\eta)l} (e^y - e^l) dl \\
&= \int_{\mathbb{R}} f_x(y) dy \left(e^y \int_{-\infty}^y e^{i(\xi+i\eta)l} dl - \int_{-\infty}^y e^{i(\xi+i\eta)l+l} dl \right)
\end{aligned}$$

Since the real part of the integrand involves $il \cdot i\eta = -l\eta$, the integral converges only if $\eta < 0$.

$$= \frac{1}{i} \int_{\mathbb{R}} f_x(y) e^{y+iy(\xi+i\eta)} \left(\frac{1}{\xi + i\eta} - \frac{1}{\xi + i(\eta-1)} \right) dy$$

$$\begin{aligned}
&= \frac{1}{i} \hat{f}_x(\xi + i(\eta - 1)) \left(\frac{1}{\xi + i\eta} - \frac{1}{\xi + i(\eta - 1)} \right) \\
&= \hat{f}_x(\xi + i(\eta - 1)) \frac{1}{\eta^2 - \eta - \xi^2 + i\xi(1 - 2\eta)}
\end{aligned}$$

Remark 3.25.1. In order to use Fubini we dominate the absolute value of the integrand for both y and l . In other words, we have to bound the quantity

$$\left| e^{i(\xi+i\eta)l} (e^y - e^l) f_x(y) \mathbb{1}_{y>l} \right|$$

Regarding the imaginary part there is no problem because the function is circular, so we can bound with 1 the exponential. Let's focus on the real part. If we replace the minus sign with the plus, we get the trivial inequality

$$\left| e^{-\eta l} (e^y - e^l) f_x(y) \mathbb{1}_{y>l} \right| \leq \mathbb{1}_{y>l} f_x(y) e^y e^{-\eta l} + \mathbb{1}_{y>l} f_x(y) e^{-l(\eta-1)}$$

Now, let's integrate with respect to l :

$$e^y f_x(y) \int_{-\infty}^y e^{-\eta l} dl + f_x(y) \int_{-\infty}^y e^{-l(\eta-1)} dl$$

We see that if $\eta < 0$ then the integrand is bounded, provided that

$$\int_{\mathbb{R}} \left(\frac{e^{y-\eta y}}{|\eta|} f_x(y) \right) < \infty$$

which is equivalent to

$$\mathbb{E} [(e^y)^{1-\eta}] < \infty$$

Since $e^y = S(T)$, this condition asks that the underlying has a finite $(1-\eta)$ -moment, with $1-\eta > 1$ (since η is negative). Once again, we face the problem of moment explosion.

So, if we take $\alpha = -\eta > 0$ such that $\mathbb{E}[S^{1+\alpha}(T)] < \infty$, then the LFT of the call price is given by

$$\begin{aligned}
\widehat{\text{Call}}(z) &= \widehat{\text{Call}}(\xi + i\eta) = \widehat{\text{Call}}(\xi - i\alpha) \\
&= e^{-r(T-t)} f_x(\xi - i(\alpha + 1)) \frac{1}{\alpha^2 + \alpha - \xi^2 + i(2\alpha - 1)}
\end{aligned} \tag{3.485}$$

In order to get the call price, we have to take the inverse LFT. Exploiting the symmetry of the integrand we can consider only the integral from 0 to ∞ :

$$\text{Call}(x, l, T) = \frac{e^{-\alpha l}}{\pi} \int_0^\infty \text{Re} \left\{ e^{-i\xi l} \widehat{\text{Call}}_x(\xi - i\alpha) \right\} d\xi \tag{3.486}$$

In summary:

- Fix $l = \ln K$
- Fix $\alpha > 0$ so that there is no moment explosion
- The price of the call at time $t = 0$ is given by

$$\begin{aligned}
\text{Call}_0 &= \frac{e^{-\alpha l - rT}}{\pi} \int_0^\infty \text{Re} \left\{ e^{-itl} \frac{\hat{f}_x(t - i(\alpha + 1)) e^{i(t-i(\alpha+1))(\ln S(0)+rT)}}{(\alpha + it)(\alpha + 1 + it)} \right\} dt \\
&= \frac{e^{-\alpha l - rT}}{\pi} \int_0^\infty \text{Re} \left\{ \frac{e^{-itl+i(t-i(\alpha+1))(\ln S(0)+rT)}}{(\alpha + it)(\alpha + 1 + it)} \hat{f}_x(t - i(\alpha + 1)) \right\} dt
\end{aligned} \tag{3.487}$$

3.26 Implementation of the Heston model

To do. Also see this article by Lewis.

3.27 Multi-dimensional Riccati equation

In the next sections we will consider the following situations:

- increase the dimension of the volatility factors. In fact, explaining the whole volatility surface in terms of the spot instantaneous variance (which is a 1-dimensional stochastic process) is quite reductive. In order to get a richer explanation we have to introduce more factors;
- If we introduce more factors we have to calibrate according to market data (usually option prices), which can be very difficult and expensive. So, we must try to be parsimonious.

So, we have a trade-off between complexity and parsimony/analytical tractability (at least keep the property of the Heston model, i.e. the possibility to compute the characteristic function). The computation of the characteristic function for the payoff is not a problem, in fact the problem is to find the characteristic function of the model. In order to do that, in the classic Heston model we have to solve the Riccati equation, which is done transforming the quadratic first order ODE into a linear second order ODE. This technique works well in the one-dimensional case but not in the multi-dimensional one.

There are many ways to extend the one-dimensional case, but the most natural just introduces a vector of volatility factors. For example, in order to deal with the two-dimensional Heston model we introduce a 2-d vector with two volatility factors, one for the short term implied volatility surface and one for the long term volatility smile. In this way we can separately manage the smile and the skew of the short term and long term expiring options. However, this is not the best procedure to follow if we want to keep a certain level of analytical tractability and moreover it introduces some intrinsic constraints. In order to understand which is the right technology to exploit and which are the potential issues in the naive generalization, we proceed step by step.

First, we need a technique to solve the Riccati ODE

$$\begin{cases} \dot{B} = \frac{1}{2}\xi^2 B^2 - (K - \rho\xi z)B - \frac{1}{2}(iz + z^2) \\ B(0) = 0. \end{cases}$$

The linearization technique says that – instead of considering that the solution is given by the ratio of the derivative of a function and the function itself – the solution is given by

$$B = H^{-1}F \tag{3.488}$$

where $H \neq 0$ and F are deterministic functions. Assume that H and F are both matrices. Then, if we multiply by H both sides of (3.488) we get

$$HB = F$$

Then, if we differentiate, we get

$$\dot{H}B + H\dot{B} = \dot{F}$$

Now, let's introduce \dot{B} as defined by the Riccati equation:

$$\dot{H}B + \frac{1}{2}\xi^2B^2H - H(K - \rho\xi iz)B - \frac{1}{2}(iz + z^2)H = \dot{F}$$

Rearranging:

$$\dot{F} - \dot{H}B = \frac{1}{2}\xi^2FB - (K - \rho\xi iz)F - \frac{1}{2}(iz + z^2)H$$

Now, let's consider the system obtained by identifying the coefficients of B and the other (constant) terms:

$$\begin{cases} \dot{H} = -\frac{1}{2}\xi^2F \\ \dot{F} = -(K - \rho\xi iz)F - \frac{1}{2}(iz + z^2)H \end{cases}$$

In this way we have transformed the quadratic first order ODE into a system of two linear ODEs, which can be rewritten as

$$\begin{pmatrix} \dot{F} \\ \dot{H} \end{pmatrix} = \begin{pmatrix} -(K - \rho\xi iz) & -\frac{1}{2}(iz + z^2) \\ -\frac{1}{2}\xi^2 & 0 \end{pmatrix} \begin{pmatrix} F \\ H \end{pmatrix} \quad (3.489)$$

So, if F and H are two functions that solve this linear system with the initial conditions

$$F(0) = 0, \quad H(0) = 1 \quad (3.490)$$

(such that $H_0^{-1}F_0 = B_0 = 0$) then we can plug them into the Riccati equation. The solution of eq. (3.489) is given by

$$\begin{aligned} \begin{pmatrix} F_\tau \\ H_\tau \end{pmatrix} &= \exp \left\{ \begin{pmatrix} -(K - \rho\xi iz)\tau & -\frac{1}{2}(iz + z^2)\tau \\ -\frac{1}{2}\xi^2\tau & 0 \end{pmatrix} \right\} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (3.491) \\ &= \begin{pmatrix} A_{11}(\tau) & A_{12}(\tau) \\ A_{21}(\tau) & A_{22}(\tau) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} A_{12}(\tau) \\ A_{21}(\tau) \end{pmatrix} \end{aligned}$$

Then, the solution of the Riccati equation at time τ will be

$$B(\tau) = A_{22}^{-1}(\tau)A_{12}(\tau) \quad (3.492)$$

which – provided that we are able to compute the exponential of the matrix in (3.491) – is completely explicit.

3.27.1 Matrix exponential

Given a $d \times d$ matrix $A \in M(\mathbb{R}_d)$, its exponential is defined as

$$e^A = \mathbb{1} + A + \frac{A^2}{2!} + \dots = \sum_{n=0}^{+\infty} \frac{A^n}{n!}.$$

If A is symmetric, $A = A^T$, all the eigenvalues are real and we can diagonalize it: there exists a matrix P such that $P^{-1} = P^T$ such that

$$A = PDP^{-1}$$

where

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{pmatrix}.$$

In this case the exponential of A is given by

$$\begin{aligned} e^A &= \mathbb{1} + PDP^{-1} + \frac{(PDP^{-1})^2}{2!} + \dots \\ &= PP^{-1} + PDP^{-1} + \frac{PD^2P^{-1}}{2!} + \dots \\ &= P \left(\mathbb{1} + D + \frac{D^2}{2!} + \dots \right) P^{-1} \\ &= Pe^D P^{-1} \\ &= P \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_d} \end{pmatrix} P^{-1}. \end{aligned}$$

Remark 3.27.1. For a symmetric positive semi-definite matrix $A \in S_d^+$ we have that

$$\begin{aligned} \sqrt{A} &= \sqrt{PDP^{-1}} = \sqrt{P\sqrt{D}\sqrt{DP^{-1}}} \\ &= \sqrt{P\sqrt{DP^{-1}}P\sqrt{DP^{-1}}} = \sqrt{(P\sqrt{DP^{-1}})^2} = P\sqrt{DP^{-1}} \end{aligned}$$

This is not the Cholesky decomposition $A = LU$. Here \sqrt{A} is also in S_d^+ .

Remark 3.27.2. In general, \sqrt{A} may not exist or be not unique. For example, $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has no square root and $A = \begin{pmatrix} 33 & 24 \\ 48 & 57 \end{pmatrix}$ has multiple square roots (for example $\begin{pmatrix} 1 & 4 \\ 8 & 5 \end{pmatrix}$ or $\begin{pmatrix} 5 & 2 \\ 4 & 7 \end{pmatrix}$). However, positive semi-definite matrices have only one positive semi-definite square root (*principal square root*).

In conclusion, care must be taken when applying to multiple arguments functions typically defined in \mathbb{R} .

