Computational Finance

Raul Kangro

Spring 2020

Contents

Introduction

In the early 1970s, Fisher Black and Myron Scholes [1] made a major breakthrough by deriving a differential equation that must be satisfied by the price of any derivative security dependent on a non-dividend-paying stock. They used the equation to obtain values for European call and put options on the stock. Their work had a huge impact on how options were viewed in the financial world. Options are now traded on many different exchanges throughout the world and are very popular instruments for both speculating and risk management. Because of the popularity of derivative securities there is a great need for good and reliable ways to compute their prices.

In order to price an option one has to complete several steps:

1. specify a suitable mathematical model describing sufficiently well the behaviour of the stock market;

2. calibrate the model to available market data;

3. derive a formula or an equation for the price of the option of interest;

4. compute the price of the option.

Very often the last step requires the usage of some numerical methods because usually explicit formulas for the price of the option are not available.

In this course we pay very little attention to the first two steps and consider mainly the last two. More precisely, there are two main approaches to completing those steps, namely probabilistic approach (where option prices are expressed as expected values of some random variables) and Partial Differential Equations (PDE) approach (where option prices are expressed as solutions to certain differential equations). This course is mainly about the PDE approach, although some aspects of the probabilistic approach are also considered.

The lecture notes are self-contained and contain (together with the lab materials) all theoretical knowledge that is required for passing the course. There is a huge number of books where the aspects of the computational finance are discussed. For additional reading I recommend [4] for an alternative introduction to mathematical finance and finite difference methods, [5] for more extensive discussion of the theory and practice of computational finance, [2] for extensive treatment of Monte-Carlo methods in finance and [3] for details of Finite Element methods.

Chapter 1

Options on one underlying

This chapter is devoted to financial option contracts whose values depend on one underlying stock or other tradeable financial asset.

1.1 Definitions and examples

Unfortunately there is a lot of controversy in descriptions of financial options in textbooks and other literature (including Wikipedia and Investopedia and other sources). Usually such descriptions start by stating that financial options are about the right to buy or sell something in the future, but later examples of options are given which do not satisfy this condition (see, for example, the definition of digital options). Therefore we adopt a non-standard, but quite general definition of financial options which covers all known option types.

Definitsioon 1 *An* option *is a contract giving it's holder the right to receive in the future a payment which amount is determined by the behaviour of the stock market up to the moment of executing the contract.*

Option contracts are classified according to several characteristics including

- possible execution times (a fixed date vs a time interval),
- the number of underlying assets,
- how the value of option depends on the asset prices (depending on the price at the execution time vs a path dependent value of the asset prices).

In order to clarify the meaning of the definition, let us look at some examples.

Example 2 The right to buy 100 Nokia shares for 350 Euros exactly after 3 months (say, March 31st, 2020) .

This is an *European* (with fixed execution date) *Call* (the right to buy) option, which is equivalent to the right to receive after three months the sum of $100 \cdot \max(S(T) - 5, 0)$ Euros, where $S(T)$ denotes the price of a Nokia share at the specified date.

Example 3 The right to sell one Amazon.com share during next 6 months for $1900.

This is *American* (with a free execution time) *Put* (the right to sell) option.

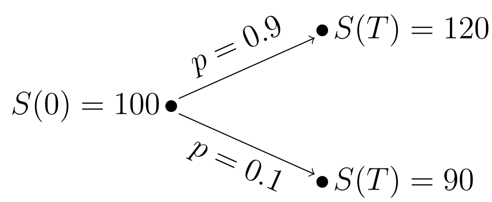
Example 4 The right to exchange after one year 10000 USD for Euros with the rate that is the average of the daily exchange rates in the one year period.

The last an *Asian* option that is an example of path dependent options.

