Study on Quantitative finance

Chi Thanh Nguyen

2020/06/24 11:50:40

Contents

Chapter 1

Introduction

This is my study note on the subject of Quantitative Finance based mainly on the book [1]. I do hope that in the very near future I can successfully move to the finance sector and work as a Quantitative Investment Researcher.

Appendix A

Brief on quantitative finance

This chapter is a very brief introduction to quantitative finance, more precisely, a collection of definitions concerning the financial markets which relate to the quantitative finance. The content of this chapter is based heavily on [1, 2].

A.1 Products and markets

A.1.1 Equities

The most basic of financial instruments is the **equity**, **stock** or **share**. This is the ownership of a small piece of a company. Shares can be bought and sold freely on a regulated stock exchanges and capital can be raised efficiently. The behavior of the stock prices is far from being predictable because the prices have a large element of randomness. Therefore, the modeling must be done in a probabilistic sense.

The value in holding stocks comes from the **dividends** and any growth in the stock's value. Dividends are lump sum payments, paid out every quarter or every six months, to the holder of stock. The amount of the dividend varies depending on the profitability of the company. When the stock is bought, it either comes with its entitlement to the next dividends (**cum**) or not (**ex**).

A stock is normally announced in a **stock-split**, e.g. a company with a stock price of $90 announces a three-for-one split which simply means that instead of holding one stock valued at $90, one holds three valued at $30 each.

A.1.2 Commodities

Commodities are usually raw products such as precious metals, oil, etc. The prices of these products are unpredictable but often show seasonal effects. Commodities are usually traded by people who have no need of the raw material. Most trading is done on the futures market, making deals to buy or sell the commodity at some time in the future. The deal is then closed out before the commodity is due to be delivered. Futures contracts are discussed later.

A.1.3 Currencies

Countries in the world use different currencies, the rate at which one currency can be exchanged for another is the **exchange rate**. This is the world of **foreign exchange**, or **Forex** or **FX** for short. Whatever the exchange rates from one currency to another, there must be consistency throughout; otherwise it is possible to make **arbitrage profits** by exploiting the mispricing.

A.1.4 Indices

For measuring how the stock market/economy is doing as whole, the stock market **indices** have been developed. A typical **index** is made up from the weighted sum of a selection or **basket** of representative stocks. The selection may be designed to represent the whole market, such as the S&P500 in the US or the FTSE100 in the UK, or a very special part of a market

A.1.5 The time value of money

The most fundamental concept in finance is the **time value of money**. $100 today is worth more than $100 in a year because of all the things we can do with $100 over the year. You can lend the money to someone who can invest it in various risky ways (e.g. banks) and give you back the money with a little bit extra, the **interest**.

Let's denote the interest rates by r and assume, for the moment, that r are constant. Two most basic types of interest are simple and **compound**. Simple interest is when the interest is based only on the initial amount, whereas compound interest is when the interest has interest also. In reality, compound interest is the only case of relevance. Compound interest has two forms, **discretely compounded** and **continuously compounded**. The difference of these two types are illustrated below.

Suppose we invest $100 in a bank at a discrete interest rate of r paid *once per annum*. After one year, the money becomes:
$$100 \times (1 + r) \tag{A.1}$$
and after n years, we have:
$$100 \times (1 + r)^n \tag{A.2}$$
Now, suppose there are m interest payments at a rate of r/m *per annum*. After one year, the money in the bank is:
$$100 \times \left(1 + \frac{r}{m}\right)^m \tag{A.3}$$
Imagine that the interest payments come at increasingly frequent intervals, but at an increasingly smaller interest rate, this will lead to a rate of interest that is paid continuously. After one year, the money in the bank is:
$$\lim_{m \to \infty} 100 \times \left(1 + \frac{r}{m}\right)^m \sim 100 \times e^r \tag{A.4}$$
And similarly, after t years, we will have the amount:
$$100 \times \left(1 + \frac{r}{m}\right)^{mt} \sim 100 \times e^{rt} \tag{A.5}$$
Equation A.5 can be derived via a differential equation. Suppose we have an amount $M(t)$ in the bank at time t. A short while later, time $t + dt$, the amount will have increased by:
$$M(t + dt) - M(t) \approx \frac{dM}{dt} dt + \cdots \tag{A.6}$$
where the rhs is a Taylor polynomial approximation. On the other hand, the interest we receive must be proportional to the amount we have M, the interest rate r and the time step dt. Therefore:
$$\frac{dM}{dt} dt = r \times M(t) dt \rightarrow \frac{dM}{dt} = r \times M(t) \tag{A.7}$$
This is the ordinary differential equation with the solution:
$$M(t) = M(0) \times e^{rt} \tag{A.8}$$
The interest rate presented above is an important factor to relate the **present value**, $M(t)$, with the **future value**, $M(T)$, of the money as follows:
$$M(t) = M(T) \times e^{-r(T-t)} \tag{A.9}$$

A.1.6 Fixed-income securities

A **fixed-income security** is an investment that provides a return in the form of **fixed periodic interest payments** and the eventual return of **principal at maturity**. Unlike **variable-income securities**, where payments change based on some underlying measure — such as **short-term interest rates** or **floating interest rates** — the payments of a fixed-income security are known in advance. **Interest rate swaps** are an exchange of a fixed rate of interest for a floating rate of interest.

There are various types of fixed income securities, such as:

- Bonds (both corporate and government) - the most common form of fixed-income securities;
- Treasury Bills;
- Money market instruments;

- Asset-backed securities.

The **index-linked bond** has been in the UK since 1981 and have provided a successful way to ensuring that income is not eroded by inflation (an inflation-proof bond). In the UK inflation is measured by the **Retail Price Index (RPI)**; in the US the inflation index is the **Consumer Price Index (CPI)**. The dynamics of the relationship between inflation and short-term interest rates is particularly interesting; the level of interest rates will affect the rate of inflation, but also interest rates are often used by central banks as a tool for keeping inflation down.

A.1.7 Forwards and futures

A **forward contract** is a private agreement between two parties giving the buyer an obligation to purchase an asset (and the seller an obligation to sell an asset) at a specified price (**the delivery price**) at a specified time (**the delivery date**) in the future. No money changes hands until the delivery date or **maturity** of the contract. The asset could be a stock, a commodity or a currency. The forward contract has a value of zero initially. As we approach maturity, the value of the contract will change from initially zero to the difference between the underlying asset and the delivery price, at maturity.

A **futures contract** is very similar to a forward contract. Futures contracts are usually traded through an exchange, which standardizes the terms of the contracts - standardization means the same thing to everyone in the market. The profit or loss from the futures position is calculated every day and the change in this value is paid from one party to the other. Thus with futures contracts there is a gradual payment of funds from initiation until maturity. Because the change in value is settled on a daily basis, the value of a futures contract at any time during its life is zero. The different between the forward contracts and the future contracts can be found here or here. A series of brief lessons for forward and future contracts can be found here.

There are two kinds of forward- and future-contract participants: hedgers and speculators. Hedgers do not usually seek a profit but rather seek to stabilize the revenues or costs of their business operations; therefore hedging is avoidance of risk. Their gains or losses are usually offset to some degree by a corresponding loss or gain in the market for the underlying asset. Speculators are usually not interested in taking possession of the underlying assets. They essentially place bets on which way prices will go.

Forward contracts tend to attract more hedgers than speculators. Speculators are often blamed for big price swings, but they also provide a lot of liquidity to the futures market. Look at section 1.9.1 in [1] for a hedged portfolio of asset and forward and the **no-arbitrage** principle.

Consider a forward contract that obliges us to hand over an amount $\$F$ at time T to receive the underlying asset. Today's date is t and the price of the asset is currently $\$S(t)$ - the **spot price**. At maturity, we will hand over the amount $\$F$ and receive the asset then worth $\$S(T)$.

Assume that we simultaneously sell the asset when we enter into the forward contract - **going short** when we sell something we don't own. Now, we have an amount of $S(t)$ in cash due to the sale of the asset, a forward contract and a short asset position. Our net position now is zero. Then, we put the cash in the bank to receive interest.

At maturity, we hand over the amount F and receive the asset of $S(T)$ - this cancels our short asset position. We are also left with *guaranteed* $-F$ in cash and the bank account with the initial investment of $S(t)$ with added interest. Our net position at maturity is then:

$$S(t) \times e^{r(T-t)} - F \tag{A.10}$$

Because we began with a portfolio worth zero and we end up with a predictable amount, that amount should also be zero (following the no-arbitrage principle). We can conclude that:

$$F = S(t) \times e^{r(T-t)} \tag{A.11}$$

This is the relationship between the spot price and the forward price. It is a linear relationship, the forward price is proportional to the spot price. If this relationship is violated then there will be an arbitrage opportunity. The standard economic argument then says that investors will act quickly to exploit the opportunity, and in the process prices will adjust to eliminate it.

Futures are usually traded publicly through an exchange which means they are very liquid instruments and have lots of rules and regulations surrounding them. Here are a few concepts related

futures contracts.

- Available assets: a futures contract will specify the asset which is being invested in, the contract also specifies how many of each asset must be delivered, and the quantity will depend on the market. As an asset, a natural **commodity** is particularly interesting because of non-uniformity in the type and quality of the asset to be delivered.

- Delivery and settlement: the futures contract will specify when the asset is to be delivered and settlement is the act of consummating the contract. The contract can be settled in three ways: (1) **closeout**, (2) **physical delivery** and (3) **cash settlement**. Most futures contracts are closed out before delivery, with the trader taking the opposite position before maturity.

- Margin: changes in the value of futures contracts are settled each day - **marking to market**. To reduce the likelihood of one party defaulting, being unable or unwilling to pay up, the exchanges insist on traders depositing a sum of money to cover changes in the value of their positions. This money is deposited in a **margin account**. Margin comes in two forms, the **initial margin** and the **maintenance margin**. The initial margin is the amount deposited at the initiation of the contract. The total amount held as margin must stay above a prescribed maintenance margin. If it ever falls below this level then more money (or equivalent in bonds, stocks, etc.) must be deposited.

- Commodity futures: futures on commodities don't necessarily obey the no-arbitrage law that led to the asset/future price relationship because of the messy topic of storage. The holder of the futures contract must compensate the holder of the commodity for his storage costs. This can be expressed in percentage terms by an adjustment s to the risk-free rate of interest. On the other hands, most people actually holding the commodity are benefiting from it in some way. And the benefit from holding the commodity is commonly measured in terms of the convenience yield c:

$$F = S(t) \times e^{(r+s-c)(T-t)} \tag{A.12}$$

Depending on the storage cost and the convenience yield, the price varies. When

$$F < S(t) \times e^{r(T-t)} \tag{A.13}$$

the market is said to be **backwardation**; when

$$F > S(t) \times e^{r(T-t)} \tag{A.14}$$

the market is said to be **contango**.

- FX futures: there are no problems associated with storage when the asset is a currency, we only need to modify the no-arbitrage result to allow for interest received on the foreign currency r_f. Then we have:

$$F = S(t) \times e^{(r-r_f)(T-t)} \tag{A.15}$$

- Index futures: futures contracts on stock indices are settled in cash. Again, there are no storage problems, but now we have dividends, which play a role similar to that of a foreign interest rate on FX futures, to contend by the dividend yield q.

$$F = S(t) \times e^{(r-q)(T-t)} \tag{A.16}$$

Here are the summary of a few loose definitions for the important financial terminologies that have been presented above.

- **Premium**: the amount paid for the contract initially.

- **Underlying (asset)** the financial instrument on which the option value depends denoted by S. The option payoff is defined as some function of the underlying asset at expiry.

- **Strike (price)** or **exercise price**: the amount for which the underlying can be bought (call) or sold (put) denoted by E. This definition only really applies to the simple calls and puts.

- **Expiration (date)** or **expiry (date)**: date on which the option can be exercised or date on which the option ceases to exist or give the holder any rights denoted by T.

- **Intrinsic value**: the payoff that would be received if the underlying is at its current level when the option expires.

- **Time value**: any value that the option has above its intrinsic value since the option value is generally different from the intrinsic value.

- **In the money**: an option with positive intrinsic value. A call option when the asset price is above the strike, a put option when the asset price is below the strike.

- **Out of the money**: an option with no intrinsic value, only time value. A call option when the asset price is below the strike, a put option when the asset price is above the strike.

- **At the money**: a call or put with a strike that is close to the current asset level.

- **Long position**: a positive amount of a quantity, or a positive exposure to a quantity.

- **Short position**: a negative amount of a quantity, or a negative exposure to a quantity. Many assets can be sold short, with some constraints on the length of time before they must be bought back. More explanation for the long and short position can be found here.

A.2 Derivatives

A.2.1 Options

In the future or forward contracts, the holder is obliged to trade at the maturity of the contract. The **option** gives the holder the *right* to trade in the future at a previously agreed price but takes away the obligation, e.g. if the stock falls, we don't have to buy it after all. A **call option** is the right to buy a particular asset for an agreed amount (the **exercise price** or **strike price**) at a specified time in the future (the **expiry** or **expiration date**). The stock on which the option is based is known as the **underlying asset**. We would exercise the option at expiry if the stock is above the strike and not if it is below. If we use S to mean the stock price and E the strike then at expiry the option is worth:

$$\max(S - E, 0) \tag{A.17}$$

This function of the underlying asset is called the **payoff function**.

A **put option** is the right to sell a particular asset for an agreed amount at a specified time in the future. The holder of a put option wants the stock price to fall so that he can sell the asset for more than it is worth. The payoff function for a put option is:

$$\max(E - S, 0) \tag{A.18}$$

And now the option is only exercised if the stock falls below the strike price.

Calls and puts have a non-linear dependence on the underlying asset. The randomness in the underlying asset and the curvature of the option value with respect to the asset are intimately related. Other terms used to describe contracts with some dependence on a more fundamental asset are **derivatives** or **contingent claims**.

An option's value at expiry is normally visualized in a **payoff diagram** in which the value of an option at expiry is a function of the underlying, or in a **profit diagram** in which the payoff subtracted by the premium originally paid for the call option is a function of the underlying. More explanation on the loss/profit diagram can be found here.

The **writer** of an option is the person who promises to deliver the underlying asset, if the option is a call, or buy it, if the option is a put. The writer is the person who receives the premium. Writing options is very risky. The downside of buying an option is just the initial premium, the upside may be unlimited. The upside of writing an option is limited, but the downside could be huge. To cover the risk of default in the event of an unfavorable outcome, the **clearing houses** that register and settle options insist on the deposit of a margin by the writers of options.

Most of the simpler options contracts are bought and sold through **exchanges**. These exchanges make it simpler and more efficient to match buyers with sellers. Part of this simplification involves the conventions about features of the contracts. Having standardized contracts traded through an exchange promotes liquidity of the instruments. Some options are an agreement between two parties, often brought together by an intermediary. These agreements can be very flexible and the contract details do not need to satisfy any conventions. Such contracts are known as **over the counter** or **OTC** contracts.

A.2.2   The value of the option

The *fair value* of the option contracts V depends on the underlying asset price S and time t, $V(S,t)$. At the expiry date $t = T$, the function V is the payoff function, e.g. for the call option:

$$V(S,T) = \max(S - E, 0) \tag{A.19}$$

Factors influencing the derivative prices consists of **variables** (e.g. the value of the underlying asset S and the time to expiry t) and **parameters** (e.g. the interest rate and the strike price). There is one important parameter which has a major impact on the option value. That parameter is the **volatility** which is a measure of the amount of fluctuation in the asset price, a measure of the randomness. Volatility is a particularly interesting parameter because it is so hard to estimate since it never stays constant and is unpredictable.

A.2.3   Speculation and gearing

If you buy a far out-of-the-money option it may not cost very much, especially if there is not very long until expiry. If the option expires worthless, then you also have not lost very much. However, if there is a dramatic move in the underlying, so that the option expires in the money, you may make a large profit relative to the amount of the investment. This is an example of **gearing** or **leverage**. The out-of-the-money option has a high gearing, a possible high payoff for a small investment. The downside of this leverage is that the call option is more likely than not to expire completely worthless and you will lose all of your investment. Further reading about gearing and leverage can be found in here, here and here

Due to the unpredictable adverse price movements in an asset, the writer of the option normally buy other contracts to offset his risk. This offsetting of risk by buying other related contracts is called **hedging**. Hedging is the practice of taking a position in one market to offset and balance against the risk adopted by assuming a position in a contrary or opposing market or investment. Further reading about hedging can be found in here and here.

A.2.4   Option strategies

Single-option

Calls and puts are the two simplest forms of option, they are often referred to as **vanilla**. Increasingly important are the **binary** or **digital options**. Binary options have an expiry date and/or time. At the time of expiry, the price of the underlying asset must be on the correct side of the strike price (based on the trade taken) for the trader to make a profit. Further explanations on the binary options can be found here.

Three other popular options in practice are:

- **American options** which can be exercised at any time up to the maturity;

- **European options** which can be exercised only on the expiration date itself; and,

- **Bermudan options** which allow exercise on specified dates or in specified periods.

Put-Call parity

Imagine that today at t you buy one European call option with a strike of E and an expiry of T and that you write a European put option with the same strike and expiry, you now have a pair of long-call and short-put. The payoff for the portfolio of the two options is the sum of the individual payoffs:

$$\max(S(T) - E, 0) - \max(E - S(T), 0) = S(T) - E \tag{A.20}$$

where $S(T)$ is the value of the underlying asset at time T, today it cost you $S(t)$.

The cash-flow of the portfolio is independent of the future behaviour of the stock and is modelled as:

$$C - P = S - Ee^{-r(T-t)} \tag{A.21}$$

where C and P are today's values of the call and the put respectively. This relationship holds at any time up to expiry and is known as the **put-call parity**. If this relationship did not hold then there would be riskless arbitrage opportunities.

Multiple-option

In a portfolio, normally several different options are combined to maximize the profit or minimize the risk. There are numerous combinations in practice, a few simple one can be listed as:

- Spread
- Bull spread
- Bear spread
- Condor spread
- Calendar spread
- Credit spread
- Debit spread
- Straddle and Strangle and their differences
- Risk reversal

Miscellaneous

- **Warrants**: a contract that is very similar to an option is a warrant.Warrants are call options issued by a company on its own equity. The main differences between traded options and warrants are the timescales involved, warrants usually have a longer lifespan, and on exercise the company issues new stock to the warrant holder. On exercise, the holder of a traded option receives stock that has already been issued.

- **Convertible bonds**: convertible bonds (or **CBs**) have features of both bonds and warrants. They pay a stream of coupons with a final repayment of principal at maturity, but they can be converted into the underlying stock before expiry. Because the CB can be converted into the asset, its value has to be at least the value of the asset. This makes CBs similar to American options

- **Over-the-Counter (OTC)** options: Not all options are traded on an exchange. Some, known as over the counter or OTC options, are sold privately from one counter-party to another.

A.3 The binomial model

A.3.1 Introduction

The binomial model is the simplest model for asset prices that can be used for pricing derivatives and to explain a couple of very important financial concepts. The original binomial concept is due to Cox

et al. [3]. More details description of the model can be found in [1, 2, 4]. Hull [2] also describes its use in the fixed-income world. Mathematical treatment of the binomial model is briefly presented in Section B.6. This section only summaries the main idea of the binomial model based on the illustration in Chapter 3 of [1].

The model is used for the behavior of an equity, e.g. a stock and to illustrate how to value options. The starting point of the model includes:

- We will have a stock currently worth \$100.
- The stock can either rise or fall by a known amount at a known probability between today and tomorrow; a rise to \$101 with a probability of 0.6 and a fall to \$99 with a probability of 0.4.
- Interest rates are zero.

We have the call option on the stock with a strike of \$100 and expires tomorrow. The question is what the option is worth today ? Sadly, not **0.6** as you may calculate with the rise/fall amounts and the corresponding probability, but **0.5**.

To understand the option value of 0.5, we must construct a portfolio consisting of one option and short $\frac{1}{2}$ of the underlying stock. By doing so, the portfolio takes the value of $-\frac{99}{2}$ regardless of whether the stock rises or falls. This portfolio is a perfectly **risk-free portfolio**. As the portfolio is worth $-\frac{99}{2}$ tomorrow and interest rates are zero, it must also be wroth $-\frac{99}{2}$ today. This is an example of **no arbitrage**.

In the above example, the value of an option does not depend on the probability of the stock rising or falling. This is equivalent to saying that the stock growth rate is irrelevant for option pricing. This is because we have **hedged** *the option with the stock* hence we do not care whether the stock rises or falls.

The quantity of stock that much be sold for hedging is $\frac{1}{2}$. This is called as **Delta-hedging**, Δ, which means choosing a proportion such that the portfolio value does not depend on the direction of the stock. In general:

$$\Delta = \frac{\text{range of option payoffs}}{\text{range of stock prices}} = \frac{1-0}{101-99} = \frac{1}{2} \tag{A.22}$$

which actually comes from the equality:

$$1 - \Delta \times 101 = 0 - \Delta \times 99 \tag{A.23}$$

Δ can represent the sensitivity of the option to changes in the stock.

How does the option value change if the interest rate is non-zero? We still delta-hedge as before to construct a risk-free portfolio, then we use exactly the same delta to present value that back in time by multiplying by a discount factor.

Assume that $r = 0.1$ then the discount factor for going back one day is:

$$\frac{1}{1 + 0.1/252} = 0.9996 \tag{A.24}$$

The portfolio value today must be the *present value* of the portfolio value tomorrow:

$$? - 0.5 \times 100 = 0 - 0.5 \times 99 \times 0.9996 \tag{A.25}$$

so that $? = 0.51963$.

The **theoretical price** derived above should be identical to the **market price** otherwise there is risk-free money to be made. If the option costs less than 0.5 simply buy it and hedge to make a profit. If it is worth more than 0.5 in the market then sell it and hedge, and make a guaranteed profit. In reality, it is possible for arbitrage opportunities to exist, and to exist for a long time. The market also really need there to be some practical mechanism for arbitrage to be 'removed'. Supply and demand should act to make the option price converge to the 0.5.

A.3.2 Risk-neutrality

In the **real world**, statistical skills have been used to estimate the future possible stock prices (99 and 101) and the probabilities of reaching them (0.4 and 0.6). Some properties of the real world can be listed as (1) we know about delta hedging and risk elimination; (2) we are sensitive to risk and expect greater return for taking risk; and (3) only the two stock prices matter for option pricing, not the probabilities.

People often refer to the **risk-neutral** world in which people don't care about risk. The risk-neutral world has the following characteristics (1) we don't care about risk, and don't expect any extra return for taking unnecessary risk; (2) we don't ever need statistics for estimating probabilities of events happening; and (3) we believe that everything is priced using **simple expectations**.

Imagine that we are in the risk-neutral world and:

- we have a stock currently worth $100;
- the stock can either rise to $101 or fall to $99.

Using simple expectation and the symmetry of rising/falling, the **risk-neutral probability** p' can be calculated from:

$$p' \times 101 + (1 - p') \times 99 = 100 \tag{A.26}$$

hence $p' = 0.5$. Please note that p' is only a mathematical term, it is not real since it is derived based on a fundamentally wrong assumption that simple expectations are used for pricing, hence no allowance has been made for risk.