3.28 Multifactor stochastic volatility models

3.28.1 2-Heston model for one asset

Consider an underlying that evolves under the risk neutral probability measure according to

$$\frac{dS(t)}{S(t)} = r dt + \sqrt{V_1(t)} dZ_1(t) + \sqrt{V_2(t)} dZ_2(t) \quad (3.493)$$

where V_1 and V_2 are respectively associated to the short term and the long term behavior of the volatility surface. The dynamics of V_1 and V_2 are given by the Heston model:

$$dV_1(t) = K_1(\theta_1 - V_1(t)) dt + \xi_1 \sqrt{V_1(t)} dW_1(t) \quad (3.494)$$

$$dV_2(t) = K_2(\theta_2 - V_2(t)) dt + \xi_2 \sqrt{V_2(t)} dW_2(t) \quad (3.495)$$

The infinitesimal generator involves the first order as well the second order derivative and the covariation. The problem is that in order to have an affine infinitesimal generator it must be

$$\langle Z_1, Z_2 \rangle = 0; \quad \langle W_1, W_2 \rangle = 0; \quad \langle Z_1, W_2 \rangle = 0; \quad \langle Z_2, W_1 \rangle = 0;$$

$$\langle Z_1, W_1 \rangle = \rho_1; \quad \langle Z_2, W_2 \rangle = \rho_2.$$

In fact, for example, if $\langle W_1, W_2 \rangle \neq 0$ then there will be a term $\sqrt{V_1 V_2}$ and the model will be not affine. However, if we introduce this constraints, there are some consequences. The fact that $W_1 \perp\!\!\!\perp W_2$ implies that the volatility factors are independent. This means that V_1 (short term behavior) and V_2 (long term behavior) are independent, which is not very reasonable. In particular, if

$$Y(t) = \ln S(t) \tag{3.496}$$

then its dynamics is

$$dY(t) = \left(r - \frac{1}{2}(V_1 + V_2) \right) dt + \sqrt{V_1} dZ_1 + \sqrt{V_2} dZ_2. \tag{3.497}$$

Regarding the Fourier pricing, since we know that

$$x_1 \perp\!\!\!\perp x_2 \quad \Rightarrow \quad \mathbb{E}[e^{\lambda_1 x_1}] \mathbb{E}[e^{\lambda_2 x_2}]$$

then we have

$$\mathbb{E}_t^{\mathbb{Q}}[e^{iz \ln S(t)}] = e^{iz \ln S(t) + A(\tau) + B_1(\tau)V_1(t) + B_2(\tau)V_2(t)} \tag{3.498}$$

where $A(\tau)$ is a constant term ($\tau = T - t$) and B_1 and B_2 are solutions of two independent Riccati ODEs. So, it is possible to find the FT of the asset price in terms of the superpositions of the solutions of the corresponding Riccati equations. The problem is that the (instantaneous) correlation between the noise of the returns and the noise of the volatility of the returns is

$$\begin{aligned} \text{Corr}(\text{Noise}(dY), \text{Noise}(\text{Vol}(dY))) &= \frac{d \langle Y, V_1 + V_2 \rangle}{\sqrt{d \langle Y \rangle} \sqrt{d \langle V_1 + V_2 \rangle}} \\ &= \frac{\rho_1 \xi_1 V_1 + \rho_2 \xi_2 V_2}{\sqrt{V_1 + V_2} \sqrt{\xi_1^2 V_1 + \xi_2^2 V_2}} \end{aligned} \tag{3.499}$$

This quantity is also called *leverage* or *stochastic skew*. Recall that in the classic 1-d Heston model the skew (i.e. the slope of the volatility smile) is driven by ρ , in fact

$$\frac{d \langle Y, V \rangle}{\sqrt{d \langle Y \rangle} \sqrt{d \langle V \rangle}} = \frac{\rho \xi V}{\sqrt{V} \sqrt{\xi^2 V}} = \rho$$

In the 2-Heston model this quantity is no more constant, so we can speak about stochastic skew. The presence of stochastic skew is more in line with the real case. The problem is that, in principle, we cannot fit the level ($\sqrt{V_1 + V_2}$, i.e. the instantaneous volatility) and the slope of the smile separately, because they are all mixed together. This means that we cannot use “dedicated” factors.

3.28.2 Multiple assets Heston model

How can we extend the Heston model to the case of multiple assets, for example two stochastic assets? Notice that

Lesson 31 of
21/05/2020

- if we correlate the Brownian motions of the 2-Heston model we have no affine property. If there is no affinity we cannot explicitly write the LFT and so the Riccati equation too. Then, if we are not able to write the Riccati equation we cannot do option pricing in a quick way;
- we want to generate dependence by keeping analytical tractability, i.e. affinity.

In order to solve this problems the idea is to introduce dependence between the two assets by taking a common volatility factor and some other specific volatility factors for each asset. Let's consider three processes

$$\begin{aligned} dX^1(t) &= (\alpha_1 X^1(t) + \beta_1) dt + \sqrt{X^1(t)} dW^1(t) \\ dX^2(t) &= (\alpha_2 X^2(t) + \beta_2) dt + \sqrt{X^2(t)} dW^2(t) \\ dX^3(t) &= (\alpha_3 X^3(t) + \beta_3) dt + \sqrt{X^3(t)} dW^3(t) \end{aligned}$$

with $W^1 \perp\!\!\!\perp W^2 \perp\!\!\!\perp W^3$. Then, let's consider the dynamics of two assets

$$\begin{aligned} \frac{dS^1(t)}{S^1} &= r dt + \sqrt{X^1(t)} dZ^1(t) + \sqrt{X^3(t)} dZ^3(t) \\ \frac{dS^2(t)}{S^2} &= r dt + \sqrt{X^2(t)} dZ^2(t) + \sqrt{X^3(t)} dZ^3(t) \end{aligned}$$

so that X^3 is the common volatility factor while X^1 and X^2 are the specific factors related to the two assets. In this way we obtain a Heston-like model (it is as we consider a 2-Heston for S^1 and a 2-Heston for S^2) which keeps analytical tractability. This means that correlations between W and Z are allowed:

$$Z^i = \rho_i W^i + \sqrt{1 - \rho_i^2} B^i, \quad W^i \perp\!\!\!\perp B^i$$

but Z^1 , Z^2 and Z^3 are independent and $Z^i \perp\!\!\!\perp W^j$ for $i \neq j$.

This model is affine, so we can compute the joint characteristic function of the log-prices

$$\mathbb{E}^{\mathbb{Q}} \left[e^{i(Z^1 \ln S^1(T) + Z^2 \ln S^2(T))} \right] = e^{A(\tau) + B^1(\tau) X^1(t) + B^2(\tau) X^2(t) + B^3(\tau) X^3(\tau)}$$

where B^1 , B^2 and B^3 solve a 1-dimensional Riccati ODE each one.

In order to see which are the consequences of these choices, let's consider the following covariations

$$\begin{aligned} \frac{d\langle S^1, S^1 \rangle}{(S^1)^2} &= (X^1(t) + X^3(t)) dt \\ \frac{d\langle S^2, S^2 \rangle}{(S^2)^2} &= (X^2(t) + X^3(t)) dt \\ \frac{d\langle S^1, S^2 \rangle}{S^1 S^2} &= X^3(t) dt > 0 \end{aligned}$$

Focus on the third equation. We have that the covariance between the two assets is constrained to be positive (or negative, if we change the order). The important thing is that it cannot change sign). This becomes a problem in pricing *correlation products*, i.e. products for which the correlation between assets plays an important role (for example the *rainbow options* or the *basket options*), where the correlation may change sign.

So, we need to go beyond the usual way to define affine positive factors. The first intuition is to consider a new process which implements the “positiveness” in a more general way.

3.28.3 Wishart processes

The starting point is the squared Bessel process of dimension $n > 1$:

$$dX(t) = 2\sqrt{X(t)} dW(t) + n dt$$

Recall that if the dimension is an integer, it is possible to show that

$$X(t) = B(t)^T B(t)$$

where $B(t) \in \mathbb{R}^n$ is a Brownian motion. We also know that for $n = \delta \geq 0$, $\delta \in \mathbb{R}$ it is still possible to define the notion of squared Bessel process BESQ^δ with dynamics

$$dX(t) = 2\sqrt{X(t)} dW(t) + \delta dt$$

with the difference that now we cannot refer to the norm of a Brownian motion vector. This equation admits a strong solution for every $\delta \geq 0$. The corresponding infinitesimal generator is given by

$$\mathcal{A} = 2XD^2 + \delta D, \quad \text{where } D = \frac{d}{dx}.$$

For fixed t , the “matrix extension” of $X(t)$ follows the *Wishart distribution*. If B is a $n \times d$ matrix of gaussian random variables:

$$B = \begin{pmatrix} B_{11} & \cdots & B_{1d} \\ \vdots & \ddots & \vdots \\ B_{n1} & \cdots & B_{nd} \end{pmatrix} \quad (3.500)$$

then

$$(B^T B)_{d \times d} \sim \text{Wishart} \quad (3.501)$$

i.e. $B^T B$ follows a Wishart distribution, which is the matrix extension of the χ^2 .

Example 3.28.1

For $d = 1$ we have

$$B^T B = B_{11}^2 + \cdots + B_{n1}^2 = \text{BESQ}^n$$

where $\text{BESQ}^n \sim \chi_n^2$.

For $n = 1$ we have

$$B^T B = \begin{pmatrix} B_{11} \\ \vdots \\ B_{1n} \end{pmatrix} (B_{11} \ \cdots \ B_{1d}) = \begin{pmatrix} \\ \\ \end{pmatrix}_{d \times d}$$

with $\text{rank}(B^T B) = 1$. This means that the information carried out by the random variable may be degenerate. Thus this leads to a degenerate distribution, which can even be not continuous with respect to the Lebesgue measure. However, in order to avoid this degeneration, we will consider only cases with $n \geq d$.

Now, let's give the definition of Wishart process.

Definition 3.28.1 (Wishart process). A Wishart process of dimension $d \geq 1$, index $n \geq d$ and initial state $s_0 = C^T C \in S_d^+$ is denoted as $\text{WIS}(n, d, s_0)$ and is the matrix process

$$S(t) = N(t)^T N(t) \quad (3.502)$$

where $N(t)$ is a Brownian motion taking values in the set of rectangular matrices $M_{n \times d}$ with initial value equal to a constant matrix C , $N(0) = C$.

Then, we have the first important property.

Theorem 3.28.1. *The following hold:*

1. *The dynamics of a Wishart process is given by*

$$dS(t) = \sqrt{S(t)} dB(t) + dB(t)^T \sqrt{S(t)} + n\mathbf{1}_{d \times d} dt. \quad (3.503)$$

where $B(t) \in M_{d \times d}$ is a Brownian motion.