1.2 Strange things about pricing options.

If an investor makes a decision about buying or selling a financial instrument, it is customary to consider the expected return and risk of the investment. Investment decision is usually made based on those quantities and investor's risk tolerance, so there is a different "right price" for each investor. It turns out that usual thinking models do not help to determine the right price for an option contract. In order to clarify this point, let us consider some simple toy models for stock price behaviour.

First, suppose that the at a future time T there are only two possible stock prices:



Assume for simplicity that the risk free interest rate is 0 (meaning that it is not possible to ear interest by depositing money in bank and it is possible to borrow money so that you have to pay back exactly the sum you borrowed). Let us consider an option to buy a at time T 10 shares of stock for 99. The value of this option at time T is 210 if $S(T) = 120$ (since you can buy 10 shares for 99 when the market price is 120) and it is worthless, if $S(T) = 0$. So buying the option seems to be a good investment possibility, if the price is not too high: the expected value of the option at time T is

$$0.9 \cdot 210 + 0.1 \cdot 0 = 189.$$

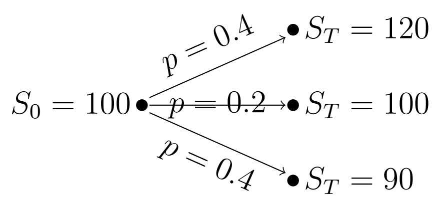
Note also that if the probability of $S(T) = 120$ is 0.1, then the expected value is only 21, so the expected value of the option depends strongly on the market probabilities. So it is natural to think that the fair price of the option should also depend on market probabilities.

But before deciding to buy the option we can consider alternative investment possibilities. It turns out that we can achieve exactly the same outcome by forming an investment portfolio consisting from a loan of 630 and 7 shares of stock: if $S_T = 120$, then the value of the portfolio is $7 \cdot 120 - 630 = 210$ and if $S_T = 90$, then the value is $7 \cdot 90 - 630 = 0$. The cost of forming the portfolio at $t = 0$ is 700-630=70 (we get 630 from the loan and have to add 70 of our own money to buy 7 shares for 100 each). So it is clear that at this market no sensible person pays more than 70 for the option. Moreover, if there exists any person willing to buy the option for more than 70, there is a possibility for anyone to earn money at the market without any risk of losing anything: one should just sell the option for the price and use 70 to set up the portfolio to cover the liabilities at time T. Such opportunities are called arbitrage opportunities and usually it is assumed that there are no arbitrage opportunities at the market.

If we allow portfolios with a negative number of shares (short selling) we can argue that the price of the option can not be less than 70. Otherwise we could buy the option, short sell 7 shares of stock for 100 each, deposit 630 in a bank and use the remaining money as we want. At time T we can in both cases use the bank deposit and the money we get by exercising the option to buy back 7 shares of the stock we borrowed earlier.

So the price of the option is completely determined by the market model. Moreover, the arguments we used did not depend on the probabilities of the up and down movements, so the option price does not depend on expected value and the risk of the option contract.

Let us consider now a similar market model with three different stock prices at time T:



It is easy to check that the price of the same option considered for the previous market model can not be larger than 70 (since the same portfolio as before requires 70 of initial investment and is worth at least as much as the option at time T for all possible values of $S(T)$. It is also possible to show that the price of the option can not be less than 10 (consider setting up the portfolio with one option, -7 shares of stock and a bank deposit of 690). Moreover, it is also possible to show that from the arbitrage principle it follows only that the value of the option is between 10 and 70 and if the option is traded for a price between those values, then arbitrage is not possible. This means that it is not possible to form a portfolio from the option, a stock holding and a loan/deposit so that the set-up cost of the portfolio is 0 but the value at time T is never negative and is positive with a positive probability (can you show it?). Consider now a second option that pays 20, if $S(T) = 100$ and 0 otherwise. If we consider this option separately from the first one, then from the arbitrage arguments if follows only that the price of the option is between 0 and 20. But if it is possible to both buy and sell any number of contracts of one of those two options then the price of the other one is completely determined. So there are strong consistency requirements between the prices of different options.