If we just continue to value the call option based on p' and the simple expectation we have:

$$? = p' \times 1 + (1 - p') \times 0 = 0.5 \tag{A.27}$$

This is called the **risk-neutral expectation** and it is correct. So actually two times of wrongly-using 'simple expectation' cancel each other out.

When interest rates are non-zero we must perform exactly the same operations but whenever we equate values at different times we must allow for present valuing. For example, with $r = 0.1$, we calculate the risk-neutral probability from:

$$\frac{1}{1 + 0.1/252} \left(p' \times 101 + (1 - p') \times 99 \right) = 100 \tag{A.28}$$

in which the first term is the discount factor as in the previous example and equal 0.9996; so $p' = 0.51984$. The expected option payoff is now:

$$0.51984 \times 1 + (1 - 0.51984) \times 0 = 0.51984 \tag{A.29}$$

and the present value of the option is:

$$0.9996 \times 0.51984 = 0.51963 \tag{A.30}$$

and this must be the option value.

It can be seen that the option values derived in the risk-neutral world are correct answers, or in the other world, in the risk-neutral world they have exactly the same price for the option (but for different reasons). In reality, risk-neutral pricing is a very powerful technique. Just remember that the risk-neutral probability is not real, it is a mathematical term only. The real probability of the stock price was always in our example 0.6, however we don't use it now.

A.3.3 The binomial tree

The binomial model allows the stock to move up or down a prescribed amount over the next time step. If the stock starts out with value S then it will take either the value S^+ or S^- after the next time step. We can extend the random walk to the further steps, i.e. after two time steps the asset will be at either S^{++} if there were two up moves, S^{+-} if an up was followed by a down or vice versa, or S^{--} if there were two consecutive down moves. One can imagine extending this random walk out all the way until expiry. The resulting structure looks like Figure A.1 where the nodes represent the values taken by the asset. This structure is called the **binomial tree**.

It can be seen that the top and bottom branches of the tree at expiry can only be reached by one path each, either all up or all down moves, whereas there will be several paths possible for each of the intermediate values at expiry. The probability of reaching a particular node in the binomial tree depends on the number of distinct paths to that node and the probabilities of the up and down moves. The binomial tree therefore contains within it an approximation to the probability density function for the lognormal random walk.
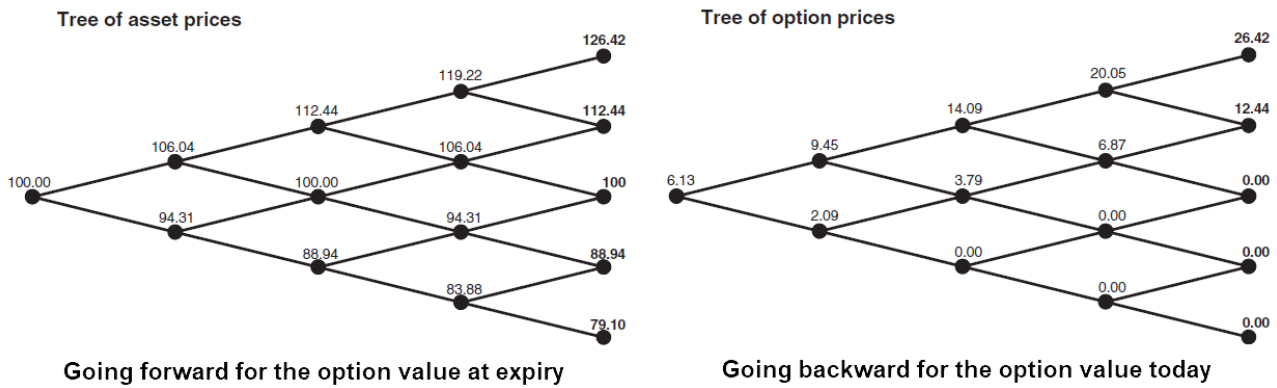
Figure A.1: The binomial tree

The binomial tree also illustrate how to fully calculate the asset prices and the option prices. By going forward, one can find the option value at expiry (V^+, V^-) at time T. As we know these values, we can find the option values one step further back in time. And by going backward to the root which is the current time and asset value, we can find the option value today. Programming of this process is given in Section B.6

The aforementioned binomial setting is for European-style exercise. The algorithm of American-style exercise is identical with one exception in the formula for V. We must ensure that there are no arbitrage opportunities at any of the nodes, i.e. the option value must be higher than the payoff value at any time.

A.3.4 Miscellaneous

The greeks are defined as derivatives of the option value with respect to various variables and parameters. These greeks will later be very important when we talk about risk management. It is important to distinguish whether the differentiation is with respect to a variable or a parameter (it could, of course, be with respect to both). If the differentiation is only with respect to the asset price and/or time then there is sufficient information in our binomial tree to estimate the derivative. It may not be an accurate estimate, but it will be an estimate. The option's delta, gamma and theta, defined below can all be estimated from the tree. On the other hand, if you want to examine the sensitivity of the option with respect to one of the parameters, then you must perform another binomial calculation.

The binomial model can also lead us to the famous Black–Scholes equation. The binomial model is a discrete-time model whereas Black–Scholes is in continuous time.

The binomial method is just a simple version of an explicit finite-difference scheme. Finite-difference methods are far more flexible and can, in many ways, incorporate dividends, allow Bermudan exercise, value path-dependent contracts and price contracts depending on other stochastic variables such as interest rates. The binomial can also do these but in a more complicated way.

Further remarks on risk neutrality by Paul Wilmott are given here:

- Hedging is used to eliminate risk.

- In simple models, hedging can be used to eliminate all risk from an option position.

- As well as eliminating risk, hedging removes dependence of an option value on the direction of an asset.

- If we don't care whether the asset price rises or falls, we shouldn't care about the probability of the rise or fall.

- *The risk-neutral random walk is one that has the same volatility as the real asset random walk but a drift rate that is the same as the risk-free interest rate and not the real drift rate.*[1]

- *The punchline is that the option value is the present value of the option payoff under a risk-neutral random walk.*

---

[1]Italic, here and down, means I still don't understand what Wilmott is talking about :).

A.4 Random behavior of assets

A.4.1 Jensen's inequality

Jensen's inequality illustrates the importance of randomness in option theory. Let's look into the example below.

> The stock price is $100 today. In one year's time, it could be $50 or 150 with both equally likely. How can we value an option on this stock, e.g. a call option with a strike of 100 expiring in one year?
>
> With those two possible scenarios we could say that we expect the stock price to be at $100 in one year, this being the average of the possible future values. The payoff for the call option would then be 0, since it is exactly at the money. And the present value of this is zero.
>
> Alternatively we could look at the two possible payoffs and then calculate that expectation. If the stock falls to 50 then the payoff is zero, if it rises to 150 then the payoff is 50. The average payoff is therefore 25, which we could present value to give us some idea of the option's value.
>
> It turns out that the second calculation is closer to what we do in practice to value options although we know that the real probabilities don't come into the calculation. But that calculation illustrates another point of great importance, that the order in which we do the payoff calculation and the expectation matters, i.e.:
>
> $$\text{Payoff} \left(\text{Expected} \left[\text{Stock price} \right] \right) = 0 \tag{A.31}$$
>
> $$\text{Expected} \left[\text{Payoff} \left(\text{Stock price} \right) \right] = 0 \tag{A.32}$$
>
> This is an example of Jensen's inequality.
>
> We can estimate how much the left-hand side is larger than the right-hand side. If we have a convex function $f(S)$ - the payoff function for a call of a random variable S - the stock price then:
>
> $$E\left[f\left(S \right) \right] \geq f\left(E\left[S \right] \right) \tag{A.33}$$
>
> As S is a random variable, let's assume that:
>
> $$S = E\left[S \right] + \epsilon = \bar{S} + \epsilon \tag{A.34}$$
>
> in which $\bar{S}$ is the mean of S and hence $E\left[\epsilon \right] = 0$.
>
> Now, we use the Taylor series approximation:
>
> $$E\left[f\left(S \right) \right] = E\left[f\left(\bar{S} + \epsilon \right) \right] = E\left[f\left(\bar{S} \right) + \epsilon f'\left(\bar{S} \right) + \frac{1}{2}\epsilon^2 f''\left(\bar{S} \right) \right] \tag{A.35}$$
>
> $$\approx f\left(\bar{S} \right) + \frac{1}{2}f''\left(\bar{S} \right) E\left[\epsilon^2 \right] \tag{A.36}$$
>
> $$= f[E\left(S \right)] + \frac{1}{2}f''\left(E\left[S \right] \right) E\left[\epsilon^2 \right] \tag{A.37}$$
>
> $E\left[\epsilon^2 \right]$ is the randomness in the underlying, and it variance. Modeling randomness is the key to modeling options.

A.4.2 Returns

This part is based on Chapter 4 of Wilmott [1]. When we invest in something, we would expect to have a comfortable return on our investment. By return, we mean the percentage growth in the value of an asset, together with accumulated dividends, over some period:

$$\text{Return} = \frac{\text{change in value of the asset} + \text{accumulated cash flows}}{\text{original value of the asset}} \tag{A.38}$$

When we model assets, it is the return that we should concentrate on. Part of the business of estimating returns for each asset is to estimate how much unpredictability there is in the asset value. It turns out that randomness plays a large part in financial markets [1].

In modeling the returns, the following concepts, terms and processes are extremely important:

- The daily returns look very much like 'noise' and is modeled as such. You, perhaps, have to wait months before you can spot the trend.

- The empirical returns are close enough to a random variable following the NORMAL distribution function with a non-zero mean and a non-zero standard deviation.

- The asset return model is similar to a model for a **random walk** of the asset price.

$$R_i = \frac{S_{i+1} - S_i}{S_i} = \mu \, \delta t + \sigma \, \phi \, \delta t^{1/2} \tag{A.39}$$

or:

$$S_{i+1} = S_i \times \left(1 + \mu \, \delta t + \sigma \, \phi \, \delta t^{1/2}\right) \tag{A.40}$$

- The parameter μ is called the drift rate, the expected return of the growth rate of the asset. In the classical option pricing theory, the drift plays almost no role.

- The parameter σ is called the volatility of the asset. The volatility is the most important and elusive quantity in the theory of derivatives. It is highly unlikely that volatility is constant for any given asset. Changing economic circumstances, seasonality, etc. will inevitably result in volatility changing with time.

- Because of their scaling with time, the drift and volatility have different effects on the asset path. The drift is not apparent over short timescales for which the volatility dominates. Over long timescales, for instance decades, the drift becomes important.

- The random walk model still have a discrete time step, this is a first step to develop a model in continuous-time. By changing from the normal distributions and discrete time to the **Wiener process**, the asset price model in the continuous-time limit using the Wiener process notation can be written as:

$$dS = \mu \, S \, dt + \sigma \, S \, dX \tag{A.41}$$

This stochastic differential equation is a continuous-time model of an asset price. It is the most widely accepted model for equities, currencies, commodities and indices and the foundation of finance theory.

They will be described in more detailed in Section B.1.

Appendix B

Basic financial simulation I

[1, 4, 5, 6, 7, 8]

B.1   Returns modeling

Return is defined as the percentage growth in the value of an asset, together with accumulated dividends, over some period [1]:

$$\text{Return} = \frac{\text{change in value of the asset} + \text{accumulated cash flows}}{\text{original value of the asset}} \tag{B.1}$$

Denoting the asset value on the i-th day by S_i then the return from day i to day $i+1$ is given by:

$$R_i = \frac{S_{i+1} - S_i}{S_i} \tag{B.2}$$

The dividends are ignored here, they are easily allowed for, especially since they only get paid two or four times a year typically.

The mean of the returns distribution is:

$$\text{mean} = \bar{R} = \frac{1}{M} \sum_{i=1}^{M} R_i \tag{B.3}$$

and the sample standard deviation is:

$$\text{sd} = \sqrt{\frac{1}{M-1} \sum_{i=1}^{M} \left(R_i - \bar{R}\right)^2} \tag{B.4}$$

where M is the number of returns in the sample (one fewer than the number of asset prices).

Figures B.1(a,b) show the quoted price of a stock and its daily return in the duration considered. The daily return looks very much like 'noise' and is modeled as such. Corresponding to the data in the example, the mean is 0.002916 and the standard deviation is 0.024521. Notice how the mean daily return is much smaller than the standard deviation. This is very typical of financial quantities over short timescales. On a day-by-day basis you will tend to see the noise in the stock price, and will have to wait months perhaps before you can spot the trend.

The frequency distribution of this time series of daily returns is plotted in Figure B.1(c), it has been scaled and translated to give it a mean zero, a standard deviation of one and an area under the curve of one. This distribution is fairly close to the standardized normal distribution of

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\phi^2} \tag{B.5}$$

where ϕ is a standardized normal variable.

Then, it is possible to assume that the returns can be modeled as a random variable drawn from a normal distribution with a know, constant, non-zero mean and a known, constant, non-zero standard deviation:

$$R_i = \frac{S_{i+1} - S_i}{S_i} = \text{mean} + \text{sd} \times \phi \qquad (B.6)$$

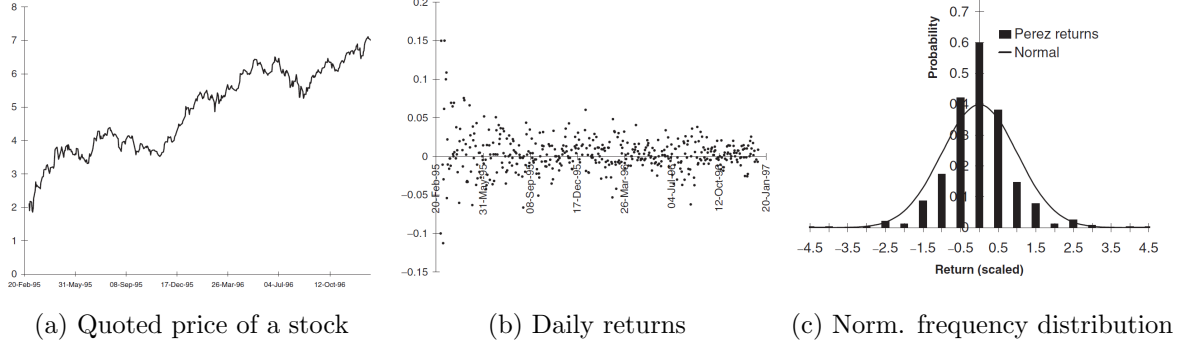| (a) Quoted price of a stock | (b) Daily returns | (c) Norm. frequency distribution |

Figure B.1: Example of a daily return

B.1.1   Discrete-time model

In reality, the assets may be measured in different time scale, i.e. hourly, daily, weekly... hence we need to take into account the time step in modeling returns. Call the time step δt. Following Equation B.3, it is reasonable to assume that the return scales with the size of the time step as follows:

$$\text{mean} = \mu\, \delta t \qquad (B.7)$$

for some μ which is assumed to be constant now. Similarly, following Equation B.4, the standard deviation of the asset return over a time step δt must be $O\left(\delta t^{1/2}\right)$:

$$\text{sd} = \sigma\, \delta t^{1/2} \qquad (B.8)$$

where σ is some parameter measuring the amount of randomness, the larger this parameter the more uncertain is the return. And the asset return model becomes:

$$R_i = \frac{S_{i+1} - S_i}{S_i} = \mu\, \delta t + \sigma\, \phi\, \delta t^{1/2} \qquad (B.9)$$

or:

$$S_{i+1} - S_i = \mu\, S_i\, \delta t + \sigma\, S_i\, \phi\, \delta t^{1/2} \qquad (B.10)$$

The left-hand side of this equation is the change in the asset price from time step i to time step $i+1$. The right-hand side is the 'model'. We can think of this equation as a model for a **random walk** of the asset price.

The parameter μ is called the **drift rate**, the **expected return** or the **growth rate** of the asset. It can be estimated by:

$$\mu = \frac{1}{M\delta t} \sum_{i=1}^{M} R_i \qquad (B.11)$$

Statistically it is very hard to measure since the mean scales with the usually small parameter δt. The unit of time that is usually used is the year, in which case μ is quoted as an annualized growth rate. In the classical option pricing theory the drift plays almost no role. So even though it is hard to measure, this doesn't matter too much [1].

The parameter σ is called the **volatility** of the asset. It can be estimated by:

$$\sigma = \sqrt{\frac{1}{(M-1)\,\delta t} \sum_{i=1}^{M} \left(R_i - \bar{R}\right)^2} \qquad (B.12)$$

Again, this is almost always quoted in annualized terms. The volatility is the most important and elusive quantity in the theory of derivatives. If δt is sufficiently small, the mean return $\bar{R}$ term can be ignored; and in this case, the following equation:

$$\sigma = \sqrt{\frac{1}{(M-1)\,\delta t} \sum_{i=1}^{M} \left(\log S(t_i) - \log S(t_{i-1})\right)^2} \tag{B.13}$$

can be used where $S(t_i)$ is the closing price on day t_i.

Because of their scaling with time, the drift and volatility have different effects on the asset path. The drift is not apparent over short timescales for which the volatility dominates. Over long timescales, for instance decades, the drift becomes important.

It is highly unlikely that volatility is constant for any given asset. Changing economic circumstances, seasonality, etc. will inevitably result in volatility changing with time. If you want to know the volatility today you must use some past data in the calculation. Unfortunately, this means that there is no guarantee that you are actually calculating today's volatility. Typically you would use daily closing prices to work out daily returns and then use the past 10, 30, 100, ... daily returns in the formula above.
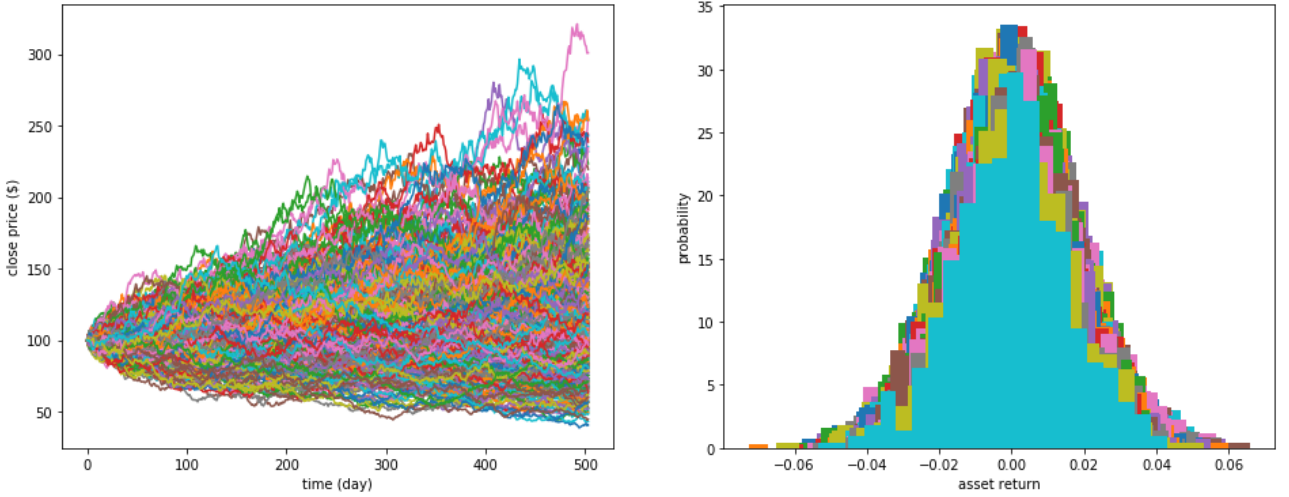

Figure B.2: Random walk model for an asset and corresponding return

Figure B.2 reproduces the random walk model presented in Section 4.8 [1] for 500 evolution scenarios of an asset with the initial value of \$100, the drift rate μ of 0.10, the volatility σ of 0.25, a daily quote in two-year duration. The corresponding asset returns of those 500 scenarios are also calculated, it can be seen that the asset returns closely follow the normal distribution as expected. Python code for this basic model can be found here.

B.1.2 Continuous-time model

The random walk model described in the previous section has a discrete time step; in this section, a brief introduction to the continuous-time limit of the random walk model is presented as a starting point for the world of stochastic modeling and Wiener processes. These topics will be presented in more detailed in Section B.2.

We rewrite the asset return model here for convenience:

$$S_{i+1} - S_i = \mu\,S_i\,\delta t + \sigma\,S_i\,\phi\,\delta t^{1/2} \tag{B.14}$$

First we introduce a quantity dS representing *the change in* the asset price in *continuous time*, i.e. we go to the limit $\delta t = 0$. So the left-hand side becomes dS. The first δt in the right-hand side can also translate into dt.

For the second term of the right-hand side, however, we can not straightforwardly translate $\delta t^{1/2}$ into $dt^{1/2}$ *otherwise any random $dt^{1/2}$ will dominate any deterministic dt term*. It turns out, we will see in Section B.2, that because the variance of the random term is $O(\delta t)$ we can make a sensible continuous-time limit of our discrete-time model. The term $\phi \delta t^{1/2}$ is rewritten as dX which can be thought as a random variable drawn from a normal distribution with mean zero and variance dt, i.e. $E[dX] = 0$ and $E[dX^2] = dt$. This is called a Wiener process. It should be noted that *this is not exactly what it is but it is close enough*, the important point is that we can build up a continuous-time theory using Wiener processes instead of normal distribution and discrete time.

The asset price model in the continuous-time limit using the Wiener process notation can be written as:

$$dS = \mu \, S \, dt + \sigma \, S \, dX \tag{B.15}$$

This is our first stochastic differential equation. It is a continuous-time model of an asset price. It is the most widely accepted model for equities, currencies, commodities and indices, and the foundation of so much finance theory. This equation will be derived following different approaches in next sections.

B.2 Elementary stochastic calculus

Main references: [1]

B.2.1 Introduction

Stochastic calculus is very important in the mathematical modeling of financial processes because of the underlying random nature of financial markets. This section provides a brief introduction to elementary stochastic calculus.

This section will follow the methodology of chapter 5 of [1] in which technical terminologies are explained by examples relating to a **coin tossing**, if possible, to make them intuitive.

Toss a coin. Every time you throw a head I give you \$1, every time you throw a tail you give me \$1. If R_i represents the random amount either \$1 or \$1 you make on the i-th toss then we have:

$$E[R_i] = 0, \quad E[R_i^2] = 1, \quad E[R_i R_j] = 0 \tag{B.16}$$

These expectations are not conditional on the past; in other words, if I threw five heads in a row it does not affect the outcome of the sixth toss.

Introduce S_i to represent the total amount of money you have won up and including the i-th toss so that:

$$S_i = \sum_{j=1}^{i} R_j \tag{B.17}$$

and assume that $S_0 = 0$, i.e. you start with no money. If we calculate the expectations of S_i it does matter what information we have:

$$E[S_i] = 0, \quad E[S_i^2] = E[R_1^2 + 2R_1 R_2 + ...] = i \tag{B.18}$$

On the other hand, if there are five tosses already, then the expectation of the sixth toss will depends on the fifth tosses. This is the conditional expectation. The expectation of S_i conditional upon the previous five tosses gives:

$$E[S_6 | R_1, R_2, ..., R_5] = S_5 \tag{B.19}$$

The property that the expected value of the random variable S_i *conditional upon all of the past events only depends on the previous value S_{i-1}* (doesn't have to be the case that the expected value of S_i is equal to S_{i-1}) is called **Markov property**. We say that the random walk has no memory beyond where it is now. Almost all of the financial models have the Markov property. This is of fundamental importance in modeling in finance.