2. *The infinitesimal generator of a Wishart process S is given by*

$$\mathcal{A} = \text{Tr}(nD + 2SD^2) \quad (3.504)$$

where

$$D = \left(\frac{\partial}{\partial S_{ij}} \right)_{ij}.$$

Proof. 1. Using Itô, we have

$$\begin{aligned} dS(t) &= d(N(t)^T N(t)) = (dN^T)N + N^T dN + \underbrace{d\langle N^T, N \rangle_t}_{n\mathbf{1}_{d \times d} dt} \\ &= dN^T N \frac{\sqrt{N^T N}}{\sqrt{N^T N}} + N^T dN \frac{\sqrt{N^T N}}{\sqrt{N^T N}} + n\mathbf{1}_{d \times d} dt. \end{aligned}$$

Now recall that $N(t)$ is a matrix Brownian motion if and only if

$$\text{Cov}_t(dN(t) \cdot \alpha, dN(t) \cdot \beta) = \mathbb{E}_t[(dN(t)\alpha)(dN(t)\beta)^T] = \alpha^T \beta \mathbf{1}_{n \times n} dt$$

for all $\alpha, \beta \in \mathbb{R}^d$. As a consequence we can write

$$\frac{N(t)^T}{\sqrt{N(t)^T N(t)}} dN(t) = dB(t)$$

in fact

$$\begin{aligned} \text{Cov} \left(\frac{N(t)^T dN(t)}{\sqrt{N(t)^T N(t)}} \alpha, \frac{N(t)^T dN(t)}{\sqrt{N(t)^T N(t)}} \beta \right) &= \mathbb{E} \left[N(t)^T \frac{(dN(t)\alpha)(dN(t)\beta)^T}{N(t)^T N(t)} N(t) \right] \\ &= \alpha^T \beta \mathbf{1}_n dt. \end{aligned}$$

2. The infinitesimal generator is given by

$$\begin{aligned} \mathcal{A} &= \text{Tr}(nD + 2SD^2) \\ &= n \frac{\partial}{\partial S_{11}} + n \frac{\partial}{\partial S_{22}} + \frac{1}{2} \left[4S_{11} \frac{\partial^2}{\partial S_{11}^2} + 4S_{22} \frac{\partial^2}{\partial S_{22}^2} + 4(S_{11} + S_{22}) \frac{\partial^2}{\partial S_{12}^2} + \right. \\ &\quad \left. + 4 \left(2S_{12} \frac{\partial^2}{\partial S_{11} \partial S_{12}} + 2S_{12} \frac{\partial^2}{\partial S_{12} \partial S_{22}} + 2 \underbrace{\langle S_{11}, S_{22} \rangle}_{=0} \frac{\partial^2}{\partial S_{11} \partial S_{22}} \right) \right] \end{aligned}$$

where we used the fact that

$$\begin{aligned} \langle S_{11} \rangle &= 4S_{11} \\ \langle S_{22} \rangle &= 4S_{22} \\ \langle S_{12} \rangle &= S_{11} + S_{22} \\ \langle S_{11}, S_{12} \rangle &= 2S_{12} \end{aligned}$$

$$\langle S_{12}, S_{22} \rangle = 2S_{12}.$$

Now let's recover the infinitesimal generator from the dynamics of $S(t)$. We know that

$$dS(t) = \dots dt + \underbrace{\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}}_{\sqrt{S(t)}} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

Since S is given by the square of the matrix σ , we can write

$$S(t) = \begin{pmatrix} \sigma_{11}^2 + \sigma_{12}^2 & \sigma_{11}\sigma_{12} + \sigma_{12} + \sigma_{22} \\ \sigma_{11}\sigma_{12} + \sigma_{12} + \sigma_{22} & \sigma_{12}^2 + \sigma_{22}^2 \end{pmatrix}.$$

Then

$$\begin{aligned} dS_{11} &= \dots dt + 2\sigma_{11} dB_{11} + 2\sigma_{12} dB_{21} \\ dS_{12} &= \dots dt + \sigma_{11} dB_{12} + \sigma_{12} dB_{22} + \sigma_{12} dB_{11} + \sigma_{22} dB_{21} \\ dS_{22} &= \dots dt + 2\sigma_{12} dB_{12} + 2\sigma_{22} dB_{22} \end{aligned}$$

and

$$\begin{aligned} d\langle S_{11} \rangle_t &= 4(\sigma_{11}^2 + \sigma_{12}^2)dt = 4S_{11} dt \\ d\langle S_{22} \rangle_t &= 4(\sigma_{12}^2 + \sigma_{22}^2)dt = 4S_{22} dt \\ d\langle S_{12} \rangle_t &= \sigma_{11}^2 + \sigma_{12}^2 + \sigma_{22}^2 = S_{11} + S_{12} \\ d\langle S_{11}, S_{12} \rangle_t &= 2\sigma_{11}\sigma_{12} + 2\sigma_{12}\sigma_{22} = 2S_{12} \\ d\langle S_{12}, S_{22} \rangle_t &= 2\sigma_{11}\sigma_{12} + 2\sigma_{12}\sigma_{22} = 2S_{12} \\ d\langle S_{12}, S_{22} \rangle_t &= 0 \end{aligned}$$

Notice that all these terms are linear in S . This is the reason why the infinitesimal generator of a Wishart process is affine, in the sense that it is written in terms of an affine combination involving the elements of the matrix S . Finally

$$\begin{aligned} 2\langle S_{11}, S_{12} \rangle \frac{\partial^2}{\partial S_{12}^2} &= \langle S_{11}, S_{12} \rangle \left(\frac{\partial^2}{\partial S_{12}^2} + \frac{\partial^2}{\partial S_{21}^2} \right) \\ 4\langle S_{11}, S_{12} \rangle \frac{\partial^2}{\partial S_{11} \partial S_{12}} &= 2\langle S_{11}, S_{12} \rangle \left(\frac{\partial^2}{\partial S_{11} \partial S_{12}} + \frac{\partial^2}{\partial S_{11} \partial S_{21}} \right) \end{aligned}$$

□

Now we would like to prove that the Wishart process is affine, in the sense that the LFT is the exponential function of deterministic matrices that satisfy a certain ODE. We will study a specific case in which we consider only the LFT of the Wishart process (without the integral) and we can find the solution without having to solve the Riccati equation. Then we will generalize this result.

Lesson 32 of
22/05/2020

Laplace transform of the Wishart process

Given a $d \times d$ Brownian motion $B(t)$ with i.i.d components, whenever the SDE

$$\begin{cases} dS(t) = \sqrt{S(t)} dB(t) + dB(t)^T \sqrt{S(t)} + n\mathbf{1} dt \\ S(0) = s_0 \end{cases} \quad (3.505)$$

has a solution in S_d^+ , the distribution of $S(t)$ for a fixed t is given by its Laplace transform

$$\mathbb{E}_{s_0} \left[e^{-\text{Tr}(uS(t))} \right] = \det(\mathbb{1} + 2tu)^{-\frac{n}{2}} \exp\{-\text{Tr}(s_0(\mathbb{1} + 2tu)^{-1}u)\} \quad (3.506)$$

for all $u \in S_d^+$ (in order to have convergence). This is a very strong result, which tells us that the Laplace transform of the Wishart process is given by a deterministic function times the exponential of a linear combination of the elements of s_0 . This means that the second term is exponentially affine with respect to S . So, we can extend to the matrix framework the affinity of the usual CIR process.

Proof. The idea behind this proof is to show that the Feynman-Kac formula holds true. That is, if we write the expected value (3.506) as a function $U(t, s)$, then it satisfies the Feynman-Kac PDE

$$\begin{cases} \frac{\partial U(t, s)}{\partial t} + \mathcal{A}U(t, s) = 0 \\ U(T, s) = e^{-\text{Tr}(u \cdot s)} \end{cases}$$

For our convenience, it is better to transform this variable in time to maturity: let's introduce the auxiliary function $V(T - t, s) = U(t, s)$, so that the Feynman-Kac formula becomes

$$\begin{cases} -\frac{\partial V(T-t, s)}{\partial t} + \mathcal{A}V(T-t, s) = 0 \\ V(0, s) = U(T, s) = e^{-\text{Tr}(u \cdot s)} \end{cases}$$

Our “candidate” solution is

$$V(t, s) = \det(\mathbb{1} + 2tu)^{-\frac{n}{2}} \exp\{-\text{Tr}(s(\mathbb{1} + 2tu)^{-1}u)\} \quad (\diamond)$$

so if we compute the derivatives they have to satisfy the Feynman-Kac PDE. In order to do that we need to differentiate the determinant of a matrix.

Lemma 3.28.2 (Jacobi formula). *The derivative of a deterministic time-dependent matrix $H(t)$ is given by*

$$\frac{d}{dt} \det(H(t)) = \det(H(t)) \cdot \text{Tr}\left(H^{-1}(t) \cdot \frac{d}{dt} H(t)\right)$$

Proof. The proof of this lemma is based on the fact that if we choose a row i then we can write the determinant as the sum of the elements H_{ik} of this row times the comatrix $\text{Com}(H_{ik})$

$$\det(H) = \sum_k H_{ik} \text{Com}(H_{ki}).$$

Recall that the $\text{Com}(M)$ of a $(n \times n)$ -matrix M is the $(n \times n)$ -matrix whose entry in (i, k) is $(-1)^{i+k}$ times the determinant of the matrix which you get by deleting in M the line l and the row k . \square

In this case

$$\text{Com}(H_{ki}) = (-1)^{k+i} \det(H \text{ without row } k \text{ and column } i)$$

Being the complement of H_{ki} , the comatrix $\text{Com}(H_{ki})$ is independent of the elements H_{ki} . This means that

$$\frac{\partial \det(H)}{\partial H_{ij}} = \text{Com}(H_{ji})$$

and then

$$\begin{aligned}\frac{d}{dt} \det(H) &= \sum_j \left(\sum_i \frac{\partial \det(H)}{\partial H_{ij}} \right) \frac{dH_{ij}}{dt} \\ &= \sum_j \sum_i \text{Com}(H_{ji}) \frac{dH_{ij}}{dt} \\ &= \text{Tr}(\text{Adj}(H)dH)\end{aligned}$$

where $\text{Adj}(H)$ is the adjugate matrix, i.e. the transpose of the comatrix.

Let's compute the derivatives of (\diamond) :

$$\begin{aligned}\frac{\partial V}{\partial t} &= -\frac{n}{2} \det(\mathbb{1} + 2tu)^{-\frac{n}{2}-1} \det(\mathbb{1} + 2tu) \text{Tr}((\mathbb{1} + 2tu)^{-1} 2u) e^{-\text{Tr}(s(\mathbb{1} + 2tu)^{-1} u)} - \\ &\quad - V(t, s) \frac{d}{dt} \text{Tr}(s(\mathbb{1} + 2tu)^{-1} u) \\ &= -nV(t, s) \text{Tr}((\mathbb{1} + 2tu)^{-1} u) + 2V(t, s) \text{Tr}(s(\mathbb{1} + 2tu)^{-2} u^2)\end{aligned}$$

Now, let's compute the infinitesimal generator:

$$\mathcal{A}V = \text{Tr}(nDV + 2SD^2V)$$

where

$$\begin{aligned}DV &= \frac{\partial V}{\partial S_{ij}} = -V(t, s)D \text{Tr}(S(\mathbb{1} + 2tu)^{-1} u) \\ &\stackrel{(a)}{=} -V(t, s)(\mathbb{1} + 2tu)^{-1} u\end{aligned}$$

where in (a) we used the fact that if we denote $(\mathbb{1} + 2tu)^{-1} u \equiv B_{ij}$, we have that

$$\text{Tr}(S(\mathbb{1} + 2tu)^{-1} u) = \sum_{i,j} S_{ij} \cdot B_{ij}$$

and the derivative is

$$D \text{Tr}(S(\mathbb{1} + 2tu)^{-1} u) = \frac{\partial}{\partial S_{ij}} \sum_{i,j} S_{ij} \cdot B_{ij} = B_{ij} = (\mathbb{1} + 2tu)^{-1} u$$

Then, the second derivative is

$$D^2V = V(t, s)(\mathbb{1} + 2tu)^{-1} u (\mathbb{1} + 2tu)^{-1} u$$

So, the infinitesimal generator reads

$$\begin{aligned}\mathcal{A}V &= -nV(t, s) \text{Tr}((\mathbb{1} + 2tu)^{-1} u) + 2V(t, s) \text{Tr}((\mathbb{1} + 2tu)^{-1} u (\mathbb{1} + 2tu)^{-1} u) \\ &\stackrel{(a)}{=} -nV(t, s) \text{Tr}((\mathbb{1} + 2tu)^{-1} u) + 2V(t, s) \text{Tr}((\mathbb{1} + 2tu)^{-1} (\mathbb{1} + 2tu)^{-1} u)\end{aligned}$$

where in (a) we use the fact that

$$u(\mathbb{1} + 2tu)^{-1} u = (\mathbb{1} + 2tu)^{-1} u$$

because the two matrices commute. Then, indeed we have

$$-\frac{\partial V(T-t, s)}{\partial t} + \mathcal{A}V(T-t, s) = 0$$

so the Wishart process is affine. \square

Now that we proved that the Wishart process is affine, let's consider some properties of this process in terms of the existence of weak or strong solutions.