Exercise 1 *Suppose one can freely trade (both buy or short sell) the second option that pays 20, if $S(T) = 100$ and 0 otherwise and that the option price is 5 at time $t = 0$. Find a portfolio consisting of a number of the contracts of the second option, a stock holding and a loan/bank deposit (with 0 interest rate) so that in the case of the last market model the portfolio replicates (has exactly the same value for all possible stock prices at the time T) as the first option (the right to buy 10 shares at time T for 99 per share). Using this portfolio, determine the price of the first option at time $t = 0$. (Hint: the portfolio is determined by 3 unknowns – the number of option holdings, the number of shares and the loan. Write down 3 equations for the portfolio value to be equal to the pay-off of the first option at the time $t = T$ and solve for the unknowns. The value of the portfolio at time 0 has then to be equal to the price of the first option.)*

Based on the two simple models we can make the following conclusions:

- Naive pricing approaches (based on the expected return and risk) do not work.
- In the case of some market models the option price is determined completely by the model (and the "no arbitrage" condition)
- There are market models, for which the option prices are not determined completely but prices of different options have to be consistent with each other.

1.3 A stock market model, no arbitrage condition

In order to use mathematics in option pricing one has to start by specifying a model for stock price evolution and describing the conditions for trading.

1.3.1 Black-Scholes model.

A relatively simple but useful market model is so called Black-Scholes model, which assumes that the stock price changes according to the stochastic differential equation

$$dS(t) = S(t)(\mu(t)\,dt + \sigma(S(t), t)\,dB(t)), \tag{1.1}$$

where $S(t)$ is the stock price at time t, μ is the average growth rate of the stock price, the parameter σ characterizing the random variability of the stock price is called the volatility and B is the standard Brownian motion. Technically correct discussion of the meaning of the equation is out of scope of this course but intuitively it means for small non-intersecting time periods (t_{i-1}, t_i) we have

$$\begin{aligned} S(t_i) &\approx S(t_{i-1}) + S(t_{i-1})(\mu(t_{i-1})h_i + \sigma(S(t_{i-1}), t_{i-1})X_i) \\ &= S(t_{i-1})(1 + \mu(t_{i-1})h + \sigma(S(t_{i-1}), t_{i-1})X_i), \end{aligned}$$

where $h_i = t_i - t_{i-1}$ and $X_j \sim N(0, \sqrt{h_i})$, $j = 1, 2, \ldots, N$ and X_i are independent normally distributed random variables. This relation enables us to simulate sample trajectories according to the market model. The figure 1.1 shows 5 stock price trajec-