After the fifth toss, you know how much money you have won. Your expected winnings after the sixth toss, and actually after any number of tosses if we keep playing, is just the amount you already hold.

$$E[S_i | R_j, j < i] = S_j \tag{B.20}$$

The property that the conditional expectation of your winnings at any time in the future is just the amount you already hold is called the **martingale property**. This is another important property in finance.

The **quadratic variation** of the random walk is defined as:

$$\sum_{j=1}^{i} (S_j - S_{j-1})^2 \tag{B.21}$$

Going back to the coin tossing example, because you either win or lose an amount \$1 after each toss, we actually have $|S_j - S_{j-1}| = 1$, thus the quadratic variation is always i.

The **mean square limit** technique states that:

$$E\left[\left(\sum_{j=1}^{n} (X(t_j) - X(t_{j-1}))^2 - t\right)\right] = O\left(\frac{1}{n}\right) \tag{B.22}$$

as $n \to \infty$ this tends to zero. Therefore, we can say that:

$$\sum_{j=1}^{n} (X(t_j) - X(t_{j-1}))^2 = t \tag{B.23}$$

in the 'mean square limit'. This is often written as:

$$\int_0^t (dX)^2 = t \tag{B.24}$$

B.2.2   Brownian motion

Now we make some changes in the coin tossing game: first, the time allowed for tossing is a period t; second, within the period t we will perform n tosses with the bet size of each throw is $\sqrt{t/n}$. This new experiment clearly still possesses both the Markov and martingale properties, and its quadratic variation measured over the whole experiment is:

$$\sum_{j=1}^{n} (S_j - S_{j-1})^2 = n \times \left(\sqrt{\frac{t}{n}}\right)^2 = t \tag{B.25}$$

To speed up the game, we will make n larger and larger, i.e. decreasing the time between tosses, with a smaller amount for each bet, i.e. still keep the bet size of $\sqrt{t/n}$. You will see that as we go to the limit $n = \infty$, the resulting random walk stays finite. It has an expectation conditional on a starting value of zero and a variance of t:

$$E\left[S(t)\right] = 0 \tag{B.26}$$
$$E\left[S(t)^2\right] = t \tag{B.27}$$

The limiting process for this random walk as the time steps go to zero is called **Brownian motion**, and it is denoted by $X(t)$. The important properties of Brownian motion, which are also very important for financial models, are as follows:

- **Finiteness**: any other scaling of the bet size or 'increments' with time step would have resulted in either a random walk going to infinity in a finite time, or a limit in which there was no motion at all. It is important that the increment scales with the square root of the time step.

- **Continuity**: the paths are continuous, there are no discontinuities. Brownian motion is the continuous-time limit of our discrete time random walk.

- **Markov**: the conditional distribution of $X(t)$ given information up until $\tau < t$ depends only on $X(\tau)$.

- **Martingale**: given information up until $\tau < t$ the conditional expectation of $X(t)$ is $X(\tau)$.

- **Quadratic variation**: if we divide the time 0 to t into $n + 1$ partition points $t_i = it/n$ then:

$$\sum_{j=1}^{n} (X(t_j) - X(t_{j-1}))^2 \to t \tag{B.28}$$

- **Normality**: over finite time increments t_{i-1} to t_i, $X(t_j) - X(t_{j-1})$ is normally distributed with mean zero and variance $t_i - t_{i-1}$.

B.2.3 Stochastic integration and differentiation

Introductory calculus courses teach differentiation first and integration second. The Itô calculus is derived and taught in the reverse order: first we need to understand the definition of an Itô (stochastic) integral, and then we can understand what it means for a stochastic process to have a differential.

The construction of the Itô integral begins with the backward Riemann sum:

$$W(t) = \int_0^t f(\tau) \, dX(\tau) = \lim_{n \to \infty} \sum_{j=1}^n f(t_{j-1}) \left(X(t_j) - X(t_{j-1}) \right) \tag{B.29}$$

with:

$$t_j = \frac{jt}{n} \tag{B.30}$$

The function $f(t)$ which is integrated is evaluated in the summation at the left-hand point t_{j-1} which means each function evaluation does not know about the random increment that multiplies it, i.e. the integration is non **anticipatory**. This choice of integration is natural in finance, ensuring that we use no information about the future in our current actions.

Taking the "differentiating" of the first equation of Equation B.29 we have:

$$dW = f(t) \, dX \tag{B.31}$$

dX can be considered as an an increment in X, i.e. a normal random variable with mean zero and standard deviation $\sqrt{dt}$. This equation looks like an ordinary differential equation but we do not make it identical to the ODE by dividing by dt because then we have the difficult task of defining $\frac{dX}{dt}$.

Pursuing this idea further, then the equation:

$$dW = g(t) \, dt + f(t) \, dX \tag{B.32}$$

is equivalent to:

$$W(t) = \int_0^t g(\tau) \, d\tau + \int_0^t f(\tau) \, dX(\tau) \tag{B.33}$$

Equations like B.33 are called **stochastic differential equations (SDE)** .

Itô's lemma

We know that in deterministic calculus if $F = X^2$ then $dF = 2X \, dX$. However this rule is not applicable in a stochastic environment. **Itô's lemma**, the most important rule of stochastic calculus, is used for this purpose since it allows us to manipulate functions of a random variable. This section provides the derivation of the Itô's lemma for an arbitrary function $F(X)$.

First we introduce a very, very small time scale:

$$\frac{\delta t}{n} = h \tag{B.34}$$

This timescale is so small that the function $F(X(t+h))$ can be approximated by a Taylor series:

$$F(X(t+h)) - F(X(t)) = (X(t+h) - X(t)) \frac{dF}{dX}(X(t)) + \frac{1}{2}(X(t+h) - X(t))^2 \frac{d^2F}{dX^2}(X(t)) \tag{B.35}$$

This can be followed as:

$$
\begin{aligned}
[F(X(t+h)) &- F(X(t))] + [F(X(t+2h)) - F(X(t+h))] + ... \\
&+ [F(X(t+nh)) - F(X(t+(n-1)h))] \\
&= \sum_{j=1}^n (X(t+jh) - X(t+(j-1))) \frac{dF}{dX}(X(t+(j-1)h)) \\
&+ \frac{1}{2} \sum_{j=1}^n (X(t+jh) - X(t+(j-1)))^2 \frac{d^2F}{dX^2}(X(t))
\end{aligned}
\tag{B.36}
$$

in which we use the approximation:

$$\frac{d^2F}{dX^2}(X(t+(j-1)h)) = \frac{d^2F}{dX^2}(X(t)) \tag{B.37}$$

which is consistent with the order of accuracy we require here.

The first term in Equation B.36 can be rewritten as:

$$F(X(t+nh)) - F(X(t)) = F(X(t+\delta t)) - F(X(t)) \tag{B.38}$$

the second term is just the definition of a normal integral:

$$\sum_{j=1}^{n}(X(t+jh) - X(t+(j-1)))\frac{dF}{dX}(X(t+(j-1)h)) = \int_{t}^{t+\delta t}\frac{dF}{dX}dX \tag{B.39}$$

and the last term, following the mean square limit (this will be presented later), can be shortened into:

$$\frac{1}{2}\sum_{j=1}^{n}(X(t+jh) - X(t+(j-1)))^2\frac{d^2F}{dX^2}(X(t)) = \frac{1}{2}\frac{d^2F}{dX^2}(X(t))\,\delta t \tag{B.40}$$

Finally we have:

$$F(X(t+\delta t)) - F(X(t)) = \int_{t}^{t+\delta t}\frac{dF}{dX}(X(\tau))\,dX(\tau) + \frac{1}{2}\int_{t}^{t+\delta t}\frac{d^2F}{dX^2}(X(\tau))\,d\tau \tag{B.41}$$

Now we extend this result over longer timescale, $0 \to t$, to get:

$$F(X(t)) = F(X(0)) + \int_{0}^{t}\frac{dF}{dX}(X(\tau))\,dX(\tau) + \frac{1}{2}\int_{0}^{t}\frac{d^2F}{dX^2}(X(\tau))\,d\tau \tag{B.42}$$

This is the integral version of **Itô's lemma** which is usually written as:

$$dF = \frac{dF}{dX}dX + \frac{1}{2}\frac{d^2F}{dX^2}dt \tag{B.43}$$

Now if we look back at the naive Taylor series expansion of F, completely disregarding the random nature of X, and treating dX as a small increment in X, we would get:

$$F(X+dX) = F(X) + \frac{dF}{dX}dX + \frac{1}{2}\frac{d^2F}{dX^2}dX^2 \tag{B.44}$$

and if we consider $F(X+dX) - F(X)$ is just the change in F we have:

$$dF = \frac{dF}{dX}dX + \frac{1}{2}\frac{d^2F}{dX^2}dX^2 \tag{B.45}$$

This equation is very similar to Equation B.43 (as Taylor series is very similar to Itô) with the only difference being that there is a dX^2 instead of a dt. However, following the mean square limit, we have:

$$\int_{0}^{t}(dX)^2 = t \tag{B.46}$$

we could write:

$$dX^2 = dt \tag{B.47}$$

This equation is technically incorrect, but using Taylor series with this "rule of thumb" in practice, we can get the right result. Besides, we shouldn't really think of dX^2 as being the square of a single normally distributed random variable, mean zero, variance dt. We should think of it as the sum of squares of lots and lots (an infinite number) of independent and identically distributed normal variables, each one having mean zero and a very, very small (infinitesimal) variance. When we add together lots of i.i.d. variables, we will get a quantity with a mean of dt and a variance which goes rapidly to zero as the 'lots' approach 'infinity'.

Now we can come back to the first example of this section, if $F = X^2$ then the stochastic differential equation of which F satisfies is:

$$dF = 2X \, dX + dt \tag{B.48}$$

In an integrated form:

$$X^2 = F(X) = F(0) + \int_0^t 2X \, dX + \int_0^t 1 d\tau = \int_0^t 2X \, dX + t \tag{B.49}$$

therefore:

$$\int_0^t X \, dX = \frac{1}{2}X^2 - \frac{1}{2}t \tag{B.50}$$

To end this section, we consider the stochastic differential equation:

$$dS = a(S) \, dt + b(S) \, dX \tag{B.51}$$

for some functions $a(S)$ and $b(S)$, and dX is the usual Brownian increment. Now if we have a function $V(S)$ of S, the stochastic differential equation applicable for V is:

$$dV = \frac{dV}{dS}dS + \frac{1}{2}b^2\frac{d^2V}{dS^2}dt \tag{B.52}$$

We can also further subtitute dS into the above equation to get an equation for dV in terms of the pure Brownian motion X:

$$dV = \left(a(S)\frac{dV}{dS} + \frac{1}{2}b(S)^2\frac{d^2V}{dS^2}\right) dt + b(S)\frac{dV}{dS}dX \tag{B.53}$$

Itô in higher dimensions

In financial problem, we often have functions of one stochastic variable S and a deterministic variable t (time) as $V(S,t)$. If:

$$dS = a(S,t) \, dt + b(S,t) \, dX \tag{B.54}$$

then the increment dV is given by:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}dt \tag{B.55}$$

This shorthand notation for the correct integrated from is written in the form of partial instead of ordinary derivatives.

Occasionally, we have a function of two or more random variables, and time as well, i.e. $V(S_1, S_2, t)$. We will write the behaviour of S_1 and S_2 in the general form:

$$dS_1 = a_1(S_1, S_2, t) \, dt + b_1(S_1, S_2, t) \, dX_1 \tag{B.56}$$
$$dS_2 = a_2(S_1, S_2, t) \, dt + b_2(S_1, S_2, t) \, dX_2 \tag{B.57}$$

in which we have two Brownian increment dX_1 and dX_2. We can think of these variables as being normally distributed with variance dt but they are *correlated*. The correlation between these two random variables is called ρ which can also be a function of S_1, S_2 and t but must satisfy:

$$-1 \leq \rho \leq 1 \tag{B.58}$$

The 'rule of thumb' for this case can be:

$$dX_1^2 = dt \quad dX_2^2 = dt \quad dX_1 dX_2 = \rho dt \tag{B.59}$$

Then Itô's lemma becomes:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S_1}dS_1 + \frac{\partial V}{\partial S_2}dS_2 + \frac{1}{2}b_1^2\frac{\partial^2 V}{\partial S_1^2}dt + \frac{1}{2}b_2^2\frac{\partial^2 V}{\partial S_2^2}dt + \rho b_1 b_2 \frac{\partial V}{\partial S_1 \partial S_2} \tag{B.60}$$

B.2.4   Some common random walks

In this section, a couple of common random walks will be presented. Recall that a stochastic differential equation model for variable S has the general form:

$$dS = \underline{\hspace{1.5cm}} dt + \underline{\hspace{1.5cm}} dX \qquad (B.61)$$

The part in front of the dt is deterministic and the part in front of the dX tells us how much randomness there is. Modelling is about choosing the functional form for the deterministic part and the functional form for the amount of randomness.

Brownian motion with drift

A Brownian motion with drift has the form:

$$dS = \mu\, dt + \sigma\, dX \qquad (B.62)$$

The point to note about this model is that S can go negative. This random walk would therefore not be a good model for many financial quantities, such as interest rates or equity prices. This stochastic differential equation can be integrated exactly to get"

$$S(t) = S(0) + \mu\, t + \sigma\, (X(t) - X(0)) \qquad (B.63)$$

The lognormal random walk

The lognormal random walk is similar to the Brownian with drift but the drift and randomness scale with S:

$$dS = \mu\, S\, dt + \sigma\, S\, dX \qquad (B.64)$$

With this model, if S starts out positive, it can never go negative, the closer that S gets to zero, the smaller the increments dS. The integral form of this stochastic differential equation follows from the stochastic differential equation for $F(S) = \log S$ (that's why it has the name "lognormal"):

$$S(t) = S(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma\, (X(t) - X(0))} \qquad (B.65)$$

This stochastic differential equation is particularly important in the modelling of many asset classes. And if we have some function $V(S,t)$ then from Itô it follows that :

$$dV = \frac{\partial V}{\partial t}\, dt + \frac{\partial V}{\partial S}\, dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\, dt \qquad (B.66)$$

A mean-reverting random walk

A typical mean-reverting random walk has the form:

$$dS = (\nu - \mu S)\, dt + \sigma\, dX \qquad (B.67)$$

If S is large, the negative coefficient in front of dt means that S will move down on average, if S is small it rises on average. There is still no incentive for S to stay positive in this random walk. Mean-reverting models are used for modeling a random variable that 'isn't going anywhere'. That's why they are often used for interest rates, e.g. governmental bonds.

B.2.5   Summary

This section revisited the most important tool of the trade, Itô's lemma. If we think of S as the value of an asset for which we have a stochastic differential equation, a 'model', then we can handle functions of the asset, and ultimately value contracts such as options.

If we use Itô as a tool we do not need to know why or how it works, only how to use it. Then, with use, familiarity breeds, if not contempt, sufficient confidence to make you believe that you understand. Some nice remarks on Itô's lemma given by Paul Wilmott [1] are presented here:

- Stochastic differential equations are like recipes for generating random walks.

- If you have some quantity, i.e. S, that follows a random walk, then any function of S is also going to follow a random walk.

- The answer for the question 'What is the random walk for this function of S?' comes from applying something very like Taylor series but with some 'rule of thumb':

  ⋆ When you do the Taylor series expansion, only keep terms of size dt or bigger $dt^{1/2}$ $(dt^2 = 0)$.
  ⋆ Every time you see a dX^2 term replace it with dt.
  ⋆ All terms with $dX \cdot dt$ varnishes, i.e. $dX \cdot dt = 0$.

Essentially all we require to successfully use the lemma is a rule of thumb, as explained in the text.

B.2.6 Exercises

This section represents some exercises from [1] to practice with stochastic calculus.

Problem

Given an Brownian motion $X(t)$, show that:

$$\int_0^t X(\tau) \, dX(\tau) = \frac{1}{2}X^2(t) - \frac{1}{2}t$$

$$\int_0^t \tau \, dX(\tau) = t \, X(t) - \int_0^t X(\tau) \, d\tau$$

$$\int_0^t X^2(\tau) \, dX(\tau) = \frac{1}{3}X^3(t) - \int_0^t X(\tau) \, d\tau$$

Problem

Given a function $f(t)$ which is continuous and bounded on $[0, t]$, prove the integration by parts:

$$\int_0^t f(\tau) \, dX(\tau) = f(t) \, X(t) - \int_0^t X(\tau) \, df(\tau) \tag{B.68}$$

Problem

Find $u(W, t)$ and $v(W, t)$ where:
$$dW(t) = u \, dt + v \, dX(t)$$

in three scenarios:

- $W(t) = X^2(t)$
- $W(t) = e^{X(t)} + t + 1$
- $W(t) = f(t) \, X(t)$

and $f(t)$ is a bounded and continuous function.

Problem

If S follows a lognormal random walk, use Itô's lemma to find the differential equations satisfied by:

- $f(S) = aS + b$
- $g(S) = S^n$
- $h(S, t) = S^n \, e^{mt}$

where a, b, m, n are constants.

Problem

If:
$$dS = \mu \, S \, dt + \sigma \, S \, dX$$
use Itô's lemma to find the stochastic differential equation satisfied by $f(S) = \log(S)$.

Problem

The change in the share price satisfies:
$$dS = A(S,t) \, dX + B(S,t) \, dt$$
for some functions A, B, find the stochastic differential equation satisfied by $f(S,t)$.

Problem

Two shares follow geometric Brownian motions as:
$$dS_1 = \mu_1 \, S_1 \, dt + \sigma_1 \, S_1 \, dX_1$$
$$dS_2 = \mu_2 \, S_2 \, dt + \sigma_2 \, S_2 \, dX_2$$

The share price changes are correlated with correlation coefficient ρ, find the stochastic differential equation satisfied by a function $f(S_1, S_2)$.

B.3 Black-Scholes model

Main references: [1, 6].

B.3.1 Black-Scholes equation

This section describes and explains the basic building blocks of derivatives theory which are delta hedging and no arbitrage.

A portfolio

The option value V can be written as:
$$V(S,t \, ; \sigma, \mu \, ; E, T \, ; r) \tag{B.69}$$
in which:

- the asset price S and the current time t are variables;
- the drift σ and the volatility μ are parameters associated with the asset price;
- the strike price E and the date of expiry T are parameters associated with the details of the particular contract;
- the interest rate r is a parameter associated with the currency in which the asset is quoted.

For the moment, we will only write $V(S,t)$ to denote the option value.

Now, consider a portfolio of one long option position and a short position in some quantity Δ of the underlying and denote Π as the value of that portfolio, we have:
$$\Pi = V(S,t) - \Delta S \tag{B.70}$$

The term ΔS is the short asset position, therefore it has a minus sign. The quantity Δ will be some constant quantity of our choosing.

Now, assume that the underlying asset follows a lognormal random walk:

$$dS = \mu S dt + \sigma S dX \tag{B.71}$$

The change in the value of the option from time t to time $t + dt$ is due partly to the change in the option value and partly to the change in the underlying:

$$d\Pi = dV - \Delta dS \tag{B.72}$$

Notice that Δ has not changed during the time step, as we assumed.

From Itô's lemma we have:

$$dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} dt \tag{B.73}$$

Thus, the portfolio changes by:

$$d\Pi = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} dt - \Delta dS \tag{B.74}$$

> The binomial analysis seems to be easier to understand than the stochastic analysis of the Black–Scholes model. In principle they are nearly identical, it's just that the math is a more abstract with the Black–Scholes model. From a technical point of view, in the binomial model we did lots of modeling, hedging, etc. first before arriving at the Black–Scholes partial differential equation by performing a Taylor series expansion. In the Black–Scholes analysis the Taylor series expansion, in its stochastic form, comes first and the hedging, etc. comes later.

Delta hedging

The right-hand side of Equation B.74 contains two types of terms, the deterministic terms being with dt and the random terms being with dS. Assume that we know V and its derivatives then we know everything on the right-hand side except for the value of dS. In fact we can never known this term in advance. The randomness terms are the risk in the portfolio. To reduce, or even eliminate this risk, we can, in theory and almost in practice, chose:

$$\Delta = \frac{\partial V}{\partial S} \tag{B.75}$$

then the randomness reduces to zero. Any reduction in randomness is generally termed **hedging**. The perfect elimination of risk, by exploiting correlation between two instruments (in this case an option and its underlying), is generally called **delta hedging**. Delta hedging is an example of a **dynamic hedging** strategy. From one time step to the next, the quantity $\frac{\partial V}{\partial S}$ changes therefore the perfect hedge must be continually rebalanced.

No arbitrage

After choosing the quantity Δ, we hold a portfolio whose value changes by the amount:

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt \tag{B.76}$$

This change is completely riskless. And in that case, the completely risk-free change $d\Pi$ must be the same as the growth we would get if we put the equivalent amount of cash in a risk-free interest-bearing account (i.e. the money-in-the-bank equation):

$$d\Pi = r\Pi dt \tag{B.77}$$

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

Black-Scholes equation

Substituting Equations B.70, B.75 and B.76 into B.77 we find that;

$$\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt = r\left(V - S\frac{\partial V}{\partial S}\right) dt \tag{B.78}$$

Dividing by dt and rearranging we get:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{B.79}$$

This is the Black–Scholes equation published in 1973 by Fischer Black and Myron Scholes. The Black–Scholes equation contains all the obvious variables and parameters such as the underlying S, time t, and volatility σ, but there is no mention of the drift rate μ since it is dropped out at the same time as we eliminated the dS component of the portfolio.

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

- As a linear function, if you have two solutions of the equation then the sum of these is itself also a solution.

- As a parabolic PDE (since it has a second derivative with respect to one variable, S, and a first derivative with respect to the other, t), it is related to the heat or diffusion equation of mechanics.

The Black–Scholes equation can be accurately interpreted as a reaction-convection diffusion equation [1]. The basic diffusion equation is a balance of a first-order t derivative and a second-order S derivative:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}$$

However, the diffusion coefficient is a function of one of the variables S, thus we really have diffusion in a non-homogeneous medium.

The first order S-derivative term:

$$rS\frac{\partial V}{\partial S}$$

can be thought of as a convection term. For example, if this equation represents the diffusion of smoke particles in the atmosphere, then the convective term would be due to a wind blowing the smoke in a preferred direction.

The final term:

$$-rV$$

is a reaction term. Balancing this term and the time derivative would give a model for decay of a radioactive body, with the half-life being related to r.

Putting these terms together and we get a reaction-convection-diffusion equation. An almost identical equation would be arrived at for the dispersion of pollutant along a flowing river with absorption by the sand. In this, the dispersion is the diffusion, the flow is the convection, and the absorption is the reaction [1].

B.3.2 Initial/final and boundary conditions

To uniquely specify a problem we must prescribe boundary conditions and an initial or final condition. Boundary conditions tell us how the solution must behave for all time at certain values of the asset.