Theorem 3.28.3. *Let $B(t)$ be a $d \times d$ Brownian motion with independent components and let $S(t)$ be the process satisfying the SDE*

$$\begin{cases} dS(t) = \sqrt{S(t)} dB(t) + dB(t)^T \sqrt{S(t)} + \alpha \mathbf{1} dt, & \alpha \in \mathbb{R} \\ S(0) = s_0 \in S_d^+. \end{cases} \quad (3.507)$$

Then:

1. if $\alpha \geq d - 1$ then there exists a weak solution for the SDE (the path-wise uniqueness of the solution is not guaranteed);
2. if $\alpha \geq d + 1$ then there exists a strong solution for the SDE, related to the fact that $\det(S(t)) > 0$;
3. if $\alpha \in \{1, \dots, d - 2\}$ (Gindikin set) we have no more a Wishart distribution, in fact the distribution is low degenerate and not absolutely continuous with respect to the Lebesgue measure. However, it is still possible to find a weak solution;
4. if $\alpha \geq d + 1$ then the eigenvalues of the matrix are all distinct for all t .
5. if $\alpha \geq d - 1$ then $\det(S(t)) \geq 0$, if $\alpha \geq d + 1$ then $\det(S(t)) > 0$ for all t .
6. if $\alpha \geq d - 1$ the eigenvalues of the matrix never collide:

$$\lambda_1(t) > \dots > \lambda_d(t) \geq 0.$$

In this case

$$d\lambda_i = 2\sqrt{\lambda_i} d\tilde{B}_i + \left(\alpha + \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) dt, \quad \forall i = 1, \dots, d$$

The term $\frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k}$ is the reason why the eigenvalues never collide (if they collide there is an explosion of the denominator).

Extension to mean reversion

Let us consider a $n \times d$ -matrix $(X(t))_{n \times d}$ which is solution of the SDE

$$\begin{cases} dX(t) = \gamma dB(t) + \beta X(t) dt \\ X(0) = x \in M_{n \times d} \end{cases} \quad (3.508)$$

where $B(t)$ is a $n \times d$ Brownian motion matrix, $\gamma \in \mathbb{R}$ and $\beta \in \mathbb{R}^-$ (in order to keep the stationarity of the process). This SDE describes a Ornstein-Uhlenbeck process, a stationary Gauss-Markov process which, over time, tends to drift towards its mean function (such a process is called *mean-reverting*). We would like to construct a Wishart process starting from this.

Theorem 3.28.4. *Set $S(t) = X(t)^T X(t)$ with initial condition $s = x^T x$. Then*

$$\begin{cases} dS(t) = \gamma(\sqrt{S} dB + dB(t)\sqrt{S}) + 2\beta S(t) + n\gamma^2 \mathbf{1}_d dt \\ S(0) = s. \end{cases} \quad (3.509)$$

Proof. Using Itô and (3.508):

$$\begin{aligned} dS(t) &= d(X(t)^T X(t)) = (dX(t))^T X(t) + X(t)^T dX(t) + d\langle X^T, X \rangle_t \\ &= \gamma X(t) dB(t)^T + \beta X(t)^T X(t) dt + X(t)^T (\gamma dB(t) + \\ &\quad + \beta X(t) dt) + \gamma^2 dB(t)^T dB(t) \end{aligned}$$

Recall that $dB(t)^T dB(t) = n\mathbb{1}_d$ and in order to change the BM multiply and divide for $\sqrt{X(t)^T X(t)}$:

$$\begin{aligned} &= \gamma \left(\sqrt{X(t)^T X(t)} \sqrt{X(t)^T X(t)}^{-1} X(t)^T dB(t) + \right. \\ &\quad \left. + dB(t)^T X(t) \sqrt{X(t)^T X(t)} \sqrt{X(t)^T X(t)}^{-1} \right) + 2\beta X(t)^T X(t) dt + \gamma^2 n\mathbb{1}_d dt \end{aligned}$$

We can define the Brownian motion $d\tilde{B}(t) \equiv \sqrt{X(t)^T X(t)}^{-1} X(t)^T dB(t)$, in fact

$$\text{Cov}(X(t)^T dB, X(t)^T dB) = X(t)^T dB dB^T X(t) = nX(t)^T X(t)$$

such that $\text{Cov}(\tilde{B}, \tilde{B}) = n\mathbb{1}$.

$$= \gamma(\sqrt{S(t)} d\tilde{B}(t) + d\tilde{B}^T \sqrt{S(t)}) + 2\beta S(t) dt + \gamma^2 n\mathbb{1}_d dt$$

□

So, we can include a mean reverting drift just by taking a Ornstein-Uhlenbeck process instead of a pure Brownian motion. Another result is the following.

Theorem 3.28.5. Let $a \in GL(d)$ (invertible $d \times d$ matrix), $b \in M_d^-$ (eigenvalues with negative real part) and $X(t)$ a Ornstein-Uhlenbeck process defined by

$$\begin{cases} dX(t) = a dB(t) + X(t)b dt \\ X(0) = x \end{cases} \quad (3.510)$$

where $B(t)$ is a $n \times d$ Brownian motion. If we define the Wishart process

$$S(t) = X(t)^T X(t), \quad s = x^T x \quad (3.511)$$

then

$$dS(t) = \sqrt{S(t)} dW(t) \sqrt{a^T a} + \sqrt{a^T a} dW(t)^T \sqrt{S(t)} + (b^T S(t) + S(t)b) dt + na^T a dt \quad (3.512)$$

where $na^T a$ can be replaced by a matrix C such that $C - (d-1)a^T a \in S_d^+$.

Proof. Homework. □

Now we have a Wishart process with mean reversion that evolves according to a matrix SDE with matrix coefficients, we can compute the Laplace transform.

Laplace transform

Let's denote our Wishart process representing the volatility as $\Sigma(t)$. Its dynamics is given by

$$d\Sigma(t) = (\beta Q^T Q + M\Sigma(t) + \Sigma(t)M^T)dt + \sqrt{\Sigma(t)} dWQ + Q^T (dW(t))^T \sqrt{\Sigma(t)} \quad (3.513)$$

where Q is a constant matrix, $M \in M^-$ (to grant stationarity) and $W(t)$ is a $d \times d$ Brownian motion. In order to get at least a weak solution, we have to impose $\beta \geq d - 1$. This is the dynamics that guarantee the symmetry and notice that the starting point is a $d \times d$ matrix BM that in principle is not symmetric (all the components are independent). This means that there are d^2 sources of noise. However, the information we can deduce from this matrix is much less because the process is symmetric. The infinitesimal generator for this process is

$$\mathcal{A}_\Sigma = \text{Tr}[(\beta Q^T Q + M\Sigma + \Sigma M^T)D + 2\Sigma D Q^T Q D] \quad (3.514)$$

Remark 3.28.1. If we write the dynamics of the elements of the matrix Σ in the case $d = 2$, we have

$$\begin{aligned} d\Sigma^{11} &= \dots dt + 2\sigma^{11}(Q_{11} dW^{11} + Q_{21} dW^{12}) + 2\sigma^{12}(Q_{11} dW^{21} + Q_{21} dW^{22}) \\ d\Sigma^{11} &= \dots dt + 2\sigma^{12}(Q_{12} dW^{11} + Q_{22} dW^{12}) + 2\sigma^{22}(Q_{21} dW^{21} + Q_{22} dW^{22}) \end{aligned}$$

Then, if we compute the quadratic covariation between the first element in the diagonal and the last one (i.e. the product of the two dynamics), we get

$$\begin{aligned} d\langle \Sigma^{11}, \Sigma^{22} \rangle_t &= 4(\sigma^{11}\sigma^{12} + \sigma^{21}\sigma^{22})(Q_{11}Q_{12} + Q_{21}Q_{22})dt \\ &= 4\Sigma^{12}(Q_{11}Q_{12} + Q_{21}Q_{22})dt \neq 0 \text{ if } Q \text{ is not diagonal.} \end{aligned}$$

When we introduced the first specification of the Wishart process we took $Q = \mathbb{1}$, which is diagonal. That's why we were not able to exploit the Wishart property that the covariation between the two positive factors is not trivial (equal to zero). If now Q is not diagonal, the covariation between $\Sigma^{11} > 0$ and $\Sigma^{22} > 0$ is not zero (like in the usual affine model, where it is the condition to have affinity) and, even more, it is given by the off-diagonal element Σ^{12} . So, *the Wishart process can be seen as a matrix fluctuating in time in a mean reversion way where there is the joint presence of the volatility factors (diagonal terms) together with the covariance (off-diagonal terms). So, also the correlations will be stochastic, even if we don't lose the analytical tractability.*

Linearization approach for the computation of the Laplace transform

Lesson 33 of
27/05/2020

Now that we know which is the infinitesimal generator for the Wishart process we would like to verify that the process is affine through a different procedure. Instead of solving the Feynman-Kac PDE just by checking if a candidate is a solution we would like to extend the procedure we gave in the context of the CIR model where we doubled the dimension of the problem and used a linear system of ODEs. Basically, we have to adapt that procedure to a matrices framework.

If we define the Laplace transform of the Wishart process Σ as the conditional expected value

$$\psi_\Sigma(\tau, \Sigma(t)) = \mathbb{E}_t \left[e^{-\text{Tr}(\Lambda\Sigma(T))} \right], \quad (3.515)$$

where $\Lambda \in S^+$ and $\tau = T - t$, then it satisfies the Feynman-Kac formula

$$\frac{\partial \psi}{\partial \tau} = \mathcal{A}_\Sigma \psi. \quad (3.516)$$

Since we know that the Wishart process is affine, we have a natural candidate solution:

$$\psi(\tau, \Sigma(t)) = e^{-\text{Tr}(A(\tau)\Sigma(t)) + a(\tau)}. \quad (3.517)$$

where $a(\tau)$ is a scalar deterministic function and $A(\tau)$ is a matrix-valued deterministic function. Replacing in (3.516), we get

$$\begin{cases} -\text{Tr}(\dot{A}(\tau)\Sigma(t)) + \dot{a}(\tau) = \text{Tr}[(\beta Q^T Q + M\Sigma + \Sigma M^T)A(\tau) + 2\Sigma A(\tau)Q^T Q A(\tau)] \\ A(0) = \Lambda \\ a(0) = 0. \end{cases}$$

By identification of the parameters, we have that

$$\begin{cases} \dot{a}(\tau) = -\text{Tr}(\beta Q^T Q A(\tau)) \\ a(0) = 0 \end{cases}$$

and

$$\begin{cases} \dot{A}(\tau) = A(\tau)M + M^T A(\tau) - 2A(\tau)Q^T Q A(\tau) \\ A(0) = \Lambda, \end{cases}$$

which is the matrix Riccati ODE. Now we linearize by setting

$$A(\tau) = F^{-1}(\tau)G(\tau), \quad F(0) = \mathbb{1}, \quad G(0) = \Lambda,$$

where $F(\tau) \in GL_d$ in order to be able to take the inverse. If we take the derivative of G with respect to τ we get

$$\frac{dF(\tau)A(\tau)}{d\tau} = \frac{dF(\tau)}{d\tau}A(\tau) + F(\tau)\frac{dA(\tau)}{d\tau}.$$

Rearranging we get

$$\dot{G} - \dot{F}A = GM + (FM^T - 2GQ^T Q)A. \quad (3.518)$$

This equation can be seen as the projection of a vector along the direction given by $(\mathbb{1}, -A)$. If we identify the coefficients of A and the other terms we get a system of two equations:

$$\begin{cases} \dot{G} = GM \\ \dot{F} = -FM^T + 2GQ^T Q, \end{cases} \quad (3.519)$$

which can be written as

$$\frac{d}{d\tau} = (G(\tau), F(\tau)) \underbrace{\begin{pmatrix} M & 2Q^T Q \\ 0 & -M^T \end{pmatrix}}_{2d \times 2d} \quad (3.520)$$