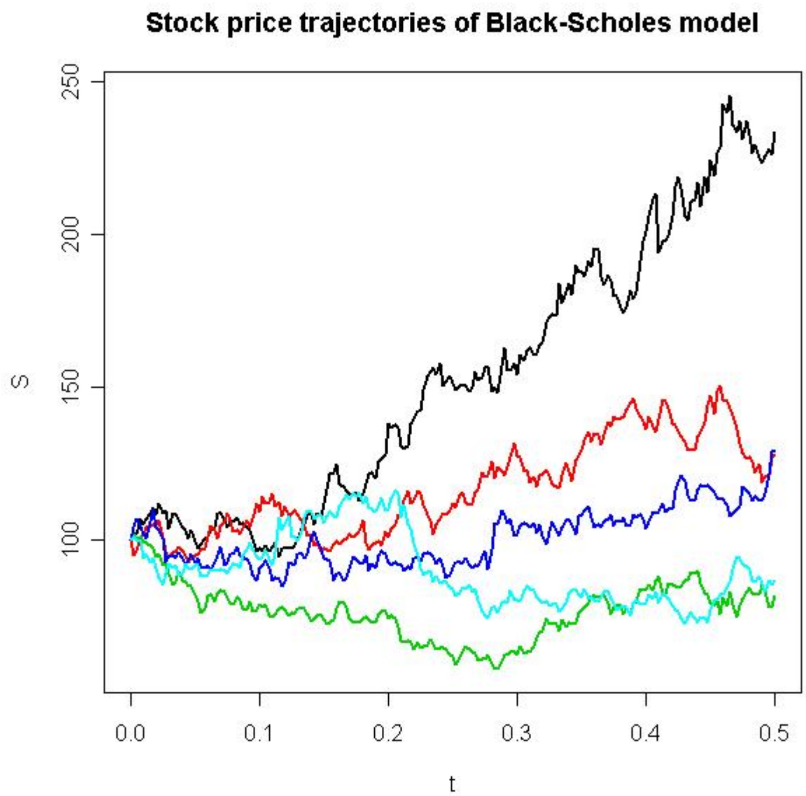


Figure 1.1: Sample trajectories of the stock price process following Black-Scholes model

tories illustrating the fact that future stock prices are random, so each time we compute a trajectory, we get a different one. In addition to the market model we make several additional simplifying assumptions:

- the risk free interest rate is a known constant r and is the same for lending and borrowing;
- it is possible to trade continuously and with arbitrarily small fractions of a stock;

- there are no transaction costs;
- it is not possible to make riskless profit by trading on the market.

It is clear, that some of the additional assumptions do not hold in practice and that the Black-Scholes model, at least with constant parameters μ and σ, is often not in a very good accordance with real market behaviour, but still it is a good starting point for mathematical modelling of the market behaviour.

1.3.2 Self-financing investment strategies

We call an *investment strategy* a rule for forming t each t in a period $[t_0, T]$ a portfolio consisting of a deposit $b(t)$ to a riskless bank account (if $b(t)$ is negative, then it corresponds to borrowing money) and of holding $\eta(t)$ shares of the stock. Both $b(t)$ and $\eta(t)$ may depend on the history up to time t (including the current value) of the stock prices but are not allowed to depend on the future values. An investment strategy is called *self-financing* if the only changes in the bank account after setting up the initial portfolio are the results of accumulation of interests of the same account, cash flows coming from holding the shares of the stock (eg dividend payments), or reflect buying or selling the shares of the stock required by changes of η, and if all cash flows that come from the changes of $\eta(t)$ are reflected in the bank account.

Let $X(t)$ denote the value of a self-financing portfolio at time t. Assume that the stock pays its holders continuously dividends with the rate D percent (realistic if the "stock" is a foreign currency, for usual stocks $D = 0$). Then in an infinitesimally small time interval dt the value of a self-financing portfolio changes according to the equation

$$dX(t) = r \cdot (X(t) - \eta(t)S(t))\,dt + D\eta(t)S(t)\,dt + \eta(t)\,dS(t). \tag{1.2}$$

The first term on the right hand side corresponds to the condition that all money that is not invested in the stock, is deposited to (or borrowed from) a bank account and bears the interest with the risk free rate r, the second term takes into account dividends and the last term reflects the change in the value of the portfolio coming from the change in the stock price. The value of a self-financing portfolio at any time $t > t_0$ is determined by the initial value $X(t_0) = X_0$ and the process $\eta(t),\ t \in [t_0, T]$.

Since nobody can borrow infinitely large sums of money, only such investment strategies for which the value of the portfolio is almost surely bounded below by a constant, are allowed.

1.3.3 No arbitrage condition.

In general, no arbitrage assumption states that it is not possible to make risk free profits by investing in the market. More precisely, it should not be possible to form a portfolio such that it does not cost any money today, the value of the portfolio is never negative during it's lifetime and has a positive value with nonzero probability at some future date. We need a corollary of the general no arbitrage condition.