In financial problems we usually specify the behavior of the solution at $S = 0$ and as $S \to \infty$. We must also tell the problem how the solution begins, i.e. a **initial condition**. The Black–Scholes equation is a backward equation, meaning that the signs of the t derivative and the second S derivative in the equation are the same when written on the same side of the equals sign. We therefore have to impose a **final condition**. This is usually the payoff function at expiry.

Besides, the Black-Scholes equation B.79 knows nothing about what kind of option we are valuing, whether it is a call or a put, nor what is the strike and the expiry. These points are dealt with by, also, the final condition. We must specify the option value V as a function of the underlying at the expiry date T, i.e. prescribing the payoff $V(S,T)$.

- For a call option:
$$V(S,T) = \max(S - E, 0) \tag{B.80}$$

- For a put option:
$$V(S,T) = \max(E - S, 0) \tag{B.81}$$

- For a binary call:
$$V(S,T) = \mathcal{H}(S - E, 0) \tag{B.82}$$

  in which $\mathcal{H}(x)$ is the **Heaviside function** which is zero when its argument is negative and one when it is positive.

- For a binary put:
$$V(S,T) = \mathcal{H}(E - S, 0) \tag{B.83}$$

B.3.3 Some other ways of deriving Black-Scholes equation

The derivation of the Black–Scholes equation above is the classical one, and similar to the original Black and Scholes derivation. There are other ways of getting to the same result. Here are a few approaches without any details [1]:

- The binomial model is a discrete time, discrete asset price model for underlying and again uses hedging and no arbitrage to derive a pricing algorithm for options. In taking the limit as the time step shrinks to zero we get the continuous-time Black–Scholes equation. We will see this derivation later.
- The martingale approach [1].
- Capital Asset Pricing Model [2].

I haven't studied the two later approaches mentioned above, just name them here as examples. More details information on those approaches can be found in [9].

B.3.4 Forward and future contracts

Forward contracts

Let $V(S,t)$ be the value of the forward contract at any time during its life on the underlying asset S and maturing at time T. Set up the portfolio of one long forward contract and short Δ of the underlying asset:
$$\Pi = V(S,t) - \Delta S \tag{B.84}$$

From time t to $t + dt$, this change by an amount:
$$d\Pi = \frac{\partial V}{\partial t} dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} dt + \frac{\partial V}{\partial S} dS - \Delta dS \tag{B.85}$$

---

[1]The martingale approach follows the idea that the value of an option can be shown to be an expectation, not a real expectation but a special, risk-neutral one; this result forms the basis for pricing by simulation. The concepts of hedging and no arbitrage are still used in this derivation. The major drawback with this approach is that it requires a probabilistic description of the financial world

[2]CAPM is a model for the behavior of risky assets and a principle and algorithm for defining and finding optimal ways to allocate wealth among the assets. Portfolios are described in terms of their risk (standard deviation of returns) and reward (expected growth). If you include options in this framework then the possible combinations of risk and reward are not increased. This is because options are just functions of their underlyings. This is market completeness. The risk and reward on an option and on its underlying are related and the Black–Scholes equation follows

To eliminate risk, again we choose:

$$\Delta = \frac{\partial V}{\partial S} \tag{B.86}$$

and by applying the no-arbitrage argument, we end up with exactly the Black-Scholes PDE again.

The final condition for the equation is simply the difference between the asset price S and the fixed delivery price $\bar{S}$ - assumed for now that we know it, therefore:

$$V(S,t) = S - \bar{S} \tag{B.87}$$

The solution of the Black-Scholes with this final condition is:

$$V(S,t) = S - \bar{S}e^{-r(T-t)} \tag{B.88}$$

This is the forward contract's value during its life.

Now we will calculate back $\bar{S}$. The delivery price is set initially $t = t_0$ as the price that gives the forward contract zero value. If the underlying asset is S_0 at t_0 then:

$$0 = S_0 - \bar{S}e^{-r(T-t_0)} \tag{B.89}$$

or:

$$\bar{S} = S_0 e^{-r(T-t_0)} \tag{B.90}$$

The forward price, as quoted, is the delivery price, as varying from day to day. So the forward price for the contract maturing at T is:

$$S\ e^{r(T-t)} \tag{B.91}$$

Future contracts

Let $F(S,t)$ denote the futures price. Remember that the value of the futures contract during its life is always zero because the change in value is settled daily, this cash flow must be taken into account in our analysis. Therefore, if we set up a portfolio of one long futures contract and short Δ of the underlying, we have the value of that portfolio is:

$$\Pi = \Delta S \tag{B.92}$$

The portfolio change in value is:

$$d\Pi = dF - \Delta dS \tag{B.93}$$

in which dF represents the cash flow due to the continual settlement. Applying the Itô's lemma we have:

$$d\Pi = \frac{\partial F}{\partial t}dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}dt + \frac{\partial F}{\partial S}dS - \Delta dS \tag{B.94}$$

Again, applying the hedging and no-arbitrage argument, i.e.:

$$\Delta = \frac{\partial V}{\partial S} \quad \text{and} \quad d\Pi = r\Pi dt \tag{B.95}$$

to get:

$$\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} + rS\frac{\partial F}{\partial S} = 0 \tag{B.96}$$

This equation has only three terms and it is not the same as the Black-Scholes equation.

The final condition is:

$$F(S,T) = S \tag{B.97}$$

indicating that the future price and the underlying must have the same value at maturity. The solution is just:

$$F(S,t) = Se^{r(T-t)} \tag{B.98}$$

B.3.5 Some popular options

Option on dividend-paying equities

The first generalization we discuss is how to value options on stocks paying dividends. To keep things simple let's assume that the asset receives a continuous and constant dividend yield D. Thus in a time dt each asset receives an amount $DSdt$. This must be factored into the derivation of the Black–Scholes equation.

The change in the value of the portfolio is given by:

$$d\Pi = \frac{\partial V}{\partial t}dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt + \frac{\partial V}{\partial t}dS - \Delta dS - D\Delta S dt \qquad (B.99)$$

in which the last term is the amount of the dividend per asset $DSdt$ multiplied by the number of the asset held $-\Delta$. The Δ is still given by the rate of change of the option value with respect to the underlying, and after some simple substitutions we now get:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D)S\frac{\partial V}{\partial S} - rV = 0 \qquad (B.100)$$

Currency options

Options on currencies are handled in exactly the same way. In holding the foreign currency we receive interest at the foreign rate of interest r_f. This is just like receiving a continuous dividend, and we also have:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - r_f)S\frac{\partial V}{\partial S} - rV = 0 \qquad (B.101)$$

B.3.6 Main solution methods

The three main mathematical approaches to derivative pricing are differential equations; binomial trees and expectations. All of these methods are based on pretty much the same assumptions. All of them will therefore give the same values for a contract, if all parameter values are the same.

Analytical solutions for the Black-Scholes equation are rather complicated. A few techniques can be listed here for references [1]:

- Transformation to constant coefficient diffusion equation (see on the exercise section)
- Green's functions
- Series solutions
- Similarity reduction
- Fourier transforms
- Laplace transforms

Analytical techniques, actually, can be used to solve only a small number of realistic financial problems; therefore you won't need to find explicit solutions. In practice, in the vast majority of cases we must solve the Black–Scholes equation numerically.

Speaking of numerical methods, each of the three approaches has its own associated numerical method. Differential equations, and the Black–Scholes equation, in particular, can be solved by finite-difference methods. The binomial tree model is, interestingly, also its own numerical method. Finally, pricing by calculating risk-neutral expectations can be solved with Monte Carlo simulations. These three methods will be briefly described in this note.

B.3.7 Exercises

This section represents some exercises from [1] to practice with stochastic calculus.

Problem

Check that the following are solutions of the Black-Scholes equation:

- $V(S, t) = S$
- $V(S, t) = e^{rt}$

Problem

What is the most general solution of the Black-Scholes equation with each of the following forms:

- $V(S, t) = A(S)$
- $V(S, t) = B(S) \times C(t)$

i.e. find the most general form of functions A, B and C.

Problem

Given an European call options $C(S, t)$ with expiry at time T on an underlying share price S with no dividends, prove the following bounds:

- $C \leq S$
- $C \geq \max(S - Ee^{-r(T-t)})$
- $0 \leq C_1 - C_2 \leq (E_2 - E_1)e{-}r(T - t)$

where C_1, C_2 are calls with exercise prices E_1, E_2, respectively and $E_1 < E_2$.

Problem

Given an European put options $P(S, t)$ with expiry at time T on an underlying share price S with no dividends, prove the following bounds:

- $P \leq Ee^{-r(T-t)}$
- $P \geq Ee^{-r(T-t)} - S$
- $0 \leq P_2 - P_1 \leq (E_2 - E_1)e{-}r(T - t)$

where P_1, P_2 are puts with exercise prices E_1, E_2, respectively and $E_1 < E_2$.

Problem

Given an European call options $C(S, t)$ on an underlying share price S with no dividends, prove the following bounds:

- for C_A and C_B are calls with the same exercise price E and expiry dates T_A and T_B, respectively and $T_A > T_B$

$$C_A \geq C_B \tag{B.102}$$

- for C_1, C_2 and C_3 are calls with the same expiry T and have exercise prices E_1, E_2 and E_3, respectively, and $E_1 < E_2 < E_3$

$$C_2 \leq \frac{E_3 - E_2}{E_3 - E_1}C_1 + \frac{E_2 - E_1}{E_3 - E_1}C_3 \tag{B.103}$$

Hint: consider a portfolio of $\Pi = \alpha C_1 - C_2 + (1 - \alpha)C_3$.

Problem

Given $C(S, t)$ and $P(S, t)$ are the values of European call and put options with exercise price E and expiry at time T. Show that a portfolio of long the call and short the put satisfies the Black-Scholes equation. What boundary and final conditions hold for this portfolio ?

Problem

Find the random walk followed by a European option $V(S, t)$. Use Black–Scholes to simplify the equation for dV.

Problem

Check that u_δ satisfies the diffusion equation where:

$$u_\delta = \frac{1}{2\sqrt{\pi\tau}} e^{-\frac{x^2}{4\tau}} \tag{B.104}$$

Problem

Solve the Black-Scholes equation by transformation to constant coefficient diffusion equation (a change of variables technique) as:

- Set:

$$V(S, t) = e^{\alpha x + \beta \tau} U(x, \tau) \tag{B.105}$$

  in which:

$$\alpha = -\frac{1}{2}\left(\frac{2r}{\sigma^2} - 1\right)$$

$$\beta = -\frac{1}{4}\left(\frac{2r}{\sigma^2} + 1\right)^2$$

$$S = e^x$$

$$t = T - \frac{2\tau}{\sigma^2}$$

- Then $U(x, \tau)$ satisfies the basic diffusion equation:

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2} \tag{B.106}$$

The solution for this problem will be described in the next section.

Problem

Consider an option with value $V(S, t)$ which has payoff at time T. Reduce the Black–Scholes equation, with final and boundary conditions, to the diffusion equation, using the following transformations:

$$S = e^x$$

$$t = T - \frac{2\tau}{\sigma^2}$$

$$V(S, t) = EV(x, \tau)$$

$$v(x, \tau) = e^{\alpha x + \beta \tau} u(x, \tau)$$

for some α and β. What is the transformed payoff? What are the new initial and boundary conditions? Illustrate with a vanilla European call option.

Problem

Reduce the following parabolic equation to the diffusion equation:

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} + a\frac{\partial u}{\partial x} + b \tag{B.107}$$

where a and b are constants.

Problem

Using a change of time variable, reduce:

$$c(\tau)\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} \tag{B.108}$$

to the diffusion equation when $c(\tau) > 0$.

Consider the Black–Scholes equation, when σ and r can be functions of time, but $k = 2r/\sigma^2$ is still a constant. Reduce the Black–Scholes equation to the diffusion equation in this case.

B.4 Black-Scholes formulæ and Greeks

B.4.1 Derivation for calls, puts and simple digitals

The Black–Scholes equation has simple solutions for calls, puts and binary options whose payoff is a known function of the asset price at expiry. In this section, such solutions are presented. These will serve as the verification for the numerical solutions later.

The Black-Scholes equation is:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{B.109}$$

which holds for $S > 0$ with $t \in [0, T)$. This equation must be solved with final condition depending on the payoff, i.e. $V(S, T) = $ payoff of the derivatives, each contract will have a different functional form prescribed at expiry $t = T$.

The main steps to solve the Black-Scholes equation consist of:

1. change variable from present value to future value terms: since the payoff is received at time T but we are valuing the option at time t, we can write:

$$V(S, t) = e^{-r(T-t)}U(S, t) \tag{B.110}$$

This makes Equation B.109 becomes:

$$\frac{\partial U}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S} = 0 \tag{B.111}$$

2. change time from t to $\tau = T - t$ as we are solving a backward equation to make the equation becomes:

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S} \tag{B.112}$$

3. change variable from S to $\xi = \log(S)$ to make all of the coefficients in the equation constant, independent of the underlying, we get:

$$\frac{\partial}{\partial S} = e^{-\xi}\frac{\partial}{\partial \xi} \quad \frac{\partial^2}{\partial S^2} = e^{-2\xi}\frac{\partial^2}{\partial \xi^2} - e^{-2\xi}\frac{\partial}{\partial \xi} \tag{B.113}$$

and Equation B.109 becomes:

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\sigma^2\frac{\partial^2 U}{\partial \xi^2} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial U}{\partial \xi} \tag{B.114}$$

This is a big step forward thanks to the lognormality of the underlying asset.

4. translate the coordinate system:

$$x = \xi + \left(r - \frac{1}{2}\sigma^2\right)\tau \tag{B.115}$$

$$U(S, t) = W(x, \tau) \tag{B.116}$$

to make Equation B.109 becomes the standard heat equation which the solution can traditionally be found:

$$\frac{\partial W}{\partial \tau} = \frac{1}{2}\sigma^2\frac{\partial^2 W}{\partial x^2} \tag{B.117}$$

Those steps above can be summarised as follows:

$$
\begin{aligned}
V(S,t) &= e^{-r(T-t)}U(S,t) \\
&= e^{-r\tau}U(S, T-\tau) \\
&= e^{-r\tau}U(e^{\xi}, T-\tau) \\
&= e^{-r\tau}U\left(e^{x-\left(r-\frac{1}{2}\sigma^2\right)\tau}, T-\tau\right) \\
&= e^{-r\tau}W(x,\tau)
\end{aligned}
$$

We interest in a special solution of Equation B.117 which has the following form:

$$
W(x,\tau) = \frac{1}{\sqrt{2\pi\tau}\sigma}e^{-\frac{(x-x')^2}{2\sigma^2\tau}} \tag{B.118}
$$

where x' is an arbitrary constant. The reason for this interest is explained in [1].

Now, as the payoff $\text{Payoff}(S)$ is the value of the option at time $t = T$, we can write:

$$
V(S,T) = \text{Payoff}(S) \tag{B.119}
$$

and this is also the final condition for the function V satisfying the Black-Scholes equation. In the equation of the new variables, this final condition is:

$$
W(x,0) = \text{Payoff}(e^x) \tag{B.120}
$$

The solution of this for $\tau > 0$ is:

$$
W(x,\tau) = \int_{-\infty}^{\infty} W(x,\tau,x')\text{Payoff}(e^{x'})dx' \tag{B.121}
$$

$$
= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\tau}\sigma}e^{-\frac{(x-x')^2}{2\sigma^2\tau}}\text{Payoff}(e^{x'})dx' \tag{B.122}
$$

If we now come back to the original variables then the solution becomes:

$$
V(S,t) = \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}}\int_{0}^{\infty} e^{-\frac{\left(\log(\frac{S}{S'})+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}}\text{Payoff}(S')\frac{dS'}{S'} \tag{B.123}
$$

where we have written $x' = \log(S')$. This is the exact solution for the option value in terms of the arbitrary payoff function.

Before moving to some special payoff functions, let take a break by looking at the integral of the form:

$$
\int_{a}^{\infty} e^{-\frac{1}{2}x^2}dx \tag{B.124}
$$

for some a. This integral is rather similar to the cumulative distribution function for the standardized normal distribution defined by:

$$
N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}\phi^2}d\phi \tag{B.125}
$$

This function is the probability that a normally distributed variable is less than x. You will see, later, that this $N(x)$ function is used alot in the solution of different payoff functions. The derivative of $N(x)$ is:

$$
N'(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \tag{B.126}
$$

To calculate $N(x)$, we can use the following formula which gives an accurate, and fast, approximation to the cumulative distribution function of the standardized normal distribution. For $x > 0$:

$$
N(x) \approx 1 - \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\left(a_1 d + a_2 d^2 + a_3 d^3 + a_4 d^4 + a_5 d^5\right) \tag{B.127}
$$

in which:

$$d = \frac{1}{1 + 0.2316419x} \quad a_1 = 0.31938153 \quad a_2 = -0.356563782$$

$$a_3 = 1.781477937 \quad a_4 = -1.821255978 \quad a_5 = 1.330274429$$

And for $x < 0$ use the fact that $N(x) + N(-x) = 1$.

Formula for a call

The call option has the payoff function as:

$$\text{Payoff}(S) = \max(S - E, 0) \tag{B.128}$$

The solution following Equation B.123 can then be written as:

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_E^\infty e^{-\frac{\left(\log(\frac{S}{S'}) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}} (S' - E) \frac{dS'}{S'} \tag{B.129}$$

Return to the variable $x' = \log(S')$ we have the solution as:

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_{\log E}^\infty e^{-\frac{\left(-x' + \log S + \left(r - \frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}} (e^{x'} - E) \, dx' \tag{B.130}$$

$$= \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_{\log E}^\infty e^{-\frac{\left(-x' + \log S + \left(r - \frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}} e^{x'} \, dx' \tag{B.131}$$

$$- E\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_{\log E}^\infty e^{-\frac{\left(-x' + \log S + \left(r - \frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}} \, dx' \tag{B.132}$$

Both integrals in this expression can be written in the form of Equation B.124 and, hence, B.125, therefore the option price can be written as two separate terms involving the cumulative distribution function of a normal distribution:

$$\text{Call option value} = S\,N(d_1) - Ee^{-r(T-t)}\,N(d_2) \tag{B.133}$$

where:

$$d_1 = \frac{\log(S/E) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \tag{B.134}$$

$$d_2 = \frac{\log(S/E) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{(T-t)} \tag{B.135}$$

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

$$\text{Call option value} = Se^{-D(T-t)}\,N(d_1) - Ee^{-r(T-t)}\,N(d_2) \tag{B.136}$$

and:

$$d_1 = \frac{\log(S/E) + \left(r - D + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \tag{B.137}$$

$$d_2 = \frac{\log(S/E) + \left(r - D - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{(T-t)} \tag{B.138}$$

Formula for a put

The call option has the payoff function as:

$$\text{Payoff}(S) = \max(E - S, 0) \tag{B.139}$$

Following the same approach as described above, we get:

$$\text{Put option value } = -S\,N(-d_1) + Ee^{-r(T-t)}\,N(-d_2) \tag{B.140}$$

with exactly the same d_1 and d_2. When there is continuous dividend yield on the underlying, or it is a currency, then:

$$\text{Put option value } = -Se^{-D(T-t)}\,N(-d_1) + Ee^{-r(T-t)}\,N(-d_2) \tag{B.141}$$

Another, and more beautiful, approach to find the put option value is found thanks to the linearity of the Black-Scholes equation (i.e. a linear combination of solutions is again a solution) and the solution for a call which we have already found. Recall a payoff of a call-put parity:

$$-[\text{call payoff}] + [\text{put payoff}] + [\text{asset price}] = E \tag{B.142}$$

therefore:

$$[\text{put payoff}] = [\text{call payoff}] - S + E \tag{B.143}$$

Recall the individual solutions:

Table B.1: Individual solutions of the Black-Scholes equation

Final condition	Solution	Note
$\max(S - E, 0)$	$V_{\text{call}}(S, t)$	above
S	S	exercise
E	$Ee^{-r(T-t)}$	exercise

And using the linearity property we have:

$$\text{Final condition: } \max(E - S, 0) = \max(S - E, 0) - S + E \tag{B.144}$$

$$\text{Solution: } V_{\text{put}}(S, t) = V_{\text{call}}(S, t) - S + Ee^{-r(T-t)} \tag{B.145}$$

and if we write in the same form of the call option value we have:

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

This comes from the equality $N(x) + N(-x) = 1$.

Combined strategies
If the strategy is a linear combination of call and put options, then its price is the same linear combination of the call and put options prices, thanks to the linearity of the Black-Scholes PDE.
Example: consider a butterfly option strategy

- If we buy call options with exercise prices E_1, E_3 and sell two call options with exercise prices E_2 in which $E_1 < E_2 < E_3$ and $E_1 + E_3 = 2E_2$.
- The payoff of the strategy can be written as:

$$V(S, T) = \max(S - E_1, 0) - 2\max(S - E_2, 0) + \max(S - E_3, 0) \tag{B.146}$$

- And hence, the Black-Scholes price is:

$$V(S, t) = V_{\text{call}}(S, t, E_1) - 2V_{\text{call}}(S, t, E_2) + V_{\text{call}}(S, t, E_3) \tag{B.147}$$

Formula for a binary call

The binary call has the payoff function:

$$\text{Payoff}(S) = \mathcal{H}(S - E) \tag{B.148}$$

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

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_{\log E}^{\infty} e^{-\frac{\left(-x'-\log S-\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}} dx' \tag{B.149}$$

or:

$$e^{-r(T-t)} N(d_2) \tag{B.150}$$

Formula for a binary put

The binary put has the payoff function:

$$\text{Payoff}(S) = \mathcal{H}(E - S) \tag{B.151}$$

The binary put option value is then:

$$e^{-r(T-t)}\left(1 - N(d_2)\right) \tag{B.152}$$

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

B.4.2 Greeks

Delta

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

$$\Delta = \frac{\partial V}{\partial S} \tag{B.153}$$

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

Delta hedging is used in practice for eliminating risk, but the usage of it is not simple. Please read section 8.3 of [1] for insights of delta hedging. Formulæ for the deltas of common contracts are given below:

- A call option:
$$e^{-D(T-t)} N(d_1) \tag{B.154}$$

- A put option:
$$e^{-D(T-t)}\left(N(d_1) - 1\right) \tag{B.155}$$

- A binary call:
$$\frac{e^{-r(T-t)} N'(d_2)}{\sigma S\sqrt{T-t}} \tag{B.156}$$

- A binary put:
$$-\frac{e^{-r(T-t)} N'(d_2)}{\sigma S\sqrt{T-t}} \tag{B.157}$$

Gamma

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

$$\Delta = \frac{\partial^2 V}{\partial S^2} \tag{B.158}$$

Since the gamma is the sensitivity of the delta to the underlying it is a measure of how much or how often a position must be rehedged in order to maintain a delta-neutral position. Although the delta also varies with time this effect is dominated by the Brownian nature of the movement in the underlying. In a delta-neutral position the gamma is partly responsible for making the return on the portfolio equal to the risk-free rate, the no-arbitrage condition. The rest of this task falls to the time-derivative of the option value, the theta given later. More insights in the role and the usage of the gamma can be found in section 8.4 of [1].