This linear system can be solved by taking the exponential of the matrix:

$$\begin{aligned} (G(\tau), F(\tau)) &= (G(0) \quad F(0)) \exp \left\{ \tau \begin{pmatrix} M & 2Q^T Q \\ 0 & -M^T \end{pmatrix} \right\} \\ &= (\Lambda \quad \mathbb{1}) \exp \left\{ \tau \begin{pmatrix} M & 2Q^T Q \\ 0 & -M^T \end{pmatrix} \right\} \\ &= (\Lambda \quad \mathbb{1}) \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \\ &= \left(\underbrace{\Lambda A_{11}(\tau) + A_{21}(\tau)}_G \quad \underbrace{\Lambda A_{12}(\tau) + A_{22}(\tau)}_F \right). \end{aligned} \quad (3.521)$$

So, the solution of the Riccati ODE is

$$A(\tau) = F^{-1}G = (\Lambda A_{12}(\tau) + A_{22}(\tau))^{-1}(\Lambda A_{11}(\tau) + A_{21}(\tau)). \quad (3.522)$$

Then, we can find $a(\tau)$:

$$\begin{aligned}\dot{a}(\tau) &= -\beta \operatorname{Tr}(Q^T Q A(\tau)) \\ &= -\beta \operatorname{Tr}(Q^T Q F^{-1} G)\end{aligned}$$

From $\dot{F} = -FM^T + 2GQ^T Q$ we have that $G = \frac{1}{2}(\dot{F} + FM^T)(Q^T Q)^{-1}$:

$$= -\frac{\beta}{2} \operatorname{Tr}(Q^T Q F^{-1} (F + FM^T)(Q^T Q)^{-1})$$

Recalling that $\partial_z \operatorname{Tr}(\ln A(z)) = \operatorname{Tr}((\partial_z A)A^{-1})$ we recognize the logarithm of F :

$$= -\frac{\beta}{2} \operatorname{Tr}(d \ln(F) + M^T)$$

By integrating we get

$$a(\tau) = -\frac{\beta}{2} \operatorname{Tr}(\ln F(\tau) + M^T \tau). \quad (3.523)$$

In conclusion, we have the “explicit” (numerical) Laplace transform. Now we have to extend the Laplace transform of the Wishart to the case in which there is also a time integral of the process.

Extension to the case of time integral

Consider the case in which the Laplace transform includes also a time integral:

$$\psi_{\Sigma}(\tau, \Sigma(t)) = \mathbb{E}_t \left[e^{-\operatorname{Tr}(\Lambda \Sigma(T)) + \Gamma \int_t^T \Sigma(s) ds} \right], \quad (3.524)$$

where $\Lambda, \Gamma \in S^+$. According to the Feynman-Kac methodology ψ satisfies the PDE

$$\frac{\partial \psi}{\partial \tau} = \mathcal{A}_{\Sigma} \psi - \operatorname{Tr}(\Gamma \Sigma) \psi. \quad (3.525)$$

A candidate solution is given by

$$\psi(\tau, \Sigma(t)) = e^{-\operatorname{Tr}(A(\tau) \Sigma(t)) + a(\tau)}. \quad (3.526)$$

These equations lead to the following Riccati matrix ODEs:

$$\begin{cases} \dot{A}(\tau) = A(\tau)M + M^T A(\tau) - 2A(\tau)Q^T Q A(\tau) + \Gamma \\ A(0) = \Lambda \\ \\ \dot{a}(\tau) = -\beta \operatorname{Tr}(Q^T Q A(\tau)) \\ a(0) = 0. \end{cases}$$

As before, we set

$$A(\tau) = F^{-1}(\tau)G(\tau)$$

and we differentiate:

$$\frac{dF \cdot A}{d\tau} = \dot{F}A + F\dot{A} = \dot{F}A + F(AM + M^T A - 2AQ^T Q A + \Gamma).$$

So, we can write

$$\dot{G} - \dot{F}A = GM + (FM^T - 2GQ^T Q)A + FF\Gamma,$$

which leads to the linear system

$$\begin{cases} \dot{G} = GM + F\Gamma \\ \dot{F} = -FM^T + 2GQ^TQ, \end{cases}$$

which can be written as

$$(\dot{G} \quad \dot{F}) = (G \quad F) \begin{pmatrix} M & 2Q^TQ \\ \Gamma & -M^T \end{pmatrix}.$$

The solution is

$$A(\tau) = (\Lambda A_{12} + A_{22})^{-1}(\Lambda A_{11} + A_{21}), \quad (3.527)$$

where

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \exp \left\{ \tau \begin{pmatrix} M & 2Q^TQ \\ \Gamma & -M^T \end{pmatrix} \right\}. \quad (3.528)$$

Then we can find $a(\tau)$ as before.

3.28.4 A multi-factor Wishart volatility model

Now that we have the Laplace transform of the Wishart process we can consider a model in which the Wishart process plays the role of the volatility. We will see that this model extends the Heston model in a non-trivial way.

As usual, the starting point is the dynamics of the underlying under the risk neutral measure:

$$\frac{dS(t)}{S(t)} = r dt + \text{Tr} \left(\sqrt{\Sigma(t)} dZ(t) \right), \quad (3.529)$$

where $Z(t)$ is a Brownian motion that drives the asset returns, which is partially correlated with $W(t)$, and

$$d\Sigma(t) = (\beta Q^T Q + M\Sigma(t) + \Sigma(t)M^T)dt + \sqrt{\Sigma(t)} dW(t)Q + Q^T dW^T(t)\sqrt{\Sigma(t)}, \quad (3.530)$$

where $Q^T Q$ is invertible, M is negative semi-definite (in order to have stationarity), $\beta \geq n - 1$ (in order to have at least a weak solution) and $\Sigma(t) \in S_n^+$.

Both W and Z are $(n \times n)$ -Brownian motions, so there are n^2 possibilities of correlation. However, we have to keep the affinity property, for example setting

$$Z(t) = W(t) \cdot R^T + B(t) \sqrt{1 - RR^T}, \quad (3.531)$$

where $R \in M_{n \times n}$ and $B(t)$ is a Brownian motion such that $W(t) \perp\!\!\!\perp B(t)$.

Proposition 3.28.1. *If we define Z as above, then it is a (matrix) Brownian motion.*

Proof. We have to show that $\forall \alpha_1, \alpha_2 \in \mathbb{R}^n$

$$\text{Cov}_t(dZ\alpha_1, dZ\alpha_2) = \mathbb{E}_t[(dZ\alpha_1)(dZ\alpha_2)^T] = \alpha_1^T \alpha_2 \mathbb{1}_n dt$$

By replacing the definition of Z we get

$$\text{Cov}_t(dZ\alpha_1, dZ\alpha_2) = \mathbb{E}_t[(dWM^T\alpha_1 + dB\sqrt{1 - RR^T}\alpha_1)(dWM^T\alpha_2 + dB\sqrt{1 - RR^T}\alpha_2)^T]$$

Use the fact that $W(t) \perp\!\!\!\perp B(t)$:

$$= \text{Cov}_t(dWR^T\alpha_1, dWR^T\alpha_2) + \text{Cov}_t(dB\sqrt{1 - RR^T}\alpha_1, dB\sqrt{1 - RR^T}\alpha_2)$$

Now use the fact that W and B are true (matrix) Brownian motions:

$$\begin{aligned} &= \alpha_1^T RR^T \alpha_2 \mathbb{1} dt + \alpha_1^T (1 - RR^T) \alpha_2 \mathbb{1} dt \\ &= \alpha_1^T \alpha_2 \mathbb{1} dt \end{aligned}$$

So Z is a true Brownian motion. □

Remark 3.28.2. This is not the only possibility to correlate Z with W . This parametrization is quite natural because it guarantees linearity and it is analogous with what we did in the scalar case.

If we are interested in pricing a call:

$$\text{price}_t(\text{Call}) = -\frac{e^{-r(T-t)}}{2\pi} \int_{\mathbb{Z}} \frac{K^{iZ+1}}{Z^2 + iZ} \hat{f}(-Z) dZ$$

we need

$$\hat{f}(Z) = \mathbb{E}_t[e^{iZ \ln S(T)}].$$

Let's introduce the function

$$\psi_{\gamma,t}(\tau) = \mathbb{E}_t[e^{\gamma \ln S(t+\tau)}] = \mathbb{E}_t[e^{\gamma Y(t+\tau)}], \quad (3.532)$$

where $Y(t) = \ln S(t)$ plays the role of the log-price. From the dynamics of the asset price we can deduce the dynamics of $Y(t)$:

$$dY(t) = \left(r - \frac{1}{2} \text{Tr}(\Sigma(t)) \right) dt + \text{Tr}\left(\sqrt{\Sigma(t)}(dW(t)R^T + dB(t)\sqrt{1 - RR^T})\right).$$

By Feynman-Kac we have that ψ satisfies the PDE

$$\begin{cases} \frac{\partial \psi_{\gamma,t}}{\partial \tau} = \mathcal{A}_{Y,\Sigma} \psi_{\gamma,t} \\ \psi_{\gamma,t}(0) = e^{\gamma Y(t)}, \end{cases} \quad (3.533)$$

where $\mathcal{A}_{Y,\Sigma}$ is the joint infinitesimal generator of the whole stochastic process.

Proposition 3.28.2. *The multi-factor Wishart model is affine and its infinitesimal generator is given by*

$$\begin{aligned} \mathcal{A}_{Y,\Sigma} &= \underbrace{\left(r - \frac{1}{2} \text{Tr}(\Sigma) \right) \frac{\partial}{\partial y} + \frac{1}{2} \text{Tr}(\Sigma) \frac{\partial^2}{\partial y^2}}_{\mathcal{A}_Y} + \underbrace{\text{Tr}((\beta Q^T Q + 2M\Sigma)D + 2\Sigma D Q^T Q D)}_{\mathcal{A}_\Sigma} + \\ &\quad + \underbrace{2 \text{Tr}(\Sigma R Q D) \frac{\partial}{\partial y}}_{\mathcal{A}_{\langle Y, \Sigma \rangle}}. \end{aligned} \quad (3.534)$$

Proof. We have already proved the expressions for the infinitesimal generator for Y and Σ , so we only need to study the covariation term. Recalling that $(\sqrt{\Sigma})_{ij} = \sigma_{ij}$

$$\begin{aligned} d\langle \Sigma^{ij}, Y \rangle_t &= \mathbb{E}_t \left[\sum_{l,k=1}^n \sigma_t^{il} dW_{lk} Q_{kj} + \sum_{l,k=1}^n \sigma_t^{jl} dW_{lk} Q_{ki} \right] \left(\sum_{i,k,h=1}^n \sigma_t^{lk} dW_{kh} Q_{lh} \right) \\ &= \sum_{l,k,h=1}^n (\sigma_t^{il} Q_{kj} + \sigma_t^{jl} Q_{ki}) \sigma_t^{lk} R_{hk} dt \\ &= \sum_{k,h=1}^n \left[\left(\sum_{l=1}^n \sigma_t^{il} \sigma_t^{kl} \right) Q_{kj} + \left(\sum_{l=1}^n \sigma_t^{jl} \sigma_t^{hl} \right) Q_{hi} \right] R_{hk} dt \\ &= \sum_{k,h=1}^n \left[\Sigma^{ih} Q_{kj} + \Sigma^{jk} Q_{hi} \right] R_{hk} dt \end{aligned}$$

$$= \text{coefficient of } \frac{\partial^2}{\partial X_{ij} \partial y}$$

In fact, note that if in

$$2 \operatorname{Tr}(\Sigma R Q D) \frac{\partial}{\partial y} = 2 \sum_{i,j,k,h=1}^n D^{ij} \Sigma^{jh} R_{hk} Q_{ki} \frac{\partial}{\partial y}$$

we fix i and j we find the expression above. \square

Now we have all the ingredients to find the characteristic function (the Laplace transform) of the asset price.