Lemma 5 *(No arbitrage condition) If a self-financing portfolio produces exactly the same cash flows as holding an option, then the initial value of the portfolio and the option price have to be equal.*

Proof. If the price of the option is higher then we sell the option, form the self-financing portfolio and some money will be left for us to spend without any risk. If the option price is lower, then we buy the option and use the opposite investment strategy (having $-\eta(t)$ shares at time t). Again some money will be left over and we can spend it without any risk. Since such possibilities should not exist on a real market (at least for long), the option price and the initial value of the portfolio have to be the same.□

1.4 Itō's formula and Monte-Carlo method for pricing European options

We have specified a stochastic differential equation for the stock price evolution but it is not enough. We want also to consider functions of the stock price and differentiate them with respect to time. It turns out, that in the case of stochastic variables the usual rules of calculus do not hold and we need new differentiation rules (stochastic calculus).

1.4.1 Itō's Formula.

The following result proved by Japanese mathematician Kiyosi Itō in 1942, is of great importance in the theory of mathematical finance.

Lemma 6 Itō 's formula *Assume that $f(y,t)$ is a twice differentiable function of two variables and that a stochastic process Y satisfies the stochastic differential equation*

$$dY(t) = \alpha(t)\,dt + \beta(t)\,dB(t),$$

where α and β are continuous processes and B is the Brownian motion. Then

$$df(Y(t),t) = \left(\frac{\partial f}{\partial t}(Y(t),t) + \frac{\beta(t)^2}{2}\frac{\partial^2 f}{\partial y^2}(Y(t),t)\right)dt + \frac{\partial f}{\partial y}(Y(t),t)\,dY(t).$$

Example 7 Let us show that if μ and σ are constant then the process

$$S(t) = S(0)\cdot e^{(\mu-\frac{\sigma^2}{2})t+\sigma B(t)},\ t \in [0,T] \tag{1.3}$$

is a solution to the equation (1.1).

Denote

$$f(y,t) = e^{(\mu - \frac{\sigma^2}{2})t + \sigma y}, \; Y(t) = B(t),$$

then $S(t) = f(Y(t), t)$. Since

$$\begin{aligned}
\frac{\partial f}{\partial t}(y,t) &= (\mu - \frac{\sigma^2}{2}) f(y,t), \\
\frac{\partial f}{\partial y}(y,t) &= \sigma\, f(y,t), \\
\frac{\partial^2 f}{\partial y^2}(y,t) &= \sigma^2 f(y,t),
\end{aligned}$$

then, according to Itō's formula, we have

$$\begin{aligned}
dS(t) &= \Big((\mu - \frac{\sigma^2}{2}) S(t) + \frac{1}{2}\sigma^2 S(t))\, dt + \sigma\, S(t)\, dB(t) \Big) \\
&= S(t)(\mu\, dt + \sigma\, dB(t)).\square
\end{aligned}$$

Exercise 2 *Compute $df(B(t))$ for $f(y) = y^2$.*

Exercise 3 *Let $Y(t) = e^t \cos(B(t))$. Compute $dY(t)$.*

1.4.2 Estimating the parameters of BS model

There are two different approaches for estimating the parameters of the market model.

1. Fitting the market model to historical data.
2. Fitting the option prices derived from a market model to the actual prices of theoretical options.

Practitioners usually prefer the second approach since, according to the efficient market hypothesis, the traded options should have correct prices and it is highly desirable for an option pricing framework to produce correct prices to traded options. One has to use the first approach if the prices of traded options are not available or if we want to check the validity of our market model for a concrete stock. Let us discuss briefly both approaches.

Fitting the historical data

For simplicity, we assume that we have available n historical observation S_i, $i = 1, 2, \ldots, n$ of stock prices at equally spaced time steps (eg closing prices).