Formulæ for the gammas of common contracts are given below:

- A call option:

$$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}} \tag{B.159}$$

- A put option:

$$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}} \tag{B.160}$$

- A binary call:

$$\frac{e^{-r(T-t)}d_1 N'(d_2)}{\sigma^2 S^2(T-t)} \tag{B.161}$$

- A binary put:

$$-\frac{e^{-r(T-t)}d_1 N'(d_2)}{\sigma^2 S^2(T-t)} \tag{B.162}$$

Theta

The theta is the rate of change of the option price with time.

$$\Theta = \frac{\partial V}{\partial t} \tag{B.163}$$

The theta is related to the option value, the delta and the gamma by the Black–Scholes equation. In a delta-hedged portfolio the theta contributes to ensuring that the portfolio earns the risk-free rate. But it contributes in a completely certain way, unlike the gamma which contributes the right amount on average.

Formulæ for the thetas of common contracts are given below:

- A call option:

$$-\frac{\sigma S e^{-D(T-t)}N'(d_1)}{2\sqrt{T-t}} + DSN(d_1)e^{-D(T-t)} - rEe^{-r(T-t)}N(d_2) \tag{B.164}$$

- A put option:

$$-\frac{\sigma S e^{-D(T-t)}N'(-d_1)}{2\sqrt{T-t}} - DSN(-d_1)e^{-D(T-t)} + rEe^{-r(T-t)}N(-d_2) \tag{B.165}$$

- A binary call:

$$re^{-r(T-t)}N(d_2) + e^{-r(T-t)}N'(d_2)\left(\frac{d_1}{2(T-t)} - \frac{r-D}{\sigma\sqrt{T-t}}\right) \tag{B.166}$$

- A binary put:

$$re^{-r(T-t)}\left(1 - N(d_2)\right) - e^{-r(T-t)}N'(d_2)\left(\frac{d_1}{2(T-t)} - \frac{r-D}{\sigma\sqrt{T-t}}\right) \tag{B.167}$$

Speed

The speed of an option is the rate of change of the gamma with respect to the stock price.

$$\text{Speed } = \frac{\partial^3 V}{\partial S^3} \tag{B.168}$$

Traders use the gamma to estimate how much they will have to rehedge when the stock moves. If the stock moves by \$1 so the delta changes by whatever the gamma is. But that's only an approximation. The delta may change by more or less than this, especially if the stock moves by a larger amount, or the option is close to the strike and expiration. Hence the use of speed in a higher-order Taylor series expansion.

Formulæ for the speeds of common contracts are given below:

- A call option:

$$-\frac{e^{-D(T-t)}N'(d_1)}{\sigma^2 S^2(T-t)}\left(d_1 + \sigma\sqrt{T-t}\right) \tag{B.169}$$

- A put option:

$$-\frac{e^{-D(T-t)}N'(d_1)}{\sigma^2 S^2(T-t)}\left(d_1 + \sigma\sqrt{T-t}\right) \tag{B.170}$$

- A binary call:

$$-\frac{e^{-r(T-t)}N'(d_2)}{\sigma^2 S^3(T-t)}\left(-2d_1 + \frac{1-d_1 d_2}{\sigma\sqrt{T-t}}\right) \tag{B.171}$$

- A binary put:

$$\frac{e^{-r(T-t)}N'(d_2)}{\sigma^2 S^3(T-t)}\left(-2d_1 + \frac{1-d_1 d_2}{\sigma\sqrt{T-t}}\right) \tag{B.172}$$

Vega

The vega, a.k.a. zeta and kappa, is a very important quantity. It is the sensitivity of the option price to volatility.

$$\text{Vega} = \frac{\partial V}{\partial \sigma} \tag{B.173}$$

This is completely different from the other greeks since it is a derivative with respect to a parameter and not a variable. In practice, the volatility of the underlying is not known with certainty. Not only is it very difficult to measure at any time, it is even harder to predict what it will do in the future. One can vega-hedge to reduce sensitivity to the volatility. This is a major step towards eliminating some model risk, since it reduces dependence on a quantity that is not known very accurately. More insights in the vega can be found in section 8.7 of [1].

Formulæ for the vegas of common contracts are given below:

- A call option:

$$S\sqrt{T-t}\,e^{-D(T-t)}N'(d_1) \tag{B.174}$$

- A put option:

$$S\sqrt{T-t}\,e^{-D(T-t)}N'(d_1) \tag{B.175}$$

- A binary call:

$$-e^{-r(T-t)}N'(d_2)\frac{d_1}{\sigma} \tag{B.176}$$

- A binary put:

$$e^{-r(T-t)}N'(d_2)\frac{d_1}{\sigma} \tag{B.177}$$

Rho

The rho is the sensitivity of the option value to the interest rate used in the Black–Scholes formulæ

$$\rho = \frac{\partial V}{\partial r} \tag{B.178}$$

In practice one often uses a whole term structure of interest rates, meaning a time dependent rate $r(t)$. The rho would then be the sensitivity to the level of the rates assuming a parallel shift in rates at all times.

Formulæ for the rhos of common contracts are given below:

- A call option:

$$E(T-t)e^{-r(T-t)}N(d_2) \tag{B.179}$$

- A put option:

$$-E(T-t)e^{-r(T-t)}N(-d_2) \tag{B.180}$$

- A binary call:

$$-(T-t)e^{-r(T-t)}N(d_2) + \frac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2) \tag{B.181}$$

- A binary put:

$$-(T-t)e^{-r(T-t)}\left(1 - N(d_2)\right) - \frac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2) \tag{B.182}$$

The sensitivities of common contract to the dividend yield are given by the following formulæ:

- A call option:

$$-(T-t)Se^{-D(T-t)}N(d_1) \tag{B.183}$$

- A put option:

$$(T-t)Se^{-D(T-t)}N(-d_1) \tag{B.184}$$

- A binary call:

$$-\frac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2) \tag{B.185}$$

- A binary put:

$$\frac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2) \tag{B.186}$$

B.4.3   Implied volatility

The Black–Scholes formula for a call option takes as input the expiry T, the strike E, the underlying S and the interest rate r together with the volatility σ to output the price. All but the volatility are easily measured.

In the relationship between volatility and an option price, if we can see the price at which the option is trading, we can ask 'What volatility must I use to get the correct market price?' This is called the implied volatility. The implied volatility is the volatility of the underlying which when substituted into the Black–Scholes formula gives a theoretical price equal to the market price.

Assuming that we are using call prices to estimate the implied volatility then provided the option price is less than the asset and greater than zero then we can find a unique value for the implied volatility. Because there is no simple formula for the implied volatility as a function of the option value we must solve the equation:

$$V_{BS}\left(S_0, t_0, \sigma, r, E, T\right) = \textbf{known value} \tag{B.187}$$

for σ in which V_{BS} is the Black-Scholes formula, today's asset price is S_0, the date is $t_0 = 0$ and everything else in this equation is known except for σ. To solve this equation, one can use the

Newton-Raphson method in which the derivative of the option price with respect to the volatility (the vega) is calculated. An illustrative algorithm in this case can be given as:

Algorithm 1: Implied volatility for a simple call option

```
1 begin
2     Asset  S_0, E, r, t_0 = 0, T, V_real                                    /* Input */
3     Volatility initiated   σ = σ_0                                   /* Initialisation */
4     Accuracy required   ε
5     dσ = ε + 1                                                 /* Iterative calculation */
6     while dσ > ε do
7         d_1 = [log(S/E) + (r + ½σ²)T] / (σ√T)
8         d_2 = d_1 − σ√T
9         V_error = SN(d_1) − Ee^(−rT)N(d_2) − V_real
10        Vega = S√T N'(d_1)
11        dσ = V_error / Vega
12        σ = σ − dσ
13    end
14 end
```

In practice if we calculate the implied volatility for many different strikes and expiries on the same underlying then we find that the volatility is not constant. The implied volatilities for the calls and puts should be identical, because of put-call parity. The dependence of the implied volatility on strike and expiry can be interpreted in many ways. The easiest interpretation is that it represents the market's view of future volatility in some complex way [1].

B.4.4 Short introduction into volatility modelling

In practice, we do not know what volatility currently is or what it may be in the future. Volatility analysis and modelling takes a crucial role in finance simulation. More information from [1] is represented in Section D.1. This section describes several volatility estimation by statistical means.

Constant volatility/moving window

If we believe that volatility is constant or slowly varying, and we have N days data, we can use:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} R_i^2 \tag{B.188}$$

where:

$$R_i = \frac{S_i - S_{i-1}}{S_{i-1}} \tag{B.189}$$

is the return on the i-th day. Note that this term is not annualized (multiply by the square root of the number of trading days in a year) yet.

There are major problems associated with this volatility measure, because the returns are equally weighted you will get a plateauing effect associated with a large return. If there is a large one-day return in will increase the volatility instantaneously, but the estimate of volatility will stay raised until N days later when that return drops out of the sample. This is a totally spurious effect.

Incorporating mean reversion

Now consider time-varying volatility. We don't just have one σ but must consider σ_n, our estimate of the volatility on the n-th day, using data available up to that point. If we believe that volatility tends to vary about a long-term mean σ, then we could use:

$$\sigma_n^2 = \alpha \bar{\sigma}^2 + (1 - \alpha)\frac{1}{n} \sum_{i=1}^{n} R_i^2 \tag{B.190}$$

Here there is a weighting assigned to each long-run volatility estimate and the current estimate based on the last n returns. This is called an **ARCH** model, for **Autoregressive Conditional Heteroscedasticity**. The hidden assumption in this model is that the last n returns are equally important.

Exponentially weighted moving average

Consider this estimate for volatility:

$$\sigma_n^2 = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^{i-1} R_{n-i+1}^2 \tag{B.191}$$

The parameter λ must be greater than zero and less than one. This is an example of an exponentially weighted moving average estimate. The more recent the return, the more weight is attached. The sum extends back to the beginning of time. The coefficient of $1 - \lambda$ ensures that the weights all add up to one.

This expression can be simplified to:

$$\sigma_n^2 = \lambda \sigma_{n-1}^2 + (1 - \lambda) R_n^2 \tag{B.192}$$

Note that this uses the most recent return and the previous estimate of the volatility. This is *Risk-Metrics* volatility measure.

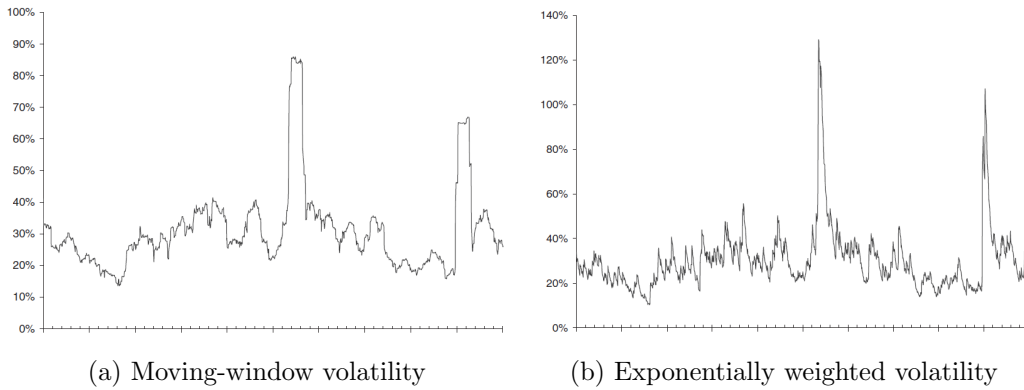(a) Moving-window volatility   (b) Exponentially weighted volatility

Figure B.3: Volatility modelling with move-window and exponentially weighted average

Figure B.3 uses the same stock price to estimate the volatility following the moving-window technique and the exponentially-weighted moving average technique. It is clear shown that the first has plateauing whereas the second doesn't have.

A simple GRACH model

Put the preceeding models together to get:

$$\sigma_n^2 = \alpha \bar{\sigma}^2 + (1 - \alpha) \left(\lambda \sigma_{n-1}^2 + (1 - \lambda) R_n^2 \right) \tag{B.193}$$

This is **GARCH** model, for **Generalized Autoregressive Conditional Heteroscedasticity**.

Range-based estimation of volatility

The problem with estimating volatility is that you need lots and lots of data to avoid sampling-error problems. But then if you use too many days worth of data you will be trying to estimate a parameter during a period when that parameter is almost certainly varying. To avoid this problem, one can use more information contained within a single day which is going down to finer timescales for the data. The problem with that is the behavior of returns over very short timescales, such as minutes, does not appear to be Normally distributed. There is even some evidence that the returns do not have a finite standard deviation. Setting aside such worries, for now.

Traditional close-to-close measure   When the drift is small:

$$\sigma_{cc}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(\log \left(\frac{C_i}{C_{i-1}} \right) \right)^2 \tag{B.194}$$

Here there is a slight change of notation from before; C_i is the closing price on the i-th day. To adjust this for the drift take:

$$\sigma_{cc}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(\left(\log \left(\frac{C_i}{C_{i-1}} \right) \right)^2 - \frac{\left(\log \left(\frac{C_n}{C_0} \right) \right)^2}{n(n-1)} \right) \tag{B.195}$$

Parkinson 1980   This estimator uses extreme value, the highs H and the lows L during the day:

$$\sigma_p^2 = \frac{1}{4n \log(2)} \sum_{i=1}^{n} \left(\log \left(\frac{H_i}{L_i} \right) \right)^2 \tag{B.196}$$

This is five times more efficient than the close-to-close estimate which means for the same amount of data the variance of the data is one fifth that of the close-to-close measure.

Garman & Klass 1980   At 7.4 times more efficient than close-to-close, we have:

$$\sigma_{gk}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(0.511 \left(\log \left(\frac{H_i}{L_i} \right) \right)^2 - 0.019 \log \left(\frac{C_i}{O_i} \right) \log \left(\frac{H_i L_i}{O_i^2} \right) - 2 \log \left(\frac{H_i}{O_i} \right) \log \left(\frac{L_i}{O_i} \right) \right) \tag{B.197}$$

Here O_i is the opening price.

Rogers & Satchell 1991   Parkinson and Garman & Klass are not independent of the drift. Our last simple volatility estimation is:

$$\sigma_{rs}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(\log \left(\frac{H_i}{C_i} \right) \log \left(\frac{H_i}{O_i} \right) + \log \left(\frac{L_i}{C_i} \right) \log \left(\frac{L_i}{O_i} \right) \right) \tag{B.198}$$

B.5   FDM solutions for Black-Scholes equation (one-factor model)

[10, 11, 12, 13, 14, 15]

B.5.1   Introduction

Not yet...

B.5.2   Differentiation using the grid

Not yet...

B.5.3   Approximating Greeks

Not yet...

B.5.4   Boundary conditions and payoffs

Not yet...

B.5.5   Explicit method

Black-Scholes equation

Not yet...

Convergence

Not yet...

European option

Not yet...

American option

Not yet...

Runge-Kutta method

Not yet...

B.5.6   Implicit method

Crank-Nicolson method

Not yet...

B.6   Binomial model

This section generalizes the numerical illustration of the binomial model for the option value given in Section A.3. Then, further applications of the binomial model are presented. The content of this section is heavily based on Chapter 3 of [1] and Chapter 10 of [6]. It is worth to repeat the Paul Wilmott's comment on the binomial model that "the binomial model is great for getting intuition about delta hedging and risk-neutral pricing but not for real-life pricing".

B.6.1   Initiation

In the binomial model, we assume that the asset, which initially has the value S, can, during a time step δt, either:

- rise or fall to a value $u \times S$ or $v \times S$, respectively with $0 < v < 1 < u$;
- the probability of a rise is p and so the probability of a fall is $1 - p$;
- the three constants (u, v, p) are chosen to give the binomial walk the same characteristics as the asset we are modeling.

For a more proper approach to quantitative finance, instead of using (u, v, p), we write everything in terms of the mean and the standard deviation of the random walk. That involves writing (u, v, p) in terms of what are known as the drift and the volatility of the asset, the drift μ is the average rate at which the asset rises, and the volatility σ is a measure of its randomness. Since we have three parameters to choose (u, v, p) and only two statistical quantities to fit (μ, σ), choices for conversion is unlimited. One of the possible conversion follows:

$$u = 1 + \sigma\sqrt{\delta t} \tag{B.199}$$

$$v = 1 - \sigma\sqrt{\delta t} \tag{B.200}$$

$$p = \frac{1}{2} + \frac{\mu\sqrt{\delta t}}{2\sigma} \tag{B.201}$$

This conversion will be described in a little bit more details in the next section. The correction of this conversion is, however, examined below.

The expected asset price after one time step is:

$$puS + (1-p)vS = \left(\frac{1}{2} + \frac{\mu\sqrt{\delta t}}{2\sigma}\right)\left(1 + \sigma\sqrt{\delta t}\right)S + \left(\frac{1}{2} - \frac{\mu\sqrt{\delta t}}{2\sigma}\right)\left(1 - \sigma\sqrt{\delta t}\right) \tag{B.202}$$

$$= (1 + \mu\delta t)\,S \tag{B.203}$$

so the expected change in the asset is $\mu S\delta t$. Because the expected return is $\mu\delta t$ so the expected asset price is corrected.

The variation of change in the asset price is:

$$S^2\left(p(u - 1 - \mu\delta t)^2 + (1-p)(v - 1 - \mu\delta t)^2\right) \tag{B.204}$$

$$= S^2\left(\left(\frac{1}{2} + \frac{\mu\sqrt{\delta t}}{2\sigma}\right)\left(\sigma\sqrt{(\delta t)} - \mu\delta t\right)^2 + \left(\frac{1}{2} - \frac{\mu\sqrt{\delta t}}{2\sigma}\right)\left(\sigma\sqrt{(\delta t)} + \mu\delta t\right)^2\right) \tag{B.205}$$

$$= S^2\left(\sigma^2\delta t - \mu^2\delta t^2\right) \tag{B.206}$$

so the standard deviation of asset changes is (approximately) $S\sigma\sqrt{\delta t}$. Because the standard deviation of return is $\sigma\sqrt{\delta t}$ so the standard deviation of asset changes is corrected.

Because of the similarity in the expected change in the asset and the standard deviation of asset changes, it can be concluded that the conversion from (u, v, p) to (μ, σ) is acceptable.

B.6.2   Value of an option

This section is a numerical generalization of Section A.3. Suppose that we have an underlying asset of value S, and we know the value of the option at time $t + \delta t$. Now construct a portfolio at time t consisting of one option and a short position in a quantity Δ of the underlying.

At time t this portfolio has value:

$$\Pi = V - \Delta S \tag{B.207}$$

where the option value V is, for the moment, unknown. At time $t + \delta t$, the option takes one of two values depending on whether the asset rises or falls:

$$V^+ \quad \text{or} \quad V^- \tag{B.208}$$

and hence, the portfolio becomes either:

$$V^+ - \Delta uS \quad \text{or} \quad V^- - \Delta uS \tag{B.209}$$

Because we know V^+, V^-, u, v, S, the values of both expressions are just linear functions of Δ.

Hedging

Having the freedom to choose Δ, we can make the value of this portfolio the same whether the asset rises or falls, i.e. hedging, which means:

$$V^+ - \Delta uS = V^- - \Delta uS \tag{B.210}$$

by choosing:

$$\Delta = \frac{V^+ - V^-}{(u - v) \times S} \tag{B.211}$$

The portfolio value is then:

$$V^+ - \Delta uS = V^+ - \frac{u(V^+ - V^-)}{u - v} \tag{B.212}$$

if the stock rises; or:

$$V^- - \Delta uS = V^- - \frac{v(V^+ - V^-)}{u - v} \tag{B.213}$$

if the stock falls. And, of course, these two expressions are the same.

No arbitrage

Since the value of the portfolio is guaranteed, following the no-arbitrage argument, its value must coincide with the value of the original portfolio plus any interest earned at the risk-free rate. Let's denote the portfolio value by:

$$\Pi + \delta\Pi \tag{B.214}$$

which just means the original portfolio value plus the change in value. Thus:

$$\delta\Pi = r\Pi\delta t \tag{B.215}$$

The full formulation for the portfolio value is then:

$$\Pi + \delta\Pi = \Pi + r\Pi\delta t = \Pi(1 + r\delta t) = V^+ - \frac{u(V^+ - V^-)}{u - v} \tag{B.216}$$

with:

$$\Pi = V - \Delta S = V - \frac{V^+ - V^-}{u - v} \tag{B.217}$$

Rearrange the above two equations as an equation for V we get:

$$(1 + r\delta t)V = (1 + r\delta t)\frac{V^+ - V^-}{u - v} + \frac{uV^- - vV^+}{u - v} \tag{B.218}$$

This is an equation for V given V^+ and V^- - the option values at the next time step, and the parameters u and v describing the random walk of the asset.

This equation can also be written in the more compacted form:

$$V = \frac{1}{1 + r\delta t}\left(p'V^+ + (1 - p')V^-\right) \tag{B.219}$$

in which:

$$p' = \frac{1}{2} + \frac{r\sqrt{\delta t}}{2\sigma} \tag{B.220}$$

. The right-hand side of Equation B.219 is just like a discounted expectation. The terms in the parentheses are the sum of probabilities multiplied by events. The probability p' is called the **risk-neutral probability**. It is like the real probability but used in the **risk-free world**. In the real world, the probability is calculated with the drift μ instead of the interest rate r:

$$p' = \frac{1}{2} + \frac{r\sqrt{\delta t}}{2\sigma} \quad \text{vs.} \quad p = \frac{1}{2} + \frac{\mu\sqrt{\delta t}}{2\sigma} \tag{B.221}$$

. Observe that the risk-free interest rate r plays two roles in option valuation. It's used once for discounting to give present value, and it's used as the drift rate in the risk-neutral asset price random walk.

Equation B.219 is equivalent to the statement "the option value at any time is the present value of the risk-neutral expected value at any later time", so it is actually the risk-neutral expectation.