Lesson 34 of
28/05/2020

Proposition 3.28.3. *The characteristic function of the asset price is given by*

$$\psi_{\gamma,t}(\tau) = \exp\{\operatorname{Tr}(A(\tau)\Sigma_t) + b(\tau)Y(t) + c(\tau)\} \quad (3.535)$$

where

- $A(\tau) = A_{22}(\tau)^{-1} A_{21}(\tau)$ is the solution of the Riccati matrix equation with

$$\begin{pmatrix} A_{11}(\tau) & A_{12}(\tau) \\ A_{21}(\tau) & A_{22}(\tau) \end{pmatrix} = \begin{pmatrix} M & -2Q^T Q \\ \frac{\gamma(\gamma-1)}{2} \mathbf{1}_n & -(M^T + 2\gamma R Q) \end{pmatrix}; \quad (3.536)$$

- $b(\tau) = \gamma$;
- $c(\tau) = \frac{\beta}{2} \operatorname{Tr}(\ln F(\tau) + (M^T + 2\gamma R Q)\tau) + \gamma r\tau$, where F comes from the linearization of the Riccati equation: $A = F^{-1}G$.

Proof. If we consider the Feynman-Kac PDE we get:

$$\begin{aligned} \frac{\partial \psi}{\partial \tau} &= \left(r - \frac{1}{2} - \operatorname{Tr}(\Sigma) \right) \frac{\partial \psi}{\partial y} + \frac{1}{2} \operatorname{Tr}(\Sigma) \frac{\partial^2 \psi}{\partial y^2} + \\ &\quad + \operatorname{Tr}[(\beta Q^T Q + M\Sigma + \Sigma M^T)D\psi + 2\Sigma DQ^T Q D\psi] + 2 \operatorname{Tr}(\Sigma R Q D) \frac{\partial \psi}{\partial y} \end{aligned}$$

Substituting the candidate solution (3.535) we get

$$\begin{aligned} 0 &= -\operatorname{Tr}(\dot{A}(\tau)\Sigma) - \dot{b}(\tau)Y - \dot{c}(\tau) + \operatorname{Tr}[(\beta Q^T Q + M\Sigma + \Sigma M^T)A(\tau) + \\ &\quad + 2\Sigma A(\tau)Q^T Q A(\tau) + 2\Sigma R Q A(\tau)b(\tau)] + \left(r - \frac{1}{2} - \operatorname{Tr}(\Sigma)\right) b(\tau) + \frac{1}{2} \operatorname{Tr}(\Sigma) b^2(\tau) \end{aligned}$$

with the boundary conditions

$$\begin{aligned} A(0) &= 0 \in M_n \\ b(0) &= \gamma \in \mathbb{R} \\ c(0) &= 0 \in \mathbb{R} \end{aligned}$$

By identifying the coefficients we get

$$\dot{b}(\tau) = 0 \quad \Rightarrow \quad b(\tau) \equiv \gamma \quad \forall t$$

$$\begin{cases} \dot{A}(\tau) = AM + (M^T + 2\gamma R Q)A + 2AQ^T Q A + \frac{\gamma(\gamma-1)}{2} \mathbf{1}_n \\ A(0) = 0 \end{cases}$$

$$\begin{cases} \dot{c}(\tau) = \text{Tr}(\beta Q^T Q A(\tau)) + \gamma r \\ c(0) = 0 \end{cases}$$

Then we linearize by taking $A = F^{-1}G$ and we solve the Riccati equation for A and plug the result into the equation for c . Recalling that

$$G = -\frac{1}{2} \left(\dot{F} + \dot{F}(M^T + 2\gamma RQ) \right) (Q^T Q)^{-1}$$

we get

$$\dot{c}(\tau) = -\frac{\beta}{2} \text{Tr}\left(F^{-1}\dot{F} + M^T + 2\gamma RQ\right) + \gamma r$$

and integrating we get

$$c(\tau) = \frac{\beta}{2} \text{Tr}\left(\ln F(\tau) + (M^T + 2\gamma RQ)\tau\right) + \gamma r\tau.$$

□

We now have a model for the underlying with the volatility driven by a matrix. What is the advantage with respect to a multi-factor Heston model?

3.28.5 Comparison between Wishard and $A_2(3)$ -Heston model

The $A_2(3)$ -Heston model is a Heston model with three factors, with two of them positive. Let us consider the 2-Heston specification with two positive factors:

$$\begin{aligned} dX_1 &= K_1(\theta_1 - X_1)dt + \epsilon_1 \sqrt{X_1} dW_1 \\ dX_2 &= K_2(\theta_2 - X_2)dt + \epsilon_2 \sqrt{X_2} dW_2 \end{aligned}$$

and introduce the log-asset price

$$\begin{aligned} dY(t) &= \left(r - \frac{1}{2}(X_1 + X_2) \right) dt + \rho_1 \sqrt{X_1} dW_1 + \rho_2 \sqrt{X_2} dW_2 + \\ &\quad + \sqrt{(1 - \rho_1^2)X_1 + (1 - \rho_2^2)X_2} dB \end{aligned}$$

which can be negative (conditionally gaussian). So, the starting point is a vector of three independent Brownian motions (B, W_1, W_2) such that if we consider the joint infinitesimal generator of the vector (X_1, X_2, Y) we get an affine process. Recall that the skew of the implied volatility surface is determined by the correlation

$$\begin{aligned} \text{Corr}_t(\text{Noise}(dY), \text{Noise}(\text{Vol}(S))) &= \frac{\langle Y, X_1 + X_2 \rangle}{\sqrt{\langle Y \rangle_t \langle X_1 + X_2 \rangle_t}} \\ &= \frac{\rho_1 \epsilon_1 X_1 + \rho_2 \epsilon_2 X_2}{\sqrt{X_1 + X_2} \sqrt{\epsilon_1 X_2 + \epsilon_2 X_1}} \end{aligned}$$

The skew is stochastic but it is a deterministic function of the volatility factors X_1, X_2 , that is of the volatility level. So, even if we have the “illusion” of having a stochastic skew, if we fix the volatility level there is no more possibility to manage the skew, which is not very good because we want a free variable to control the skew independently of the volatility.

Remark 3.28.3. If $X_2 = 0$ we get the classic Heston model, where

$$\text{skew} \simeq \frac{\rho_1 \epsilon_1 X_1}{\sqrt{X_1} \sqrt{\epsilon_1 X_1}} = \text{constant}$$

So in the classic 1-dimensional Heston model the skew is constant and it is directly related to the correlation. That's why the correlation basically manages the skew. However, this model is too poor if we want to take into account the variability of the skew.

Proposition 3.28.4. *In the Wishart case*

$$\text{Corr}_t(\text{Noise}(S), \text{Noise}(\text{Vol}(S))) = \frac{\text{Tr}(R)Q\Sigma(t)}{\sqrt{\text{Tr}(\Sigma(t))\text{Tr}(Q^T Q\Sigma(t))}}. \quad (3.537)$$

Proof. Let's start considering the asset noise. From the dynamics of the underlying we have

$$\begin{aligned} \frac{dS}{S} &= r dt + \text{Tr}\left[\sqrt{\Sigma} dZ\right] \\ &= r dt + \sqrt{\text{Tr}(\Sigma)} \frac{\text{Tr}\left[\sqrt{\Sigma} dZ\right]}{\sqrt{\text{Tr}(\Sigma)}} \end{aligned}$$

It can be shown that this leads to a scalar Brownian motion $b(t)$:

$$= r dt + \sqrt{\text{Tr}(\Sigma)} db(t)$$

In fact, if we denote with e_i the i -th component of the canonical basis,

$$\begin{aligned} d\langle b \rangle_t &= \frac{1}{\text{Tr}(\Sigma)} \sum_{i,j} \text{Cov}(e_i^T \sqrt{\Sigma} dZ e_i, e_j^T \sqrt{\Sigma} dZ e_j) \\ &= \frac{1}{\text{Tr}(\Sigma)} \sum_{i,j} \mathbb{E}_t[e_i^T \sqrt{\Sigma} dZ e_i e_j^T dZ^T \sqrt{\Sigma} e_j] \end{aligned}$$

Now, since Z is a matrix BM with independent components, we have that $dZ e_i e_j^T dZ^T = e_j^T e_i \mathbb{1}_n dt$:

$$= \frac{1}{\text{Tr}(\Sigma)} \sum_{i,j} e_i^T \Sigma e_j e_j^T e_i dt$$

Sum over j :

$$\begin{aligned} &= \frac{1}{\text{Tr}(\Sigma)} \sum_i e_i^T \Sigma e_i dt \\ &= \frac{\text{Tr}(\Sigma)}{\text{Tr}(\Sigma)} dt \\ &= dt \end{aligned}$$

Now let's consider the noise driving the volatility ($\text{Tr}(\Sigma(t))$):

$$\begin{aligned} d\text{Tr}(\Sigma) &= (\beta \text{Tr}(Q^T Q) + 2 \text{Tr}(M\Sigma))dt + 2 \text{Tr}\left(\sqrt{\Sigma} dW Q\right) \\ &= \dots dt + 2\sqrt{\text{Tr}(\Sigma Q^T Q)} \frac{\text{Tr}\left(\sqrt{\Sigma} dW Q\right)}{\sqrt{\text{Tr}(\Sigma(t) Q^T Q)}} \end{aligned}$$

We can introduce the scalar BM $v(t)$:

$$= \dots dt + 2\sqrt{\text{Tr}(\Sigma Q^T Q)} dv(t)$$

In fact

$$\begin{aligned} d\langle v \rangle_t &= \frac{1}{\text{Tr}(\Sigma Q^T Q)} \sum_{i,j} \mathbb{E}_t [e_i^T \sqrt{\Sigma} dW Q e_i e_j^T Q^T dW^T \sqrt{\Sigma} e_j] \\ &= \frac{1}{\text{Tr}(\Sigma Q^T Q)} \sum_{i,j} e_i^T \sqrt{\Sigma} \text{Tr}(Q e_i e_j^T Q^T) \sqrt{\Sigma} e_j dt \end{aligned}$$

Since the expression inside the trace commutes, we can write $Q e_i e_j^T Q^T = e_j^T Q^T Q e_i \in \mathbb{R}$. But the trace of a scalar is the scalar itself, so we can write:

$$= \frac{1}{\text{Tr}(\Sigma Q^T Q)} \sum_{i,j} e_i^T \Sigma e_j e_j^T Q^T Q e_i dt$$

Sum over j :

$$\begin{aligned} &= \frac{1}{\text{Tr}(\Sigma Q^T Q)} \sum_i e_i^T \Sigma Q^T Q e_i dt \\ &= \frac{\text{Tr}(\Sigma Q^T Q)}{\text{Tr}(\Sigma Q^T Q)} dt \\ &= dt \end{aligned}$$

In conclusion,

$$\text{Cov}_t(b(t), v(t)) = \mathbb{E}_t \left[\frac{\text{Tr}[\sqrt{\Sigma} dZ]}{\sqrt{\text{Tr}(\Sigma)}} \frac{\text{Tr}[\sqrt{\Sigma} dW Q]}{\sqrt{\text{Tr}(\Sigma Q^T Q)}} \right]$$

Introduce $Z = WR^T + B\sqrt{1 - RR^T}$:

$$= \sum_{i,j} \frac{\mathbb{E}_t [e_i^T \sqrt{\Sigma} dZ e_i e_j^T Q^T dW^T \sqrt{\Sigma} e_j]}{\sqrt{\text{Tr}(\Sigma) \text{Tr}(\Sigma Q^T Q)}}$$

The only part of Z which contributes is the one with W :

$$= \sum_{i,j} \frac{\mathbb{E}_t [e_i^T \sqrt{\Sigma} dW R^T e_i e_j^T Q^T dW^T \sqrt{\Sigma} e_j]}{\sqrt{\text{Tr}(\Sigma) \text{Tr}(\Sigma Q^T Q)}}$$

Now, let's exploit the BM property which says that $e_i^T dW A dW^T e_i = \text{Tr}(A) dt$:

$$\begin{aligned} &= \sum_{i,j} \frac{e_i^T \sqrt{\Sigma} \text{Tr}(R^T e_i e_j^T Q^T) \sqrt{\Sigma} e_j dt}{\sqrt{\text{Tr}(\Sigma) \text{Tr}(\Sigma Q^T Q)}} \\ &= \sum_{i,j} \frac{e_i^T \Sigma e_j e_j^T Q^T R^T e_i dt}{\sqrt{\text{Tr}(\Sigma) \text{Tr}(\Sigma Q^T Q)}} \\ &= \frac{\text{Tr}(\Sigma Q^T R^T)}{\sqrt{\text{Tr}(\Sigma) \text{Tr}(\Sigma Q^T Q)}} \end{aligned}$$

□

The crucial difference is that now in the numerator we have a linear combination involving all the elements of the Wishart process, not only the positive volatility factors (diagonal elements) but also their covariance (off-diagonal elements). So, we have a dedicated factor to manage the skew.