Using the Black-Scholes model (1.1) and Itō's formula, we get that

$$d(\ln S(t)) = (\mu(t) - \frac{\sigma(S(t),t)^2}{2})\, dt + \sigma(S(t),t)\, dB(t),$$

hence in the case of constant μ and σ we have for any time moments t_1 and $t_2 > t_1$ the equality

$$\ln \frac{S(t_2)}{S(t_1)} = (\mu - \frac{\sigma^2}{2})\,(t_2 - t_1) + \sigma(B(t_2) - B(t_1)).$$

Thus $x_i = \ln \frac{S_{i+1}}{S_i}$ are (if our assumption about the market model is correct) values of normally distributed iid random variables, $x_i \sim N((\mu - \frac{\sigma^2}{2})\Delta t, \sigma\sqrt{\Delta t})$, where Δt is the time interval between observations (usually measured in years). Therefore we can find estimates for μ and σ as follows:

$$\bar{\sigma} = \frac{\text{std}(x)}{\sqrt{\Delta t}},\ \ \bar{\mu} = \frac{\text{mean}(x)}{\Delta t} + \frac{\bar{\sigma}^2}{2}.$$

Unfortunately, if we test the normality of the logarithms of the quotients of the stock prices by some well-known statistical test, then it usually turns out that we have to reject the normality hypothesis.

As an example, let us consider the closing prices of a Cisco share. The price trajectory is given in the Figure 1.2. From the formulas above we get (assuming 252 working days

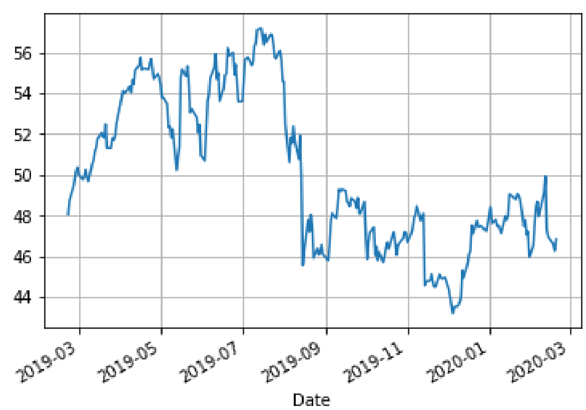


Figure 1.2: Adjusted closing prices of Cisco share, 21 February 2019 - 20 February 2020. Source of data: http://finance.yahoo.com/

per year)

$$\bar{\sigma} = 0.250,\ \ \bar{\mu} = 0.005.$$

Unfortunately it is not safe to use the Black-Scholes market model with constant parameters for pricing options on Cisco stock since the statistical test tell us that we can not believe the validity of the normality assumptions. For example, Shapiro-Wilk normality test (see Wikipedia!) gives for logarithmic returns the following result:

```
(0.9012666940689087, 8.953015412371812e-12)
```

Hence, the probability to get stock prices similar to the actual ones when Black-Scholes market model with constant coefficients holds, is extremely small (less than 0.00000000000895), so it is not reasonable to believe in the validity of this simple market model. If we do not want to assume that the parameters are constant, we may start with approximating the market model or the model for the logarithm of the stock price. Let us consider approximating the market model directly. By replacing the differentials by differences between the time moments t_{i-1} and t_i and taking the values of the functions appearing at the right hand side of the market model at the time t_{i-1} we get the approximate equalities

$$\frac{S(t_i) - S(t_{i-1})}{S(t_{i-1})} \approx \mu(t_{i-1})\,(t_i - t_{i-1}) + \sigma(S(t_{i-1}), t_{i-1})(B(t_i) - B(t_{i-1})).$$

Next, we introduce a finite number of unknown parameters $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$ and make an assumption how the functions $\mu = \mu_\theta$ and $\sigma = \sigma_\theta$ depend on those parameters. One way to find those parameters is to maximize the log-likelihood function: if Y_i are random variables with (conditional) probability density functions f_i, then the log-likelihood function of the values y_i is

$$\sum_i \ln f_i(y_i).$$

Recall that the probability density function of the normal distribution $N(\mu, \sigma)$ is