Binomial tree

The binomial tree presents the possible development of the asset value following the random walk of the binomial model all the way until expiry. An example of the binomial tree is given in Figure B.4 where the nodes represent the values taken by the asset. Observe how the tree bends due to the geometric nature of the asset growth although it is normally presented in an equal-distance manner as in Figure B.5.
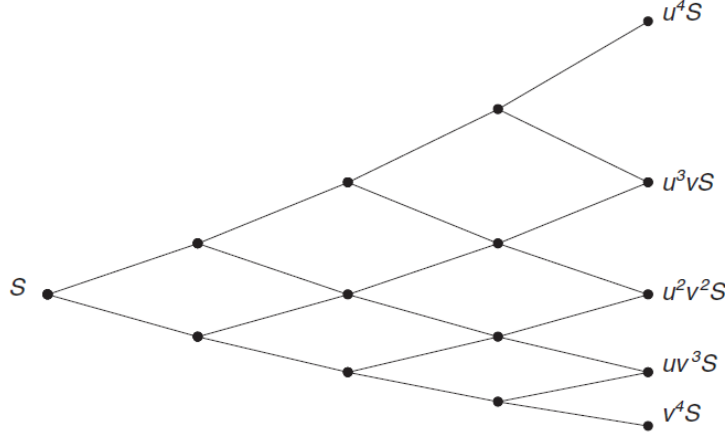
Figure B.4: A binomial tree

The probability of reaching a particular node in the binomial tree depends on the number of distinct paths to that node and the probabilities of the up and down moves. The binomial tree therefore contains within it an approximation to the probability density function for the lognormal random walk, Figure B.5.
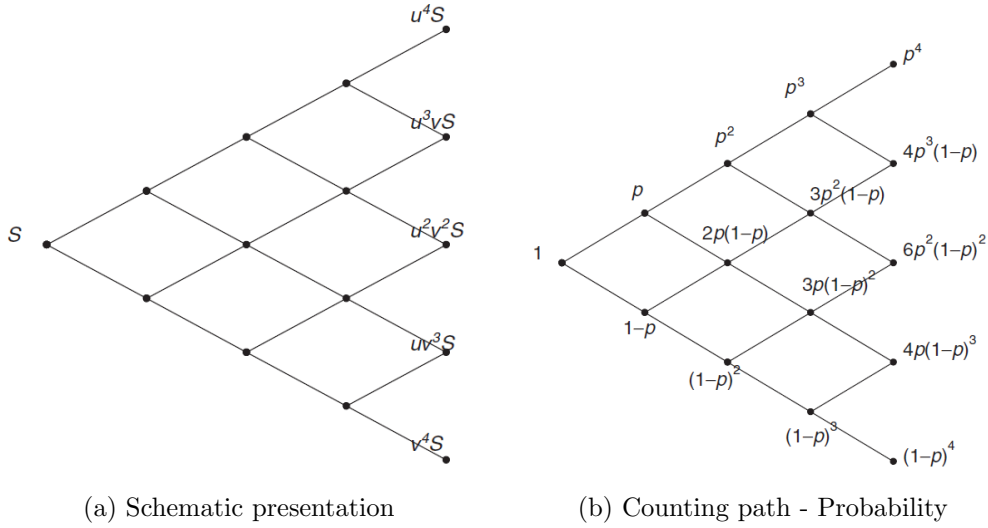

(a) Schematic presentation   (b) Counting path - Probability

Figure B.5: A binomial tree

The Greeks

The greeks are defined as derivatives of the option value with respect to various variables and parameters. It is important to distinguish whether the differentiation is with respect to a variable or a parameter. If the differentiation is only with respect to the asset price and/or time then there is sufficient information in our binomial tree to estimate the derivative. It may not be an accurate estimate, but it will be an estimate. The option's Δ, Γ and Θ can all be estimated from the tree. On the other hand, it is harder if you want to examine the sensitivity of the option with respect to one of the parameters. They are harder to calculate in the sense that you must perform a second binomial calculation. This applies to the option's Vega and ρ.

The option's delta is defined by:

$$\Delta = \frac{V^+ - V^-}{(u - v) \times S} \tag{B.222}$$

As all quantities on the right-hand side are known, we can calculate the delta directly from the tree.

In the limit as the time step approaches zero, the delta becomes:

$$\Delta = \frac{\partial V}{\partial S} \tag{B.223}$$

The gamma of the option is also defined as a derivative of the option with respect to the underlying:

$$\Gamma = \frac{\partial^2 V}{\partial S^2} \tag{B.224}$$

Gamma is a measure of how much we must rehedge at the next time step. Using the binomial tree, the gamma is just the change in the delta from one of these to the other divided by the distance between them. It, however, will be much easier when we use a finite-difference grid to estimate this quantity.

The theta of the option is the sensitivity of the option price to time, assuming that the asset price does not change. Again, this is easier to calculate from a finite difference grid. An obvious choice for the discrete time definition of theta is to interpolate between V^+ and V^- to find a theoretical option value had the asset not changed and use this to estimate:

$$\Gamma = \frac{\partial V}{\partial t} \tag{B.225}$$

which results in:

$$\Gamma = \frac{\frac{1}{2}\left(V^+ - V^-\right)}{\delta t} \tag{B.226}$$

Estimating the other type of greeks, the ones involving differentiation with respect to parameters, requires performing a second binomial calculation. The calculation of the option's vega is presented here for illustration. The vega is the sensitivity of the option value to the volatility:

$$\text{Vega} = \frac{\partial V}{\partial \sigma} \tag{B.227}$$

Suppose we want to find the option value and the vega when the volatility is 20%. The most efficient way to do this is to calculate the option price twice, using a binomial tree, with two different values of σ. Calculate the option value using a volatility of $\sigma \pm \varepsilon$, for a small number ε; call the values you find $V_\pm$. The option value is approximated by the average value:

$$V = \frac{1}{2}\left(V_+ - V_-\right) \tag{B.228}$$

and the vega is approximated by:

$$\text{Vega} = \frac{V_+ - V_-}{2\varepsilon} \tag{B.229}$$

The idea can be applied to other greeks.

B.6.3 Implementation

(u, v, p) **to** (μ, σ)

The three constants (u, v, p) are chosen to give the binomial walk the same drift and standard deviation (μ, σ) as the stock we are trying to model. Our starting point, the lognormal random walk

$$dS = \mu\, S\, dt + \sigma\, S\, dX \tag{B.230}$$

has the solution:

$$S(t) = S(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma\phi\sqrt{t}} \tag{B.231}$$

where ϕ is a standardized normal random variable.

For the binomial random walk to have the correct drift over a time period of δt we need (details can be found in [6]):

$$puS + (1 - p)vS = S \times E\left[e^{\left(\mu - \frac{1}{2}\sigma^2\right)\delta t + \sigma\phi\sqrt{\delta t}}\right] = S \times e^{\mu\delta t} \tag{B.232}$$

i.e.

$$pu + (1-p)v = e^{\mu\delta t} \tag{B.233}$$

Rearranging this equation gives us:

$$p = \frac{e^{\mu\delta t} - v}{u - v} \tag{B.234}$$

For the binomial random walk to have the correct variance we need (details can be found in [6]):

$$pu^2 + (1-p)v^2 = e^{(2\mu+\sigma^2)\delta t} \tag{B.235}$$

Having only two equations for the three parameters gives us one degree of freedom in this choice. As Equations B.234 and B.235 determine all the statistically important properties of the discrete random walk, the choice of the third equation is somewhat arbitrary. This degree of freedom is often used to give the random walk the further property that after an up and a down movement (or a down followed by an up) the asset returns to its starting value S. This gives us the requirement that:

$$v(uS) = u(vS) = S \tag{B.236}$$

i.e.

$$uv = 1 \tag{B.237}$$

This choice leads to a tree in which the starting asset price reoccurs every even time-step and which is symmetric about this price, Figure B.6(a). Another popular choice is:

$$p = \frac{1}{2} \tag{B.238}$$

which leads to a tree oriented in the direction of the drift, Figure B.6(b) since the probabilities of an up-jump and a down-jump are equal.

(a) A tree with $u \times v = 1$ (b) A tree with $p = \frac{1}{2}$
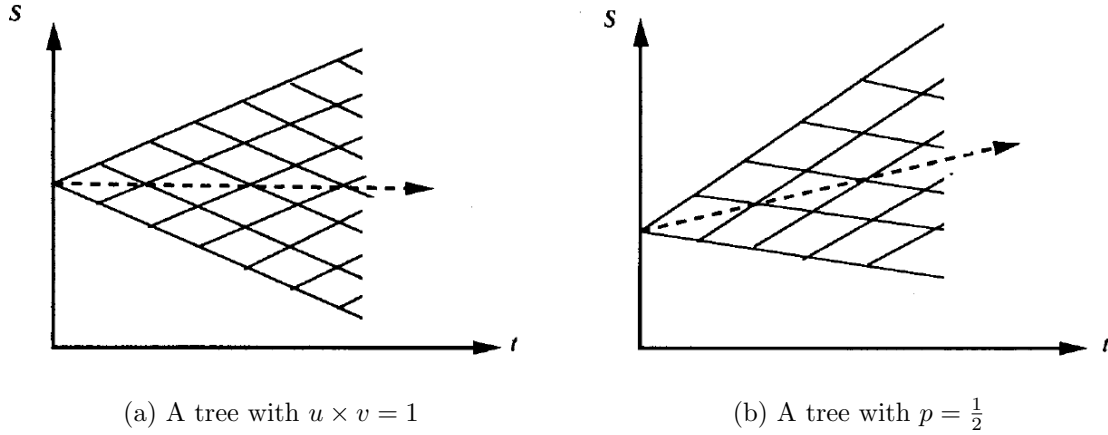
Figure B.6: Different binomial trees corresponding to the third equation

This note only consider the first choice of freedome; correspondingly, Equations B.234, B.235 and B.237 can be solved to give:

$$u = \frac{1}{2}\left(e^{-\mu\delta t} + e^{(\mu+\sigma^2)\delta t}\right) + \frac{1}{2}\sqrt{\left(e^{-\mu\delta t} + e^{(\mu+\sigma^2)\delta t}\right)^2 - 4} \tag{B.239}$$

The following approximations are good enough for most purposes:

$$u \approx 1 + \sigma\delta t^{1/2} + \frac{1}{2}\sigma^2\delta t \tag{B.240}$$

$$v \approx 1 - \sigma\delta t^{1/2} + \frac{1}{2}\sigma^2\delta t \tag{B.241}$$

$$p \approx \frac{1}{2} + \frac{\left(\mu - \frac{1}{2}\sigma^2\right)\delta t^{1/2}}{2\sigma} \tag{B.242}$$

If the above derivation is being used for pricing options, the μ should be replaced by r everywhere.

Data structures

The asset price and the option value are stored in two-dimensional array for clarity. The notation S_i^n means the asset price at the n-th time step and at the node i from the bottom, $0 \leq i \leq n$. In the lognormal model:

$$S_i^n = S \times u^i v^{n-i} \tag{B.243}$$

The notation V_i^n means the option value at the same node. Our goal is to find V_0^0 knowing the payoff, i.e. knowing V_i^N for all $0 \leq i \leq N$ where N is the number of time steps.

Payoff function

Assuming that we know the payoff function for our derivative security, and that it depends only on the values of the underlying asset at expiry, we are able to value it at expiry, i.e. time-step $N\delta t$.

If we are considering a put, we find that:

$$V_i^N = \mathbf{max} \left(E - S_i^N, 0 \right) \quad i = 0, 1, 2..., N \tag{B.244}$$

where E is the exercise price (strike). For a call, we find that:

$$V_i^N = \mathbf{max} \left(S_i^N - E, 0 \right) \quad i = 0, 1, 2..., N \tag{B.245}$$

Implementation

The implementation is quite straight-forward for the European-style exercise. First, we build a tree of possible asset prices; then find the values of the option at expiry using the **payoff_function**; finally, calculate the present values of the expected values of the option price under a risk-neutral random walk from expiry back until the present. The algorithm is presented below for the asset price $Asset$, the strike value E, the interest rate r, the volatility σ, the expiry T and the number of time step N. The code can be found here.

Algorithm 2: European-style exercise

```
1  begin
2      Asset, E, r, σ, T, N                                              /* Input */
3      Time step   δt = T/N                                     /* Initialisation */
4      Discount rate   dcr = exp(−rδt)
5      Calculate (u, v, p) following B.239, B.237 and B.234, respectively
6      2D-array of asset value S = 0 and option value V = 0
                                                        /* Calculate asset values */
7      Present asset price   S_0^0 = Asset
8      for  i = 1 → N  do
9          S_0^i = S_0^{i-1} × v
10         for  j = i → 1  do
11             S_j^i = S_{j-1}^{i-1} × u
12         end
13     end
                                                       /* Calculate option values */
14     for  i = 1 → N  do
15         V_i^N = payoff_function (S_i^N, E)
16     end
17     for  i = N-1 → 0  do
18         for  j = 0 → i+1  do
19             V_j^i = (p × V_{j+1}^{i+1} + (1 − p) × V_j^{i+1}) × dcr
20         end
21     end
22     Present option value   Option = V_0^0
23 end
```

The implementation for the American-style exercise is almost identical to that for the European-style exercise except a minor modification to address the early exercise to avoid arbitrage, i.e. the

option value should be larger or equal to the payoff at any time. If our theoretical value calculated following B.219 falls below the payoff then it is time to exercise.

$$V_j^i = \max\left(\frac{p \times V_{j+1}^{i+1} + (1-p) \times V_j^{i+1}}{\exp(r\delta t)}, \textbf{payoff\_function}\left(S_j^i, E\right)\right) \tag{B.246}$$

The algorithm for the American-style exercise is then only different from Algorithm 2 in calculating the intermediate option values, i.e. the option values is calculated by taking the maximum of the expected values and the payoff for early exercise at each time-step and the asset price.

Algorithm 3: American-style exercise

```
1  begin
2      for  i = N-1 → 0  do
3          for  j = 0 → i+1  do
4              Theoretical option value  tv = (p × V_{j+1}^{i+1} + (1 − p) × V_j^{i+1}) × dcr
5              Early exerice value  ev = payoff_function (S_j^i, E)
6              V_j^i = max (tv, ev)
7          end
8      end
9      Present option value   Option = V_0^0
10 end
```

Figure B.7 illustrate the option value for European- and American-put options corresponding to different numbers of time steps N. The input is as follows $Asset = 100, E = 110, r = 0.06, \sigma = 0.3, T = 0.25 \& 1$ (3-month and 1-year duration).

(a) Expiry in 3-month   (b) Expiry in 1-year
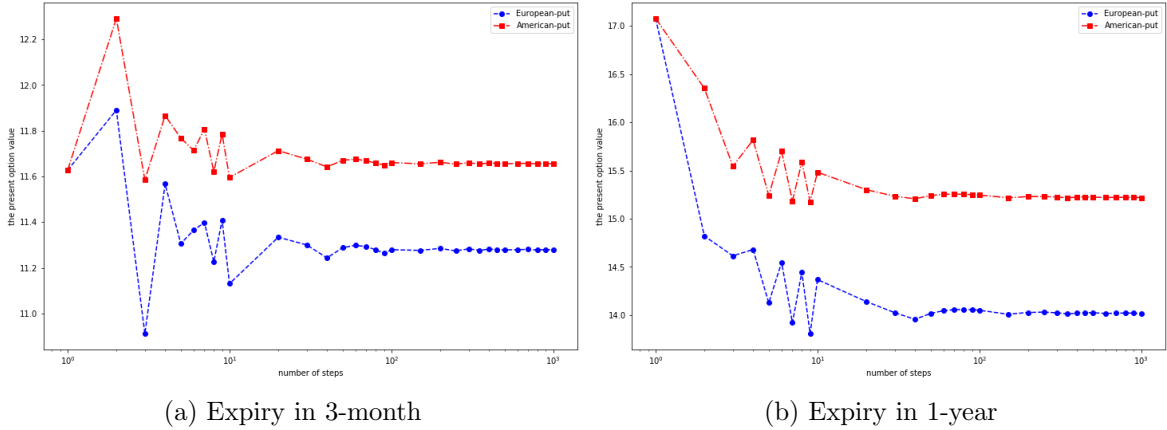
Figure B.7: Binomial model for European- and American-put options

Dividend yields

The binomial model can easily accommodate a constant dividend yield D_0 paid on the underlying. The effective risk-free growth rate of the asset becomes $\mu - D_0$ instead of μ, that is:

$$dS = (\mu - D_0)\ S\ dt + \sigma\ S\ dX \tag{B.247}$$

therefore we should replace μ by $\mu - D_0$ in the derivation of (u, v, p). For the case $uv = 1$, the formulations of (u, v, p) become:

$$A = \frac{1}{2}\left(e^{-(\mu-D_0)\delta t} + e^{(\mu-D_0+\sigma^2)\delta t}\right) \tag{B.248}$$

$$u = A + \sqrt{A^2 - 1} \tag{B.249}$$

$$v = A - \sqrt{A^2 - 1} \tag{B.250}$$

$$p = \frac{e^{(\mu-D_0)\delta t} - v}{u - v} \tag{B.251}$$

The only effect of continuous dividend yields is to modify the probability p and the jump sizes u and v. Therefore, the same algorithm given in the previous section for both the European- and American-option values can still be used by simply modifying the parameters (u, v, p) according to the above equations.

B.6.4 Continuous-time limit

The binomial model is a discrete-time model for the option pricing. From this model, we can develop a continuous-time model (the famous Black-Scholes model) by let $\delta t \to 0$.

We will start with the three approximation:

$$u \approx 1 + \sigma\sqrt{\delta t}$$
$$v \approx 1 - \sigma\sqrt{\delta t}$$
$$p \approx \frac{1}{2} + \frac{\mu\sqrt{\delta t}}{2\sigma}$$

as introduced at the beginning of this section.

Next, we write:

$$V = V(S, t) \tag{B.252}$$
$$V^+ = V(uS, t + \delta t) = V(S + (u - 1)S, t + \delta t) \tag{B.253}$$
$$V^- = V(vS, t + \delta t) = V(S + (v - 1)S, t + \delta t) \tag{B.254}$$

and apply the Taylor series approximation with small δt for V^+ and V^- following:

$$V(S + \delta S, t + \delta t) \approx V(S, t) + \delta t \frac{\partial V}{\partial t} + \delta S \frac{\partial V}{\partial S} + \frac{1}{2}\delta S^2 \frac{\partial^2 V}{\partial S^2} + ... \tag{B.255}$$

This series goes on for ever, but we only use the largest and most important terms, those which are required for the Black–Scholes analysis.

By substituting V^+ and V^- into the equation:

$$\Delta = \frac{V^+ - V^-}{(u - v) \times S}$$

to achieve:

$$\Delta \approx \frac{\partial V}{\partial S} \quad \text{as } \delta t \to 0 \tag{B.256}$$

Then, by substituting V, V^+ and V^- into the equation:

$$(1 + r\delta t)\,V = pV^+ + (1 - p)V^-$$

to find:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{B.257}$$

This is the Black–Scholes equation. Again, the drift rate μ has disappeared from the equation.

The famous Black–Scholes equation will be derived later using stochastic calculus rather than via the binomial model. The stochastic calculus derivation is far more useful than the binomial and being far easier to generalize.

B.6.5 Miscellaneous [FUTURE]

- Replicating portfolio
- Asset price process

B.7 Monte-Carlo solution for Geometric Brownian motion

B.8 A trading game

Appendix C

Basic financial simulation II

C.1   Time series analysis

C.2   Statistical techniques

C.3   Machine learning techniques

Appendix D

Basic financial simulation III

D.1 Volatility modelling

In general, the Black–Scholes model is very robust and does a decent job of pricing derivatives. One of the most important flaws in the model concerns the behavior of volatility. In fact, we do not know what volatility currently is as well as what it may be in the future; whereas the correct pricing of derivatives requires us to know what the volatility is going to be. For this reason, volatility analysis and modeling takes a prominent role in financial simulation.

D.1.1 Different types of volatility

Volatility is difficult to estimate, the best we can hope to do is to measure it statistically. But such a measure is necessarily backwards looking, and we really want to know what volatility is going to be in the future. There are a few types of volatility:

- Actual volatility

  ⋆ This is the measure of the amount of randomness in an asset return at any particular time. It is very difficult to measure, but is supposed to be an input into all option pricing models. In particular, the actual (or 'local') volatility goes into the Black–Scholes equation.
  ⋆ There is no 'timescale' associated with actual volatility, it is a quantity that exists at each instant, possibly varying from moment to moment.
  ⋆ Example: the actual volatility is now 20%... now it is 22.5%...

- Historical or realized volatility

  ⋆ This is a measure of the amount of randomness over some period in the past. The period is always specified, and so is the mathematical method for its calculation. Sometimes this backward-looking measure is used as an estimate for what volatility will be in the future. In pricing an option we are making an estimate of what actual volatility will be over the lifetime of the option. After the option has expired we can go back and calculate what the volatility actually was over the life of the option. This is the realized volatility.
  ⋆ There are two 'timescales' associated with historical or realized volatility: one short and one long.
  ⋆ Example: The 60-day volatility using daily returns. Perhaps of interest if you are pricing a 60-day option, which you are hedging daily.

- Implied volatility

  ⋆ The implied volatility is the volatility which when input into the Black–Scholes option pricing formulæ gives the market price of the option. It is often described as the market's view of the future actual volatility over the lifetime of the particular option. However, it is also influenced by other effects such as supply and demand.

⋆ There is one 'timescale' associated with implied volatility: expiration.

- Forward volatility

  ⋆ The adjective 'forward' can be applied to many forms of volatility, and refers to the volatility (whether actual or implied) over some period in the future.
  ⋆ Forward volatility is associated with either a time period, or a future instant.

D.1.2 To be continued

- Volatility estimation by statistical means
- Maximum likelihood estimation
- Different approaches to modelling volatility

D.2 Delta hedging

D.2.1 Classification of hedging strategies

'Hedging' in its broadest sense means the reduction of risk by exploiting relationships or correlation between various risky investments. The reason for hedging is that it can lead to an improved risk/return.

Two main classifications

Probably the most important distinction between types of hedging is between model-independent and model-dependent hedging strategies.

Model-independent hedging   An example of such hedging is put-call parity. There is a simple relationship between calls and puts on an asset (when they are both European and with the same strikes and expiries), the underlying stock and a zero-coupon bond with the same maturity. This relationship is completely independent of how the underlying asset changes in value.

Model-dependent hedging   Most sophisticated finance hedging strategies depend on a model for the underlying asset. The obvious example is the hedging used in the Black–Scholes analysis that leads to a whole theory for the value of derivatives. In pricing derivatives we typically need to at least know the volatility of the underlying asset.

To be continued

- Delta hedging
- Gamma hedging
- Vega hedging
- Static hedging
- Margin hedging
- Crash (Platinum) hedging

D.2.2 To be continued

- Implied versus actual volatilities
- Hedge with actual volatility
- Hedge with implied volatility
- Hedge with different volatilities

- Pros and cons of hedging with each volatility

- How does implied volatility behave

D.3   Not yet covered

- Multi-asset options

    - ⋆ Multidimensional lognormal random walk
    - ⋆ Measuring correlations
    - ⋆ Options of many underlying
    - ⋆ And many more

- Exotic and path-dependent options

- Barrier options

- Fixed-income products

- Swaps

- One-factor interest rate modeling

- Interest rate derivatives

Appendix E

Intermediate financial simulation

- Simulation of portfolio
- Simulation of processes with different levels of returns
- Simulation of processes with time-varying volatility
- Simulation of price processes
- Calibrating option pricing models

Appendix F

Advanced financial simulation

- Portfolio optimization with alternative risk measures
- Portfolio optimization with heuristics
- Portfolio optimization under Value-At-Risk
- Back-testing
- Back-testing portfolio strategies
- Econometric models

Appendix G

Mathematical tools for finance

G.1 Probability

G.1.1 Set theory

Under moving...

G.1.2 Permutation and Combination

Under moving...

G.1.3 Probability

Basic

Under moving...

Conditional probability and Bayes's formula

Under moving...

G.1.4 Random variables

Basis

Under moving...

Properties

Under moving...

G.1.5 Discrete random variables

Basis

Under moving...

Uniform discrete random variable

Under moving...

Binomial random variable

Under moving...

Geometric random variable

Under moving...