3.29 Wishart affine stochastic correlation model

The Wishart affine stochastic correlation model (WASC) is a multiple-asset model (there are at least two risky assets).

Let's begin with a preliminary remind. We have seen the Heston model with two risky assets. We can write their dynamics introducing two volatility factors:

$$\begin{aligned} dS_1 &= S_1(r dt + \sqrt{X_1(t)} dZ_1 + \sqrt{X_3(t)} dZ_3) \\ dS_2 &= S_2(r dt + \sqrt{X_2(t)} dZ_1 + \sqrt{X_3(t)} dZ_3) \end{aligned}$$

where X_1, X_2 and X_3 are three independent (otherwise we loose affinity, i.e. analytical tractability) CIR processes. We can also introduce the leverage effect by allowing correlations between the Brownian motions Z and the ones driving the CIR processes:

$$Z_i = \rho_i W_i + \sqrt{1 - \rho_i^2} B_i$$

where W_i is the Brownian motion driving X_i and B_i is its orthogonal part.

If we consider the (normalized) quadratic covariation, since all the Brownian motions are independent except for the one of X_3 , we get

$$\frac{d\langle S_1, S_2 \rangle}{S_1 S_2} = X_3(t) dt > 0$$

because X_3 is a CIR process. This is a strong constraint, in fact it is reasonable to expect stochastic correlations among the assets. So, if we naively try to extend the Heston model to the three factors case we end up with something which is very limited, because there is this constraint on the sign of the correlation.

This is the preliminary remark that motivates the introduction of a more sophisticated model. The only assumption that we require is that the stochastic covariance matrix evolves according the dynamics of the Wishart process:

$$d\Sigma(t) = (\beta Q^T Q + M\Sigma + \Sigma M^T)dt + \sqrt{\Sigma} dWQ + Q^T dW^T \sqrt{\Sigma} \quad (3.538)$$

with $Q \in GL_n$ in order to be invertible, $\beta \geq n - 1$ in order to have at least a weak solution and $M \in M^-$ in order to have stationarity. The assets evolve according a vector dynamics

$$dS = \text{diag}(S)(r \mathbf{1} dt + \sqrt{\Sigma(t)} dZ(t)) \quad (3.539)$$

where Z is a vector Brownian motion.

Example 3.29.1: 2-dimensional case

For $n = 2$ we write (3.539) as

$$d \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix} \left(\begin{pmatrix} r dt \\ r dt \end{pmatrix} + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \begin{pmatrix} dZ_1 \\ dZ_2 \end{pmatrix} \right)$$

Remark 3.29.1. The dynamics of each asset is driven by only one row of $\Sigma(t)$, not by the whole matrix as in the 1-asset case considered before, where

$$\frac{dS}{S} = r dt + \text{Tr}(\Sigma(t) dZ). \quad (*)$$

Remark 3.29.2. It is simple to extend the WASC model in order to consider the dynamics $(*)$ for each asset. For example, by considering the projection along a certain direction:

$$\frac{dS_i}{S_i} = r dt + \text{Tr}(A_i \sqrt{\Sigma(t)} dZ).$$

Remark 3.29.3. If we consider, in analogy with the multiple-asset Heston model, the quadratic covariation between the first two assets, we get

$$\frac{d\langle S_1, S_2 \rangle}{S_1 S_2} = \Sigma_{12} dt \geq 0$$

so we have a stochastic correlation.

Remark 3.29.4. The general way to correlate the matrix BM W and the vector BM Z is to introduce a matrix $R_k \in M_n$ for $k = 1, \dots, n$. So, in principle, there are many parameters because any element of Z can be correlated with any element of W :

$$dZ_k = \text{Tr}(R_k dW^T) + \sqrt{1 - \text{Tr}(R_k R_k^T)} dB_k(t) \quad (3.540)$$

In other words, the element R_k^{ij} plays the role of the covariation between Z_k and W_{ij} : $R_k^{ij} \cong \langle Z_k, W_{ij} \rangle$.

Lesson 35 of
29/05/2020

So, we need many parameters to describe the correlation between the vector Brownian motion driving the asset prices and the matrix Brownian motion driving the volatility. Luckily, the model has the following property.

Proposition 3.29.1. *The WASC model is affine if and only if the matrix R_k has the following form:*

$$R_k = \begin{pmatrix} 0 & \cdots & 0 \\ \rho_1 & \cdots & \rho_n \\ 0 & \cdots & 0 \end{pmatrix} \leftarrow k\text{-th row} \quad (3.541)$$

with $\rho^T \rho \leq 1$. So, we can write the Brownian motion Z as the projection along the direction given by the vector ρ :

$$dZ(t) = dW(t)\rho + \sqrt{1 - \rho^T \rho} dB(t) \quad (3.542)$$

where $Z(t)$ and $B(t)$ are vectors and $W(t)$ is a matrix (independent of B).

So, we need only n parameters to describe the correlation.

Proof. The starting point consists in taking the log-price. Adopting the notation

$$\text{vec}(\Sigma_{ii}) = \begin{pmatrix} \Sigma_{11} \\ \vdots \\ \Sigma_{nn} \end{pmatrix},$$

we can write it as

$$dY(t) = r\mathbb{1} - \frac{1}{2}\text{vec}(\Sigma_{ii})dt + \sqrt{\Sigma} dZ(t).$$

The idea is to show when $\langle Y_h, \Sigma_{ij} \rangle$ is affine in Σ (if this is true then the whole infinitesimal generator will be affine). By Itô we get

$$\begin{aligned} d\langle Y_h, \Sigma_{ij} \rangle &= \sum_{k=1}^n (\sigma_{hk} dZ_k) d\Sigma_{ij} \\ &= \sum_{k=1}^n \sigma_{hk} \text{Tr}(R_k dW(t)^T) \left(\sum_{l,m=1}^n \sigma_{im} dW_{ml} Q_{lj} + \sigma_{jm} dW_{ml} Q_{li} \right) \end{aligned}$$

$$= \sum_{k=1}^n \sigma_{hk} \sum_{p,q=1}^n R_k^{pq} dW_{pq}(t) \left(\sum_{l,m=1}^n \sigma_{im} dW_{ml} Q_{lj} + \sigma_{jm} dW_{ml} Q_{li} \right).$$

Now, let's exploit the information that we have on the matrix Brownian motion W . Since all the elements are independent we have

$$d\langle W_{pq}, W_{ml} \rangle_t = \delta_{pm} \delta_{ql} dt$$

As a consequence, we get

$$\begin{aligned} d\langle Y_h, \Sigma_{ij} \rangle &= \sum_{m,l,k=1}^n \sigma_{hk} R_k^{ml} (\sigma_{im} Q_{lj} + \sigma_{jm} Q_{li}) dt \\ &= \sum_{m,l,k=1}^n (\sigma_{hk} R_k^{ml} \sigma_{im} Q_{lj} + \sigma_{hk} R_k^{ml} \sigma_{jm} Q_{li}) dt \end{aligned}$$

We would like to recognize the product $\sigma_{ij} \sigma_{ml}$ in such a way that by summing up over all the indices it leads to Σ_{pq} . The only possibility is to put $R_k^{ml} = 0$ for all $k \neq m$ and $\forall l = 1, \dots, n$. We end up with the desired condition:

- R_k must have all the elements equal to zero except for the k -th row;
- The rows of R_k have the same elements for all k , so ρ does not depend on k .

Therefore

$$d\langle Y_h, \Sigma_{ij} \rangle = \sum_{m,l,k=1}^n (\sigma_{hk} \sigma_{im} Q_{lj} + \sigma_{hk} \sigma_{jm} Q_{li}) \rho_l dt$$

Fixing l and summing over m and k we recognize the developement of the matrix Σ_{hi} and Σ_{hj} :

$$\begin{aligned} &= \sum_{l=1}^n (\Sigma_{hi} Q_{lj} + \Sigma_{hj} Q_{li}) \rho_l dt \\ &= \left(\sum_{hi} \text{Tr}(R_j Q) + \Sigma_{hj} \text{Tr}(R_i Q) \right) dt \end{aligned}$$

We end up with a linear combination in Σ , so the model is affine. \square

3.29.1 Pricing in a multiple asset model

Given a payoff $F(Y(T), T) \in L^2(\mathcal{F}(T))$, its initial price is given by

$$\begin{aligned} \text{price}_0 &= e^{-rT} \mathbb{E}^{\mathbb{Q}}[F(Y(T), T)] \\ &= e^{-rT} \frac{1}{(2\pi)^n} \mathbb{E}^{\mathbb{Q}} \left[\int_{\mathbb{Z}} e^{-i\langle Z, Y \rangle} \left(\int_{\mathbb{R}^n} e^{i\langle Z, Y \rangle} F(y, T) dy \right) dZ \right] \\ &= e^{-rT} \frac{1}{(2\pi)^n} \int_{\mathbb{Z}} \mathbb{E}^{\mathbb{Q}} \left[e^{-i\langle Z, Y \rangle} \right] \left(\int_{\mathbb{R}} e^{i\langle Z, Y \rangle} F(y, T) dy \right) dZ \end{aligned}$$

Once again we have that the price of a contingent claim can be written as the numerical integral of a product of two elements: the first one being the characteristic function of the asset prices (or log-prices) and the second being the characteristic function of the particular payoff we are interested in. The only difference is that now we are reasoning in multiple dimensions, where we can introduce the following notation:

$$= \frac{1}{(2\pi)^n} \int_{\mathbb{Z}} \psi_{Y(0), \Sigma(0)}(-iz, 0, 0, T) \hat{F}(Z) dZ$$

where

$$\hat{F}(Z) = \int_{\mathbb{R}} e^{i\langle Z, Y \rangle} F(Y(T), T) dY(T) \quad (3.543)$$

is the Fourier transform of the payoff and

$$\psi_{Y(t), \Sigma(t)}(\gamma, \Gamma, t, \tau) = e^{-r\tau} \mathbb{E}^{\mathbb{Q}} \left[e^{\langle \gamma, Y(t+\tau) \rangle + \text{Tr}[\Gamma \Sigma(t+\tau)]} \right] \quad (3.544)$$

This is the general Fourier transform: if we want to price calls or puts we can skip the second term of the exponential whereas if we are interested in pricing volatility models we can skip the first term (in fact in (3.543) we set $\Gamma = 0$). By Feynman-Kac we can write

$$\frac{\partial \psi_{Y, \Sigma}}{\partial \tau} = (\mathcal{A}_{Y, \Sigma} - r)\psi. \quad (3.545)$$

Proposition 3.29.2. *The infinitesimal generator of the WASC model is given by*

$$\begin{aligned} \mathcal{A}_{Y, \Sigma} &= \underbrace{\text{Tr}((\beta Q^T Q + M\Sigma + \Sigma M^T)D + 2\Sigma D Q^T Q D)}_{\text{Wishart}} + \\ &\quad + \underbrace{\nabla_Y \left(r\mathbb{1} - \frac{1}{2}\text{vec}(\Sigma_{ii}) \right) + \frac{1}{2}\nabla_Y \Sigma \nabla_Y^T + 2\text{Tr}(D Q^T \rho \nabla_Y \Sigma)}_{\text{log-price } Y} \end{aligned} \quad (3.546)$$

where ρ is a column vector and ∇_Y is a row vector, so that $\rho \nabla_Y$ is a $n \times n$ matrix.