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}.$$

Since in our case the random variables $Y_i = \frac{S_i - S_{i-1}}{S_{i-1}}$ are, according to the approximate market model, normally distributed with mean $\mu_\theta(t_{i-1})\Delta t$ and standard deviation $\sigma_\theta(S_{i-1}, t_{i-1})\sqrt{\Delta t}$ we have to maximize the function

$$f(\theta) = -\sum_i \left(\frac{(Y_i - \mu_\theta(t_{i-1})\Delta t)^2}{2\sigma_\theta(S_{i-1}, t_{i-1})^2 \Delta t} + \ln \sigma_\theta(S_{i-1}, t_{i-1}) + \ln \sqrt{2\pi\Delta t} \right).$$

or minimize the negative of the function. Since f is usually a quite complicated function of the parameter vector θ, it may have several local extrema points, so one has to be careful in accepting an output of an optimization procedure as the solution of our parameter estimation problem.

Fitting the data of traded options

Starting from a market model we derive prices of various options. In the simplest cases we have explicit formulas, in more complicated cases we have to solve certain equations to get the option prices, but always we may thing that there is a function depending on market parameters that gives us the option prices. Suppose that we know the current prices $V_1, V_2, \dots, V_m$ of m different options and that $f_i(\theta)$ are the functions that give the option prices for (unknown) market parameters θ. Then we have m equations:

$$f_i(\theta) = V_i, \ i = 1, \dots, m.$$

Usually the number of unknown market parameters is much smaller than the number of available option prices, so the system of equations is solved in the least squares sense by minimizing the function

$$F(\theta) = \frac{1}{2} \sum_{i=1}^{m} (f_i(\theta) - V_i)^2.$$

Again there may be several local minima, so one should check carefully a possible candidate for the optimal solution.

Let us consider again the example of Black-Scholes market model with constant coefficients. It is known that in the case of this model the prices of European put and call options with exercise date T and strike price E at time t can be computed by Black-Scholes formulas as follows:

$$C(S,t,T) = Se^{-D(T-t)}\Phi(d_1) - Ee^{-r(T-t)}\Phi(d_2),$$
$$P(S,t,T) = -Se^{-D(T-t)}\Phi(-d_1) + Ee^{-r(T-t)}\Phi(-d_2),$$

where

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

and Φ is the cumulative distribution function of the standard normal distribution. Here D is the rate of proportional dividend payments, r is the risk free interest rate and σ is the volatility of the stock. So if we assume that r, D, t, T are fixed then we have functions that for any given volatility and exercise price give us the values of corresponding options. Since options are traded on the market, the prices of standard call and put options are available for several exercise prices. So we can try to pick the value of σ so that we get the observed prices from Black-Scholes formula. Moreover, if the Black-Scholes market model with constant coefficients holds, we should be able to find a value of σ that gives the observed prices for all strike prices for which we have data. Usually this is not the case: for each strike price we get a different value of σ (so called volatility smile effect). If this is the case, then we can be sure that Black-Scholes market model with constant volatility does not hold.

As a concrete example, let us consider finding the volatility from the market prices of call options for AstraZeneca PLC shares. Part of the data available for options expiring on July 17, 2020 was on March 6, 2020 as follows (source: `http:/finance.yahoo.com`, the price data was computed as the average of the bid and ask prices)

E	35.0	40.0	42.5	45.0	47.5	50.0	52.5	55.0	57.5	60.0
Price	13.25	8.75	6.35	5.0	3.5	2.4	1.325	0.975	0.425	0.2

The share price was at that moment \$47.52 and the share does not pay proportional dividends, so $D = 0$. Time to expiry was computed as the proportion of the working days from March 6 to July 17 (which is 93) to the total number of US working days in 2020 (which is 252) . For the risk free interest rate we used $r = 0.01$. Let us consider first the strike price $E = 35$, the graph of theoretical prices as a function of the volatility is given at figure 1.3. the horizontal line indicates the observed price \$13.25. From