Negative binomial random variable

Under moving...

Hyper-geometric random variable

Under moving...

Poisson random variable

Under moving...

G.1.6 Continuous random variables

Basis

Under moving...

Expected value

If X is a continuous random variable having a pdf $f(x)$ then the **expected value** of X is given by:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx \tag{G.1}$$

Cumulative distribution function (cdf)

The **cumulative distribution function**, F, of the random variable X is defined for all real numbers b, by:

$$F(b) = \mathbf{P}\{X \leq b\} \tag{G.2}$$

Probability density function (pdf)

We say X admits a **probability density function** or **density** if:

$$\mathbf{P}\{X \leq b\} = F(b) = \int_{-\infty}^{b} f(x) dx \tag{G.3}$$

for some non-negative function f.

Continuous uniform distribution function

Under moving...

Normal random variable

X is a **normal random variable** with parameter μ and σ^2 if the density of X is given by:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty \tag{G.4}$$

Thus, the cdf of a **standard random variable**, i.e. one with mean $\mu = 0$ and variance $\sigma^2 = 1$, is given by:

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{\frac{-y^2}{2}} dy \tag{G.5}$$

The random variable X is *log-normally distributed* if for some normally distributed variable Y, $X = e^Y$, that is $\ln(X)$ is normally distributed.

Exponential random variable

Under moving...

Gamma distribution

Under moving...

G.1.7 Relationship between random variables

Basis

Under moving...

Independent random variables

Under moving...

Conditional distribution

Under moving...

Joint pdf of functions

Under moving...

Expected value of a function

Under moving...

Covariance

Under moving...

Coefficient of correlation

Under moving...

G.1.8 Miscellaneous

Central tendency

Under moving...

Moment generating functions

Under moving...

Central limit theorem

Under moving...

Special series

Under moving...

G.2   Stochastic processes

G.2.1   Basic

Under moving...

G.2.2   Bernoulli stochastic process

Under moving...

G.2.3   Poisson stochastic process

Under moving...

G.2.4   Discrete-time Markov chains

Discrete-time Markov chains

Not yet...

Classification of states

Not yet...

Steady-state behavior

Not yet...

Absorption probabilities and expected time to absorption

Not yet...

Random walk models

Not yet...

G.2.5   Continuous-time Markov chains

Continuous-time Markov chains

Not yet...

Brownian motion

Not yet...

Birth-Death process

Not yet...

G.3   Statistics

Not yet...

G.4   General numerical methods

G.4.1   Polynomial approximation

A polynomial approximation is an approximation of a curve or a function with a polynomial. In this section we will see how to find a family of polynomials widely used to estimate functions, known as Taylor polynomials or Taylor series.

Linear and quadratic approximation

Let's start with approximating a function using its tangent line - a key idea in Euler's method. Assume that we know the function value and its derivative at some point (e.g. $f(a), f'(a)$). We can use these to approximate $f(x)$ for other points x relatively near a by:

$$f(x) \approx P_1(x) = f(a) + f'(a)(x - a) \tag{G.6}$$

This is called the first order Taylor polynomial of $f(x)$.

To get a better approximation, we can try to use a quadratic polynomial. Another thing we can try is to find a polynomial that has the same value as the function at some point a and the same first and second derivatives there. We can do both by defining the second order Taylor polynomial for $f(x)$ near the point $x = a$ as follows:

$$f(x) \approx P_2(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2 \tag{G.7}$$

Higher-order approximation

By involving more derivatives and getting higher order polynomial we get better and better polynomial approximation. The n-th order Taylor polynomial for x near a is given by:

$$P_n(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2 + \sum_{i=3}^{n} \frac{f^{(i}(a)}{i!}(x - a)^i \tag{G.8}$$

As n goes to infinity, we have the Taylor series. The Taylor series is defined as follows:

> The Taylor series of a real or complex-valued function $f(x)$ that is infinitely differentiable at a real or complex number a is the power series:
>
> $$P_n(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n \tag{G.9}$$
>
> where $f^{(n)}(a)$ denotes the n-th derivatives of f evaluated at the point a. The derivative of order zero of f is defined to be f itself.

Examples of Taylor series for several functions at $a = 0$ are given below:

- Single function:

  ★ $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + R_4(x)$

  ★ $\ln(1 + x) = \sum_{n=0}^{\infty} (-1)^{n+1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} + R_4(x)$

  ★ $\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$

- A function of a function (derivation can be found in Section 14 of this link:

  ★ $\frac{e^x}{\cos(x)} = 1 + x + x^2 + \frac{2x^3}{3} + \frac{x^4}{2} + R_5(x)$

  ★ $\ln(1 + (\cos(x) - 1)) = -\frac{x^2}{2} - \frac{x^4}{12} - \frac{x^6}{45} + R_8(x)$

The notation $R_n(x)$ is presented in the next section.

Residual

To evaluate how good the approximation is (or how big the error is) we need to calculate the difference between the approximation value and the exact answer. When using the n-th order Taylor series $P_n(x)$ centered at a to approximate $f(x)$, the error is:

$$E = f(x) - P_n(x) \tag{G.10}$$

Often we are only interested in the magnitude of the error $|E|$. Fortunately, there is a simple formula for a bound on the size of the error $|E| = |R_n(x)|$, the bound is called the residual and is given as:

$$|R_n(x)| = \left| \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \right| \tag{G.11}$$

where c is between x and a. Normally, we will chose the value of c in the range $(a;x)$ to maximize the value of $f^{(n+1)}(c)$; or, just choose a value that we know is surely reasonably larger than $f^{(n+1)}(c)$ for all $c \in (a;x)$.

Example
Let's consider the magnitude of $R_5(x)$ of function $f = \cos(x)$ at $a = \pi$. We have $f^{(6)(x)} = -\sin(x)$ and $-1 \leq \sin(x) \leq 1$ hence:

$$|f(x) - P_5(x)| \leq |R_n(x)| = \left| \frac{f^{(6)}(c)}{6!}(x-\pi)^6 \right| = \frac{1}{6!}(x-\pi)^6$$

If we approximate $\cos(3)$ by the fifth order Taylor series centered at π then we have an error of at most:

$$\frac{1}{720}(\pi-3)^6 \approx 1.11 \times 10^{-9}$$

But if we approximate $\cos(1)$ also by the fifth order Taylor series centered at π then we have an error of at most:

$$\frac{1}{720}(\pi-1)^6 \approx 3.0 \times 10^{-4}$$

Multivariate Taylor series

The Taylor series may also be generalized to functions of more than one variable as done in Section 19 of this link. Since the generalized equation is quite lengthy, an example for a two-variable function is given here.

Let's a function $f(x,y)$ depend on two variables x and y, then the second order Taylor series at point (a,b) is:

$$f(a,b) + (x-a)f_x(a,b) + (y-b)f_y(a,b)$$
$$+ \frac{1}{2!}\left((x-a)^2 f_{xx}(a,b) + 2(x-a)(y-b)f_{xy}(a,b) + (y-b)^2 f_{yy}(a,b)\right) \tag{G.12}$$

where the subscripts denote the respective partial derivatives. A second order Taylor series expansion of a scalar-valued function of more than one variable can be written compactly as:

$$T(\mathbf{x}) = f(\mathbf{a}) + (\mathbf{x}-\mathbf{a})^T Df(\mathbf{a}) + \frac{1}{2!}(\mathbf{x}-\mathbf{a})^T \left\{D^2 f(\mathbf{a})\right\}(\mathbf{x}-\mathbf{a}) \tag{G.13}$$

where $Df(\mathbf{a})$ is the gradient of f evaluated at $\mathbf{x} = \mathbf{a}$ and $D^2 f(\mathbf{a})$ is the Hessian matrix. And from these two equations, you can deduce how the higher order Taylor series expansions look like for multivariate functions.

Example

Let's practice the Taylor series expansion for a two-variable function $f(x, y) = e^x \ln(1 + y)$ centered at $(a, b) = (0, 0)$.

First we need to calculate the partial derivatives of $f(x, y)$ and evaluate them at the origin (a, b).

$$
\begin{aligned}
f_x &= e^x \ln(1 + y) && \to f_x(0,0) = 0 \\
f_y &= \frac{e^x}{1 + y} && \to f_y(0,0) = 1 \\
f_{xx} &= e^x \ln(1 + y) && \to f_{xx}(0,0) = 0 \\
f_{yy} &= -\frac{e^x}{(1 + y)^2} && \to f_{yy}(0,0) = -1 \\
f_{xy} = f_{yx} &= \frac{e^x}{1 + y} && \to f_{xy}(0,0) = f_{xy}(0,0) = 1
\end{aligned}
$$

Substituting these values in to the general formula to get:

$$
e^x \ln(1 + y) = T(x, y) = y + xy - \frac{y^2}{2} + \dots
$$

Now if we assume $(x, y) = (0.2, 0.2)$ and use this formula to approximate $e^x \ln(1 + y)$ we will have a value of 0.220, the error is 2.69×10^{-3}.

G.4.2 Polynomial interpolation

This section describes the Lagrange polynomials which are widely used for polynomial interpolation. For a given set of points (x_j, y_j), the Lagrange polynomial is the polynomial of lowest degree that assumes at each value x_j there is a corresponding value y_j, so that the functions coincide at each point.

Given a set of $k + 1$ data points:

$$
(x_0, y_0), (x_1, y_1), \dots\dots, (x_k, y_k) \tag{G.14}
$$

where no two x_i are the same, the interpolation polynomial in the Lagrange form is a linear combination:

$$
L(x) = \sum_{j=0}^{k} y_j l_j(x) \tag{G.15}
$$

of Lagrange basis polynomials:

$$
l_j(x) = \prod_{o \le m \le k} \frac{x - x_m}{x_j - x_m} \tag{G.16}
$$

where $0 \le j \le k$ and $m \ne j$.

It can be seen that, for all $j \ne i$, we have $l_j(x_i) = 0$ whereas $l_j(x_j) = 1$. It follows that $y_j l_j(x_j) = y_j$ so at each point x_j we have $L(x_j) = y_j$ showing that L interpolates the function exactly.

Example

We wish to interpolate function $y = f(x)$ over the range $1 \le x \le 3$ given these three points:

$$
\begin{aligned}
x_0 &= 1 && \to f(x_0) = 1 \\
x_1 &= 2 && \to f(x_1) = 4 \\
x_2 &= 3 && \to f(x_2) = 9
\end{aligned}
$$

The interpolating polynomial is:

$$
L(x) = 1 \times \frac{x - 2}{1 - 2} \frac{x - 3}{1 - 3} + 4 \times \frac{x - 1}{2 - 1} \frac{x - 3}{2 - 3} + 9 \times \frac{x - 1}{3 - 1} \frac{x - 2}{3 - 2} = x^2
$$

G.4.3   Numerical integration

Definite integrals

The definite integral of a function $f(x)$ over an interval $[a; b]$ is the limit:

$$\int_a^b f(x)dx = \lim_{N \to \infty} \sum_{i=1}^{N} f(x_i^*)(x_i - x_{i-1}) \quad x_i^* \in [x_{i-1}; x_i] \tag{G.17}$$

and for each N:

$$x_0 = a < x_1 < .. < x_N = b \tag{G.18}$$

is a partition of $[a; b]$ with N subintervals and the values $x_i^* \in [x_{i-1}; x_i]$ chosen in each subintervals is arbitrary. This definition is actually based on a simple idea is that the value of the definite integral represents the net area under the curve of the graph of $y = f(x)$ on the interval $[a; b]$. The term "net" means that the areas above and under x-axis are positive and negative, respectively.

Although many integrals can be exactly solved by the fundamental theorem of calculus, most definite integrals are actually impossible to solve exactly. For example, the famous error function in probability:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{G.19}$$

is a definite integral which cannot be evaluated in closed form. Because of this reason, numerical integration becomes a good tool in mathematical modelling. The idea of numerical integration comes directly from the definition of integral which is using geometric shapes to approximate the area under the curve $y = f(x)$ to estimate definite integrals.

In this section, the simplest methods of numerical integration are represented:

- Riemann sums
- Newton-Cotes formula
- Monte-Carlo integration

The implementation of these algorithms can be found here.

Riemann sums

A Riemann sum is a certain kind of approximation of an integral by a finite sum. A Riemann sum of a function $f(x)$ over a partition:

$$x_0 = a < x_1 < x_2 < ... < x_{N-1} < x_N = b \tag{G.20}$$

is defined as a sum of the form:

$$\sum_{i=1}^{N} f(x_i^*)(x_i - x_{i-1}) \quad x_i^* \in [x_{i-1}; x_i] \tag{G.21}$$

where each value $x_i^* \in [x_{i-1}; x_i]$ in each subinterval is abitrary. Riemann sums are important because they provide an easy way to approximate a definite integral:

$$\int_a^b f(x)dx \approx \sum_{i=1}^{N} f(x_i^*)(x_i - x_{i-1}) \quad x_i^* \in [x_{i-1}; x_i] \tag{G.22}$$

As the product $f(x_i^*)(x_i - x_{i-1})$ is the area of a rectangle of height $f(x_i^*)$ and width $x_i - x_{i-1}$, a Riemann sum can be considered as the area of N rectangles with heights determined by the graph of $y = f(x)$.

The value x_i^* chosen in each subinterval is arbitrary, there are four specific choices giving us different types of Riemann sum:

- $x_i^* = x_{i-1}$: left Riemann sum
- $x_i^* = x_i$: right Riemann sum
- $x_i^* = \frac{x_{i-1}+x_i}{2}$: midpoint Riemann sum
- trapezoidal sum: technically not a Riemann sum but the average of the left and right Riemann sums

Those four methods are usually approached with partitions of equal size. The interval $[a; b]$ is divided into N subintervals, each of length:

$$\Delta x = \frac{b-a}{N} \tag{G.23}$$

Hence, the points in the partition will be:

$$a, a + \Delta x, a + 2\Delta x, ..., a + (N-1)\Delta x, b \tag{G.24}$$

Left Riemann sum

Approximating the function by its value at the left-end points we have:

$$\int_a^b f(x)dx = [f(a) + f(a + \Delta x) + ... + f(b - \Delta x)] + \text{Residual} \tag{G.25}$$

The error of this formula will be:

$$\text{Residual} = \left| \int_a^b f(x)dx - \sum f_{\text{left}} \right| \leq \frac{M_1(b-a)^2}{2N} \tag{G.26}$$

where M_1 is the maximum value of the absolute value of $f'(x)$ on the interval.

Right Riemann sum

Approximating the function by its value at the right-end points we have:

$$\int_a^b f(x)dx = [f(a + \Delta x) + f(a + 2\Delta x) + ... + f(b)] + \text{Residual} \tag{G.27}$$

The error of the right Riemann sum is identical to that of the left Riemann sum.

Midpoint Riemann sum

Approximating the function by its value at the midpoints of intervals we have:

$$\int_a^b f(x)dx = \left[f(a + \frac{\Delta x}{2}) + f(a + \frac{3\Delta x}{2}) + ... + f(b - \frac{\Delta x}{2}) \right] + \text{Residual} \tag{G.28}$$

The error of this formula will be:

$$\text{Residual} = \left| \int_a^b f(x)dx - \sum f_{\text{midpoint}} \right| \leq \frac{M_2(b-a)^3}{24N^2} \tag{G.29}$$

where M_2 is the maximum value of the absolute value of $f^{(2)}(x)$ on the interval.

Trapezoidal sum

In this case, the values of the function $f(x)$ on an interval are approximated by the values at the left- and right-end points, i.e. the area of the trapezoid formed by these two points. Following the same approach above we have:

$$\int_a^b f(x)dx = \frac{1}{2}\Delta x \left[f(a) + 2f(a + \Delta x) + 2f(a + 2\Delta x) + ...2f(b - \Delta x) + f(b) \right] + \text{Residual} \tag{G.30}$$

The error of this formula will be:

$$\text{Residual} = \left| \int_a^b f(x)dx - \sum f_{\text{midpoint}} \right| \leq \frac{M_2(b-a)^3}{12N^2} \qquad \text{(G.31)}$$

Higher dimensions

In two dimensions, the domain A can be divided into a number of cells A_i and each cell can be interpreted as having an "area" denoted by ΔA_i. The Riemann sum is then:

$$S = \sum_{i=1}^N f(x_i^*, y_i^*)\Delta A_i \qquad \text{(G.32)}$$

where $(x_i^*, y_i^*) \in A_i$.

Similarly, the domain V in three dimensions can be divided into a number of cells V_i and each cell can be interpreted as having a "volume" denoted by ΔV_i. The Riemann sum is then:

$$S = \sum_{i=1}^N f(x_i^*, y_i^*, z_i^*)\Delta V_i \qquad \text{(G.33)}$$

where $(x_i^*, y_i^*, z_i^*) \in V_i$.

Figure G.1 presents the Riemann sum of an integral:

$$\int_0^5 x^{\sin(x)} + x^{\cos(x)} - \sqrt{x}\,dx \qquad \text{(G.34)}$$

with a small number of interval, $N = 30$, to illustrate the approximation. This example can be found here.

(a) Left Riemann sum

(b) Right Riemann sum

(c) Midpoint Riemann sum

(d) Trapezoidal sum
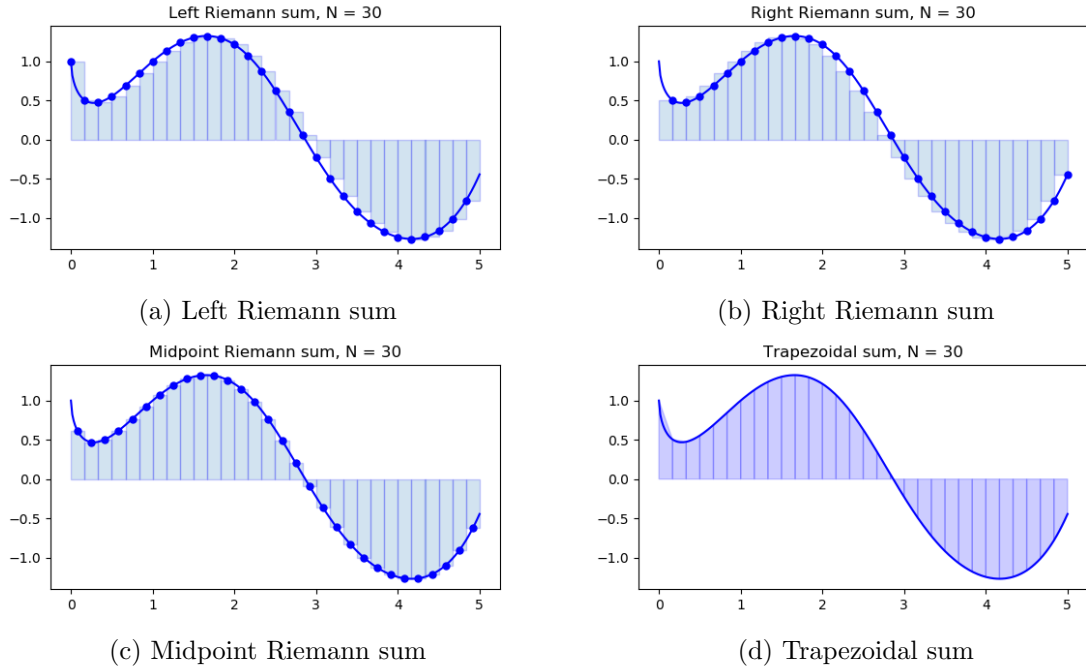
Figure G.1: Riemann integration

Newton-Cotes formulas

Newton–Cotes formulas, also called the Newton–Cotes (quadrature) rules, are a group of formulas for numerical integration (also called quadrature) based on evaluating the integral at equally spaced

points. If it is required that the points at which the integral is evaluated are not equally spaced, then other methods such as Gaussian quadrature are probably more suitable.

It is assumed that the value of a function f defined on $[a; b]$ is known at equally spaced points x_i, for $i = 0, ..., n$, where $x_0 = a$ and $x_n = b$. There are two types of Newton–Cotes formulas, the "closed" type which uses the function value at all points, and the "open" type which does not use the function values at the endpoints. The Newton–Cotes formulas of degree n are stated as:

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} w_i f(x_i) \quad \text{"close"} \tag{G.35}$$

$$\int_a^b f(x)dx \approx \sum_{i=1}^{n-1} w_i f(x_i) \quad \text{"open"} \tag{G.36}$$

where $x_i = ih + x_0$ with $h = \frac{x_n - x_0}{b} = \frac{b-a}{n}$ is called the step size. The w_i are called weights and are derived from the Lagrange basis polynomials. They depend only on the x_i and not on the function f. Let $L(x)$ be the interpolation polynomial in the Lagrange form for the given data points $(x_0, f(x_0)), ..., (x_n, f(x_n))$, then:

$$\int_a^b f(x)dx \approx \int_a^b L(x)dx = \int_a^b \left(\sum_{i=0}^{n} f(x_i) l_i(x) \right) dx = \sum_{i=0}^{n} f(x_i) \underbrace{\int_a^b l_i(x)dx}_{w_i} \tag{G.37}$$

The Newton-Cotes formulas, in fact, cover some other rules such as Riemann midpoint, trapezoidal sum, Simpson's rule... The following table lists some of the Newton-Cotes formulas (the notation f_i is the shorthand for $f(x_i) = f(a + i\frac{b-a}{n})$):

Table G.1: Numerical integration with Newton-Cotes formulas

Closed Newton-Cotes formulas

Degree n	Step size h	Common name	Formula	Error bound
1	$b - a$	Trapezoid rule	$\frac{h}{2}(f_0 + f_1)$	$\frac{1}{12}h^3 f^{(2)}(\xi)$
2	$\frac{b-a}{2}$	Simpson's rule	$\frac{h}{3}(f_0 + 4f_1 + f_2)$	$\frac{1}{90}h^5 f^{(4)}(\xi)$
3	$\frac{b-a}{3}$	Simpson's 38 rule	$\frac{3h}{8}(f_0 + 3f_1 + 3f_2 + f_3)$	$\frac{3}{80}h^5 f^{(4)}(\xi)$

Open Newton-Cotes formulas

Degree n	Step size h	Common name	Formula	Error bound
2	$\frac{b-a}{2}$	Midpoint rule	$2hf_1$	$\frac{1}{3}h^3 f^{(2)}(\xi)$
3	$\frac{b-a}{3}$	Trapezoid rule	$\frac{3h}{2}(f_1 + f_2)$	$\frac{1}{4}h^3 f^{(2)}(\xi)$
4	$\frac{b-a}{4}$	Milne's rule	$\frac{4h}{3}(2f_1 - f_2 + 2f_3)$	$\frac{28}{90}h^5 f^{(4)}(\xi)$

Monte-Carlo integration

Monte-Carlo integration uses random sampling of a function to numerically compute an estimate of its integral. Suppose that we want to integrate the one-dimensional function $f(x)$ from a to b:

$$F = \int_a^b f(x)dx \tag{G.38}$$

We can approximate this integral by averaging samples of the function f at uniform random points within the interval. Given a set of N uniform random variables $X_i \in [a; b]$ with a corresponding PDF of $\frac{1}{b-a}$, the Monte Carlo estimator for computing F is:

$$\langle F^N \rangle = (b - a)\frac{1}{N-1} \sum_{i=0}^{N} f(X_i) \tag{G.39}$$

The random variable $X_i \in [a;b)$ can be constructed by warping a random number uniformly distributed between zero and one, $\xi \in [0;1)$:

$$X_i = a + \xi(b - a) \tag{G.40}$$

Using this construction, we can expand the estimator to:

$$\langle F^N \rangle = (b-a)\frac{1}{N}\sum_{i=0}^{N-1} f(a + \xi(b-a)) \tag{G.41}$$

Sine $\langle F^N \rangle$ is a function of X_i, it is also a random variable; and therefore, $\langle F^N \rangle$ is an approximation of F using N samples.