Proof. We already know the infinitesimal generator of the Wishart process, so let's look for the other two. For the vector $Y(t)$ the infinitesimal generator is given by the Itô formula:

$$\mathcal{A}_Y = \sum_{i=1}^n \left(r - \frac{1}{2}\Sigma_{ii} \right) \frac{\partial}{\partial y_i} + \frac{1}{2} \sum_{i,j=1}^n \Sigma_{ij} \frac{\partial^2}{\partial y_i \partial y_j}$$

In terms of the covariation, we can write

$$\begin{aligned} \mathcal{A}_{\langle Y, \Sigma \rangle} &= 2 \sum_{i=1}^n \text{Tr}(\Sigma R_i Q D) \frac{\partial}{\partial y_i} \\ &= 2 \text{Tr} \left[\left(\sum_{i=1}^n R_i \frac{\partial}{\partial y_i} \right) Q D \Sigma \right] \\ &= 2 \text{Tr} \left[\begin{pmatrix} \rho_1 \frac{\partial}{\partial y_1} & \cdots & \rho_n \frac{\partial}{\partial y_1} \\ \vdots & \ddots & \vdots \\ \rho_1 \frac{\partial}{\partial y_n} & \cdots & \rho_n \frac{\partial}{\partial y_n} \end{pmatrix} \right] \\ &= 2 \text{Tr} [\Sigma \nabla_y^T \rho^T Q D] \end{aligned}$$

□

Proposition 3.29.3. *The Laplace transform of the joint process is given by*

$$\psi_{Y(t), \Sigma(t)}(\gamma, \Gamma, t, \tau) = e^{-r\tau} \exp \{ \text{Tr}(A(\tau) \Sigma(t)) + \gamma^T Y(t) + c(\tau) \} \quad (3.547)$$

where

$$A(\tau) = (\Gamma A_2^1(\tau) + A_2^2(\tau))^{-1} (\Gamma A_1^1(\tau) + A_1^2(\tau)) \quad (3.548)$$

is the solution of the Riccati equation and

$$\begin{pmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{pmatrix} = \exp \left\{ \tau \begin{pmatrix} M + Q^T \rho \gamma^T & -2Q^T Q \\ \frac{1}{2} (\gamma \gamma^T - \sum_{i=1}^n Y_i e_{ii}) & -(M^T + \gamma \rho^T Q) \end{pmatrix} \right\}, \quad (3.549)$$

$$c(\tau) = -\frac{\beta}{2} \text{Tr} [\ln(\Gamma A_2^1(\tau) + A_2^2(\tau)) + \tau M^T + \tau \gamma \rho^T Q] + \tau r(\gamma^T \mathbf{1} - 1). \quad (3.550)$$

Proof. Putting the candidate into the PDE we get the Riccati equation

$$\begin{cases} \dot{A}(\tau) = AM + M^T A - \frac{1}{2} \sum_{i=1}^n \gamma_i e_{ii} + 2AQ^T QA + \frac{1}{2} \gamma \gamma^T + AQ^T \rho \gamma^T + \gamma \rho^T QA \\ A(0) = \Gamma \end{cases}$$

together with

$$\begin{cases} \dot{c}(\tau) = \text{Tr}(r \mathbf{1} \gamma^T + \beta Q^T QA) - r \\ c(0) = 0 \end{cases}$$

The linearization procedure is standard (express A as a ratio involving the derivative of the denominator, ecc.). Homework \square

Now we would like to see which are the other advantages of introducing a Wishart model instead of a multiple Heston. We have already seen that the Wishart process allows to consider stochastic correlations between the assets. Now we analyze what happens in the skew of the vanillas. If we are interested in pricing a call option on the first asset, in principle we can forget about the other assets. However we will see that in this model all the assets communicate with each other, because in the matrix σ there is a “cocktail” involving all the volatility parameters.

3.29.2 Leverage in single name vanillas

If we consider the correlation between the noise driving the first (log-)asset price and the noise driving its volatility:

$$\text{Corr}_t(\text{Noise}(Y_1), \text{Noise}(\text{Vol}(S_1))) = \frac{\langle Y_1, \Sigma_{11} \rangle_t}{\sqrt{\Sigma_{11}(t)} \sqrt{\langle \Sigma_{11} \rangle_t}} \quad (3.551)$$

$$= \frac{2\Sigma_{11} \text{Tr}(R_1 Q)}{\sqrt{\Sigma_{11}(t)} \sqrt{4\Sigma_{11}(Q_{11}^2 + Q_{21}^2)}} \quad (3.552)$$

$$= \frac{\text{Tr}(R_1 Q)}{\sqrt{Q_{11}^2 + Q_{21}^2}} \quad (3.553)$$

Recall that $R_1 = \begin{pmatrix} \rho_1 & \rho_2 \\ 0 & 0 \end{pmatrix}$:

$$= \frac{Q_{11}\rho_1 + Q_{21}\rho_2}{\sqrt{Q_{11}^2 + Q_{21}^2}} \quad (3.554)$$

So, we have a constant skew, like in the classic Heston model. This is due to the fact that we decided not to adopt the richest parametrization of the WASC model (which considers the trace for any asset).

Remark 3.29.5. If the matrix Q is diagonal

$$\text{Corr}_t(\text{Noise}(Y_1), \text{Noise}(\text{Vol}(S_1))) = \rho_1$$

$$\text{Corr}_t(\text{Noise}(Y_2), \text{Noise}(\text{Vol}(S_2))) = \rho_2$$

as in the classic Heston model. If Q is not diagonal, the skew of the implied volatility on asset 1 depends on the volatility of asset 2.

In some sense, in this framework vanillas are “basket options”, in the sense that there are dependences between assets through their volatilities.

3.29.3 Asymmetric conditional correlation effect

In the stochastic volatility framework for a single asset, if the market rises up we observe that the volatility σ decreases while if the market crushes σ increases. In the multiple assets case:

- if the market rises up, the correlation ρ between the assets decreases, in the sense that there is more diversification in the market;
- if the market crushes, the correlation between the assets increases. There is a contagion effect, in the sense that the credit risk transfers from one asset to another. This means that there will no more diversification in the market (all the assets behave more or less in the same way).

We would like to see if there exists a model that can describe this effect. In the WASC model we can find a closed form expression of the covariation between the asset Y_i and the cross correlation

$$\rho_{Y_1, Y_2}(t) = \rho_{12}(t) = \frac{\Sigma_{12}(t)}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}. \quad (3.555)$$

Proposition 3.29.4. *The covariation between the i -th asset price and the cross correlation ρ_{12} is given by*

$$d\langle Y_i, \rho_{12} \rangle = \frac{\sqrt{\Sigma_{ii}}}{\Sigma_{jj}} (1 - \rho_{12}^2) \text{Tr}(R_j Q) dt \quad (3.556)$$

(where in our case $i, j = 1, 2$).

Proof. By definition,

$$\rho_{12}\sqrt{\Sigma_{11}\Sigma_{22}} = \Sigma_{12}.$$

By differentiating, we get

$$d\Sigma_{12} = \sqrt{\Sigma_{11}\Sigma_{22}} d\rho_{12} + \rho_{12} d\left(\sqrt{\Sigma_{11}\Sigma_{22}}\right) + (\dots)dt.$$

By Itô

$$d\left(\sqrt{\Sigma_{11}\Sigma_{22}}\right) = \frac{1}{2}\sqrt{\frac{\Sigma_{22}}{\Sigma_{11}}} d\Sigma_{11} + \frac{1}{2}\sqrt{\frac{\Sigma_{11}}{\Sigma_{22}}} d\Sigma_{22} + (\dots)dt.$$

$$d\rho_{12} = \frac{1}{\sqrt{\Sigma_{11}\Sigma_{22}}} d\Sigma_{12} - \frac{1}{2}\rho_{12} \left(\frac{1}{\Sigma_{11}} d\Sigma_{11} + \frac{1}{\Sigma_{22}} d\Sigma_{22} \right) + (\dots)dt.$$

Then

$$\begin{aligned} d\langle Y_1, \rho_{12} \rangle_t &= \frac{1}{\sqrt{\Sigma_{11}\Sigma_{22}}} d\langle Y_1, \Sigma_{12} \rangle - \frac{1}{2}\rho_{12} \left(\frac{1}{\Sigma_{11}} d\langle Y_1, \Sigma_{11} \rangle + \frac{1}{\Sigma_{22}} d\langle Y_1, \Sigma_{22} \rangle \right) \\ &= \frac{1}{\sqrt{\Sigma_{11}\Sigma_{22}}} (\Sigma_{11} \text{Tr}(R_2 Q) + \Sigma_{12} \text{Tr}(R_1 Q)) dt - \\ &\quad - \rho_{12} \left(\text{Tr}(R_1 Q) + \frac{\Sigma_{12}}{\Sigma_{22}} \text{Tr}(R_2 Q) \right) dt \\ &= \text{Tr}(R_2 Q) \left(\frac{\Sigma_{11}}{\sqrt{\Sigma_{11}\Sigma_{22}}} - \rho_{12} \frac{\Sigma_{12}}{\Sigma_{22}} \right) dt \\ &= \text{Tr}(R_2 Q) \frac{\sqrt{\Sigma_{11}}}{\Sigma_{22}} (1 - \rho_{12}). \end{aligned}$$

□

We have that $(1 - \rho_{12}) \geq 0$ and $\frac{\sqrt{\Sigma_{ii}}}{\Sigma_{jj}} \geq 0$, so

$$\frac{\sqrt{\Sigma_{ii}}}{\Sigma_{jj}}(1 - \rho_{12}^2) \geq 0.$$

Conversely, the term $\text{Tr}(R_j Q)$ is related to the skew of the vanilla on the asset j , which typically is negative:

$$\text{Tr}(R_j Q) \leq 0 \quad \Rightarrow \quad d \langle Y_i, \rho_{12} \rangle \leq 0$$

The correlation of every combination of the assets with the cross correlation among the assets is typically negative. So, if we define and index as the weighted sum of the log-assets

$$\text{Index} = \sum_i \alpha_i Y_i$$

when this index decreases, due to the fact that $\langle \text{Index}, \rho_{ij} \rangle < 0$, the cross correlation ρ_{12} increases. This is the meaning of the contagion effect.

Bibliography

- [1] Tomas Bjork. *Arbitrage Theory in Continuous Time*. 3rd ed. Oxford University Press, 2009. URL: <https://EconPapers.repec.org/RePEc:oxp:obooks:9780199574742>.
- [2] John C. Hull. *Options, futures, and other derivatives*. 6. ed., Pearson internat. ed. Upper Saddle River, NJ [u.a.]: Pearson Prentice Hall, 2006. XXII, 789. ISBN: 978-0-13-197705-1. URL: http://gso.gbv.de/DB=2.1/CMD?ACT=SRCHA&SRT=YOP&IKT=1016&TRM=ppn+563580607&sourceid=fbw_bibsonomy (cit. on p. 87).
- [3] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance, Second Edition*. Chapman & Hall/CRC Financial Mathematics Series. Taylor & Francis, 1996. ISBN: 9780412718007. URL: https://books.google.it/books?id=61zI%5C_o-pIkEC (cit. on p. 28).