It is possible to show that:

$$E\left[\langle F^N \rangle\right] = F \quad \text{note that } pdf(F) = \frac{1}{b-a} \tag{G.42}$$

As we increase the number of samples N, the estimator $\langle F^N \rangle$ becomes closer approximation of F. Due to the *Law of Large Numbers*, in the limit we can guarantee that we have the exact solution:

$$Pr\left\{\lim_{N\to\infty}\langle F^N \rangle = F\right\} = 1 \tag{G.43}$$

We can estimate convergence of this estimation by determining the convergence rate of the estimators variance. It is possible to show that:

$$\sigma\left[\langle F^N \rangle\right] \propto \frac{1}{\sqrt{N}} \tag{G.44}$$

This means that we must quadruple the number of samples in order to reduce the error by half. Standard integration techniques exist which converge much faster in one dimension, that's why Monte-Carlo integration is not a preference in one-dimensional problem. However the convergence rate of basic techniques becomes exponentially worse with increased dimensions; whereas the basic Monte Carlo estimator above can easily be extended to multiple dimensions, Algorithm 4, and the convergence rate for Monte Carlo is independent of the number of dimensions in the integral. This makes Monte Carlo integration the only practical technique for many high dimensional integration problems. More details information on the Monte-Carlo integration can be found here and here.

Algorithm 4: Monte-Carlo integration

```
1  begin
2  |    x_min, x_max, y_min, y_max                                    /* Input */
3  |    Function f(x,y), Number of sample N
4  |    Set temporary variable tmp = 0                            /* Initialisation */
5  |    for i = 1 → N do
6  |    |    x = random.uniform ∈ [x_min; x_max]
7  |    |    y = random.uniform ∈ [y_min; y_max]
8  |    |    tmp += f(x,y)
9  |    end
10 |    estimator = (x_max − x_min)(y_max − y_min) tmp/N
11 end
```

Another way of implementing the Monte-Carlo integration, following closely a counting-approach and geometrical definition of the integral can be found in Algorithm 5. This algorithm is, however, harder to generalize into multiple dimensions.

Algorithm 5: Monte-Carlo integration by geometrical counting

```
 1 begin
 2    x_min, x_max                                                     /* Input */
 3    Function f(x), Number of sample N
 4    Finding bounds of f(x) = f_min, f_max                      /* Initialisation */
 5    Set counting variable count = 0, indicator vector z (can be zero-initialized here)
 6    for i = 1 → N do
 7        x = random.uniform ∈ [x_min; x_max]
 8        fx = f(x)
 9        ft = random.uniform ∈ [f_min; f_max]
10        if |ft| ≤ |fx| then
11            if ft > 0 and fx > 0 and ft < fx then
12                count += 1                              /* area over x-axis is positive */
13                z[i] = 1
14            end
15            if ft < 0 and fx < 0 and ft ≥ fx then
16                count -= 1                             /* area under x-axis is negative */
17                z[i] = -1
18            end
19        else
20            z[i] = 0
21        end
22    end
23    estimator = (x_max - x_min)(y_max - y_min) count/N
24    Plotting with the indicator vector z
25 end
```

Figure G.2 presents the Monte-Carlo integration following Algorithm 5 for an integral:

$$\int_{10}^{20} x^2 \sin(x)^3 dx \tag{G.45}$$

The results for a sample of 100.000 sample points following Algorithm 4 and 5 are -181.04 and -182.12, respectively. The results from WolframAlpha is -181.16. The approximation of this integral by the Riemann sum with 10.000 interval yields the result of -181.16. This example can be found here.
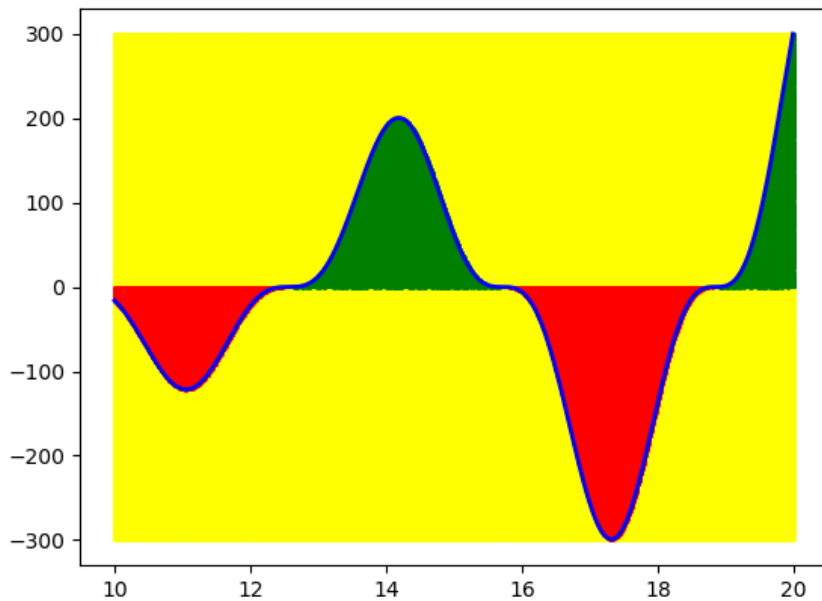
Figure G.2: Monte-Carlo integration

G.4.4 Numerical differentiation

Derivative and differentiation

The derivative of a function of a real variable measures the sensitivity to change of the function value (output value) with respect to a change in its argument (input value). Differentiation is the action of computing a derivative. The derivative of a function $y = f(x)$ of a variable x is a measure of the rate at which the value y of the function changes with respect to the change of the variable x. By definition we have:

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \tag{G.46}$$

When the limit exists, f is said to be differentiable at a. The derivative satisfies the property that:

$$\lim_{h \to 0} \frac{f(a+h) - (f(a) + f'(a) \times h)}{h} = 0 \tag{G.47}$$

which means:

$$f(a+h) \approx f(a) + f'(a)h \tag{G.48}$$

for f near a. This interpretation can be extended further easily. For example, if f is twice differentiable then:

$$f(a+h) \approx f(a) + f'(a)h + \frac{1}{2}f''(a)h^2 \tag{G.49}$$

If f is infinitely differentiable, then this is the beginning of the Taylor series for f evaluated at $a + h$ around a. Consequently, the simplest method to estimate the derivatives to an arbitrary order of accuracy is to use finite difference approximations which can be central, forward or backward.

First derivative

Following the finite difference approximations, the first derivative is calculated as follows:

- The forward difference formula with step size h

$$f'(a) \approx \frac{f(a+h) - f(a)}{h} \tag{G.50}$$

- The backward difference formula

$$f'(a) \approx \frac{f(a) - f(a-h)}{h} \tag{G.51}$$

- The central difference formula is the average of the forward and backwards difference formulas

$$f'(a) \approx \frac{f(a+h) - f(a-h)}{2h} \tag{G.52}$$

It can be seen that the finite difference approximations are based on the Taylor series, therefore the error in the finite difference approximations can be derived from Taylor's theorem:

- The error of the forward difference formula

$$\left| \frac{f(a+h) - f(a)}{h} - f'(a) \right| \leq \frac{hF}{2} \tag{G.53}$$

with $|f''(x)| \leq F$ for all $x in [a; a+h]$. This error formula is hold for the backward difference formula as well.

- The error of the central difference formula

$$\left| \frac{f(a+h) - f(a-h)}{2h} - f'(a) \right| \leq \frac{h^2 C}{6} \tag{G.54}$$

with $|f'''(x)| \leq C$ for all $x in [a - h; a + h]$.

An important consideration in practice when the function is calculated using floating-point arithmetic is the choice of step size h. If h is too small, the subtraction will yield a large rounding error due to cancellation. If h is too large, apparently, the accuracy reduces. For the numerical derivative formula evaluated at x and $x + h$, a choice for h that is large enough to avoid the large rounding error of $\sqrt{\varepsilon}x$, where the machine epsilon $\varepsilon = 2.2 \times 10^{-16}$ but small enough to avoid the secant error for optimum accuracy; such an option can be:

$$h = \sqrt{\varepsilon \left| \frac{f(x)}{f''(x)} \right|} \tag{G.55}$$

In the family of the finite difference approximations, you can achieve approximation with higher order of accuracy by involving more terms of the Taylor series or finite difference coefficient.

Higher derivative

Approximations of higher derivatives $f''(x), f'''(x), \ldots$ can easily be obtained in by using the Talor series expansion as above. For example, the central difference formula for the second derivative is as follows:

$$f(a+h) = f(a) + hf'(a) + h^2 \frac{f''(a)}{2!} + h^3 \frac{f'''(a)}{3!} + h^4 \frac{f^{(4)}(\xi_+)}{4!} \ldots$$

$$f(a-h) = f(a) - hf'(a) + h^2 \frac{f''(a)}{2!} - h^3 \frac{f'''(a)}{3!} + h^4 \frac{f^{(4)}(\xi_-)}{4!} \ldots$$

$$\rightarrow f(a+h) + f(a-h) = 2f(a) + h^2 f''(a) + h^4 \frac{f^{(4)}(\xi_+) + f^{(4)}(\xi_-)}{24}$$

$$\rightarrow f''(a) \approx \frac{f(a+h) - 2f(a) + f(a-h)}{h^2} \tag{G.56}$$

and the error term is:

$$h^2 \frac{f^{(4)}(\xi_+) + f^{(4)}(\xi_-)}{24} \tag{G.57}$$

in which $\xi_{+/-}$ is a small number between a and $a + h$.

Other forms with the same or higher order of accuracy of finite difference formulas can be found here.

Partial derivatives

Given a function of multiple variables, differentiation generalises in a simple way: differentiating the resulting function as a function of one variable while keeping other variable fixed. Finite difference formulas for partial derivatives follows the same approach. For example, the central difference formulas for a function of two variables are as follows:

$$\frac{\partial f}{\partial x}(a, b) \approx \frac{f(a+h, b) - f(a-h, b)}{2h} \tag{G.58}$$

$$\frac{\partial f}{\partial y}(a, b) \approx \frac{f(a, b+k) - f(a, b-k)}{2k} \tag{G.59}$$

$$\frac{\partial^2 f}{\partial x^2}(a, b) \approx \frac{f(a+h, b) - 2f(a, b) + f(a-h, b)}{h^2} \tag{G.60}$$

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) \approx \frac{f(a+h, b+k) - f(a+h, b-k) - f(a-h, b+k) + f(a-h, b-k)}{4hk} \tag{G.61}$$

Sometimes, the notation:

$$f(a+h, b+k) = f_{1,1} \qquad f(a+h, b-k) = f_{1,-1}$$
$$f(a-h, b+k) = f_{-1,1} \qquad f(a-h, b-k) = f_{-1,-1}$$

is used in practice, then the finite difference formula for partial derivatives can be shortened, for example:

$$\frac{\partial^3 f}{\partial x^2 \partial y} \approx \frac{f_{2,1} - 2f_{0,1} + f_{-2,1} - f_{2,-1} + f_{0,-1} - f_{-2,-1}}{8h^2 k} \tag{G.62}$$

> **Example**
>
> Suppose $u = u(x, y)$ is a function of two variables that we only know at grid points (x_i, y_j). The values of u at (x_i, y_j) are recorded in the matrix:
>
> $$u(x_i, y_j) = \begin{pmatrix} 5.1 & 6.5 & 7.5 & 8.1 & 8.4 \\ 5.5 & 6.8 & 7.8 & 8.3 & 8.9 \\ 5.5 & 6.9 & 9.0 & 8.4 & 9.1 \\ 5.4 & 9.6 & 9.1 & 8.6 & 9.4 \end{pmatrix}$$
>
> Suppose that the grid points are evenly spaced in both $x-$ and $y-$direction with $h = 0.5$ and $k = 0.2$. Then $u_y(x_2, y_4)$ can be approximated by the central difference as follows:
>
> $$u_y(x_2, y_4) \approx \frac{u_{2,5} - u_{2,3}}{2k} \approx \frac{8.9 - 7.8}{2 \times 0.2} = 2.75$$
>
> Similarly, $u_{xy}(x_2, y_4)$ can be approximated as:
>
> $$u_{xy}(x_2, y_4) \approx \frac{u_{3,5} - u_{3,3} - u_{1,5} + u_{1,3}}{4hk} \approx \frac{9.1 - 9.0 - 8.4 + 7.5}{4 \times 0.2 \times 0.5} = -2.0$$

Figure G.3 presents the first and second derivative of a function calculated by the numerical differentiation with the finite difference approach. The implementation can be found here.
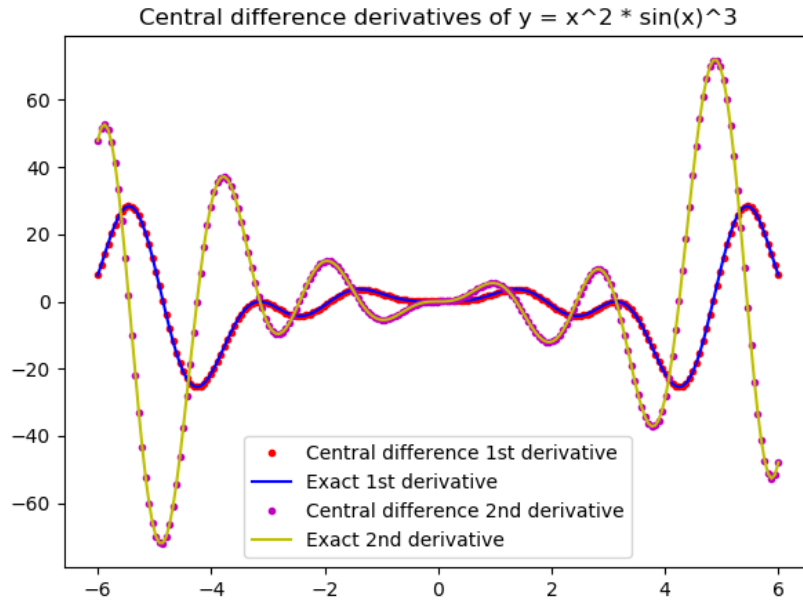
Figure G.3: First and second derivatives calculated with numerical differentiation

G.4.5 Root finding

In mathematics and computing, a root-finding algorithm is an algorithm for finding zeroes of continuous functions, i.e. a number/vector x for which $f(x) = 0$. Generally, the zeroes of a function cannot be computed exactly nor expressed in closed form, root-finding algorithms provide approximations to those. There are many algorithms for finding roots, this section only describes three basic algorithms, namely:

- Bisection method
- Newton-Raphson method
- Secant method

The implementation of these algorithms can be found here.

Bisection method

The simplest root-finding algorithm is the bisection method which can apply to any continuous function $f(x)$ on an interval $[a;b]$ where the value of the function $f(x)$ changes sign from a to b, i.e. $f(a) \times f(b) < 0$. The idea is simple: divide the interval in two, a solution must exist within one sub-interval, select the sub-interval where the sign of $f(x)$ changes and repeat until the root is found. Following this algorithm, we can easily estimate the absolute error of the approximation after N iteration as:

$$|x_{\text{exact}} - x_N| \leq \frac{b-a}{2^{N+1}} \tag{G.63}$$

From this equation, we can also calculate the minimum number of iteration required to achieve the absolute error less than a certain threshold ϵ as follows:

$$N \geq \frac{\log \frac{b-a}{\epsilon}}{\log(2)} - 1 \tag{G.64}$$

Although Bisection method is easy to implement and quite robust, generalization of Bisection method to a multivariate problem is not straightforward.

Newton-Raphson method

Newton–Raphson method is a root-finding algorithm which produces successively better approximations to the roots of a real-valued function.

Single-variable function

Given a single-variable function $f(x)$, the function's derivative $f'(x)$ and an initial guess x_0 for a root of $f(x) = 0$. If the initial guess is close enough, then:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{G.65}$$

is a better approximation of the root than x_0. Geometrically, $(x_1, 0)$ is the intersection of the x-axis and the tangent of the graph f at a point $(x_0, f(x_0))$, therefore, the improved guess is the root of the linear approximation at the initial point. The process is repeated as:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{G.66}$$

until $|x_{n+1} - x_n| \leq \epsilon$ in which ϵ is a certain predefined precision. Newton-Raphson method will usually converge with a quadratic rate provided that the initial guess is close enough to the unknown root and that $f'(x_0) \neq 0$.

Multivariate problem

Given k multivariate functions $\mathbf{f} = (f_1, f_2, ...f_k)$ of k variables $\mathbf{x} = (x_1, x_2, ...x_k)$. One can find the root $\mathbf{x}$ satisfying $\mathbf{f}(\mathbf{x}) = 0$ by following the iteration:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{J_f}(\mathbf{x}_n)^{-1} \mathbf{f}_{\mathbf{x}_n} \tag{G.67}$$

in which $\mathbf{J_f}$ is the Jacobian matrix which is defined as:

$$\mathbf{J_f} = \left[\frac{\partial \mathbf{f}}{\partial x_1} \cdots \frac{\partial \mathbf{f}}{\partial x_n} \right] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \tag{G.68}$$

Rather than computing the inverse of the Jacobian matrix, we can save time by solving the system of linear equations:

$$\mathbf{J_f}(\mathbf{x}_n) (\mathbf{x}_{n+1} - \mathbf{x}_n) = -\mathbf{f}(\mathbf{x}_n) \tag{G.69}$$

for the unknown $\mathbf{x}_{n+1} - \mathbf{x}_n$.

Minimization and maximization problems

Newton-Raphson method can be used to find a minimum or maximum of a function $f(x)$. The derivative is zero at a minimum or maximum, so local minima and maxima can be found by applying Newton-Raphson method to the derivative. The iteration becomes:

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)} \tag{G.70}$$

Practical considerations

Newton-Raphson method is an extremely powerful technique, in general the convergence is quadratic. However, there are a couple of difficulties with the method:

- Difficulty in calculating derivative of a function: we can use numerical approximation of the derivative or use secant method;
- Failure to converge to the root;
- Overshoot then divergence;
- Poor initial estimation then divergence;
- and some other minor problems.

This mean we need pay enough attention on using Newton-Raphson method. More information on the problems with Newton-Raphson method and how to (partially) overcome them can be found here.

Secant method

The secant method is very similar to the bisection method except instead of dividing each interval by choosing the midpoint the secant method divides each interval by the secant line connecting the endpoints. The secant method can be thought of as a finite-difference approximation of Newton-Raphson method although it is invented long before. The secant method always converges to a root of $f(x) = 0$ provided that $f(x)$ is continuous on $[a; b]$ and $f(a) \times f(b) < 0$.

The secant method can be birefed as follows: given a function $f(x)$ and the interval $[a_0; b_0]$, compute $f(x_0)$ where x_0 is given by the secant line:

$$x_0 = a_0 - f(a_0) \frac{b_0 - a_0}{f(b_0) - f(a_0)} \tag{G.71}$$

Then determine the next subinterval which should be $[a_0; x_0]$ if $f(a_0) \times f(x_0) < 0$ or $[x_0; b_0]$ if $f(x_0) \times f(b_0) < 0$. Repeat this procedure until the root is found.

Table G.2 presents the performance of the aforementioned root-finding algorithms in finding the root of function:

$$f(x) = \sqrt{x} * \log(x) - 5$$

in the interval of $[1; 10]$ with the threshold for the absolute error of $\epsilon = 10^{-10}$. The implementation of this illustration can be found here.

Table G.2: Performance of root-finding algorithms

Method	Iteration number	Root	Error
Bisection method	33	6.8017009	3.9394×10^{-11}
Newton-Raphson numerical	4	6.8017009	4.3432×10^{-13}
Newton-Raphson analytical	4	6.8017009	8.8818×10^{-16}
Secant method	11	6.8017009	4.4653×10^{-11}

G.5 Finite difference method

G.5.1 Introduction

Illustration with the heat or diffusion equation having the form:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x}$$ (G.72)

where u is the temperature, x is a spatial coordinate and t is time. Physical meaning of this equation is as follows: (1) consider the flow into and out of a small section of the bar; (2) the flow of heat along the bar is proportional to the spatial gradient of the temperature, hence the derivative of this, the second derivative of the temperature (right-hand side), is the heat retained by the small section; (3) this retained heat causes a change in the temperature (left-hand side).

- Local truncation error
- Numerical stability

G.5.2 Explicit

Not yet...

G.5.3 Implicit

Not yet...

G.5.4 Crank-Nicolson

Not yet...

G.6 Optimization

G.6.1 Unconstrained optimization

- Line search method
- Steepest descent method
- Newton-Raphson method
- Conjugate gradient method
- Quasi Newton-Raphson method
- Choice of stepsize

Not yet...

G.6.2 Constrained optimization

- Linear programming
- Quadratic programming

Not yet...

Index

Bibliography

[1] Paul Wilmott. *Introduces Quantitative Finance*. John Wiley & Sons Ltd, 2007.

[2] John C. Hull. *Options, futures, and other derivatives*. Pearson, 2015.

[3] J.C. Cox, S. Ross, and M. Rubinstein. Option pricing: a simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

[4] Paul Wilmott. *Paul Wilmott on Quantitative Finance*. John Wiley & Sons Ltd, 2006.

[5] Paul Wilmott, Jeff Dewynne, and Sam Howison. *Option pricing: Mathematical models and computation*. Oxford Financial Press, 1994.

[6] Paul Wilmott, Sam Howison, and Jeff Dewynne. *The mathematics of financial derivatives*. Cambridge University Press, 1995.

[7] Paul Wilmott. *Derivatives: the theory and practice of financial engineering*. John Wiley & Sons Ltd, 1998.

[8] Manfred Gilli, Dietmar Maringer, and Enrico Schumann. *Numerical Methods and Optimization in Finance*. Academic Press, 2019.

[9] Jesper Andreasen, Bjarke Jensen, and Rolf Poulson. Eight valuation methods in financial mathematics: the black–scholes formula as an example. *Mathematical Scientist*, 23:18–40, 1998.

[10] Peter E. Kloeden and Eckhard Platen. Numerical solution of stochastic differential equations. 1977.

[11] G. N. Milstein. Numerical integration of stochastic differential equations. 1994.

[12] Desmond J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43:525–546, 2001.

[13] Hagen Gilsing and Tony Shardlow. Sdelab: stochastic differential equations with matlab. 2006.

[14] Simon J. A. Malham and Anke Wiese. An introduction to sde simulation. 2010.

[15] Mustafa Bayram, Tugcem Partal, and Gulsen Orucova Buyukoz. Numerical methods for simulation of stochastic differential equations. *Advances in Difference Equations*, 2018:1–10, 2018.