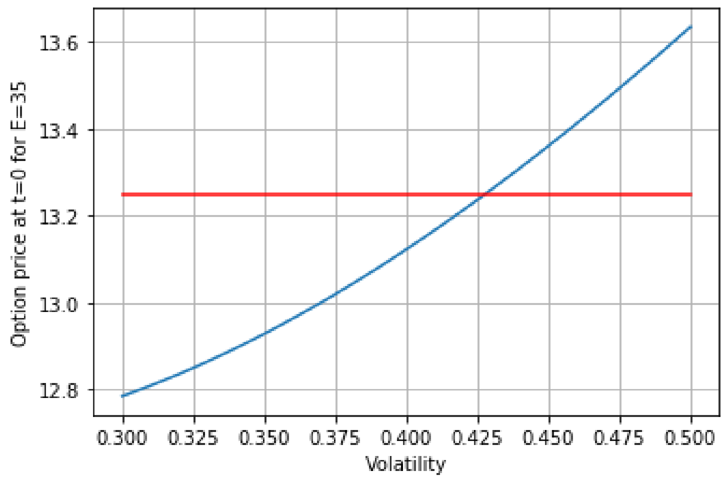


Figure 1.3: Call option price as a function of volatility for strike price $E = \$35$

the graph we see that there is a single volatility that gives us the observed price, the approximate value of the volatility is 0.425. By using a numerical solver we get that the volatility giving us the observed price is 0.427756.

Similarly we can find the volatilities that correspond to the other observed option prices for different strike prices. The results obtained are given in the figure 1.4. As we see, the volatilities that correspond to different strike prices are not equal. Thus either Black-Scholes market model with constant volatility does not hold, or there are arbitrage possibilities at the market. It is safer to assume that the model does not hold, so a better model is needed for pricing real options.

1.4.3 Monte-Carlo method for computing the prices of European options.

Suppose we know that an European option can be replicated (exactly the same outcome can be achieved by) a self-financing trading strategy. Let us recall, that the value of a portfolio corresponding to a self-financing trading strategy, satisfies the equation

$$dX(t) = r \cdot (X(t) - \eta(t)S(t))\,dt + D\eta(t)S(t)\,dt + \eta(t)\,dS(t).$$

Note that we can rewrite the equation in the form

$$d(e^{-rt}X(t)) = \eta(t)e^{-rt}S(t)((\mu(t) - r + D)\,dt + \sigma(S(t), t)\,dB(t)).$$

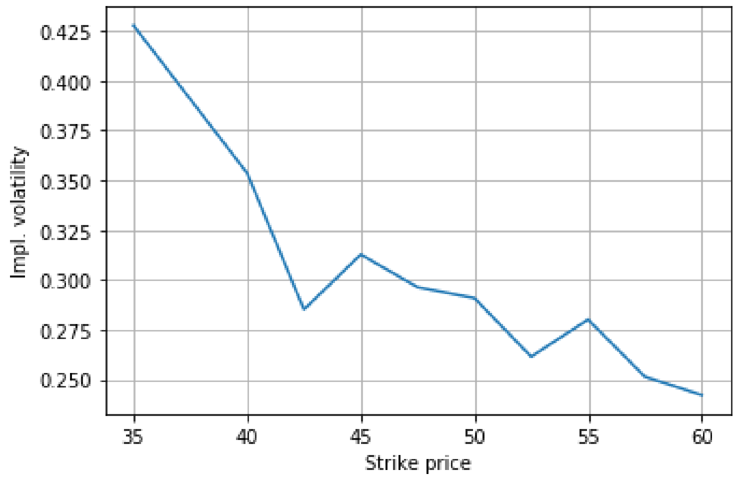


Figure 1.4: Implied volatilities for different exercise prices

Consider the case $\mu(t) \equiv r - D$. Then we have on the right hand side only the term with $dB(t)$ and, according to the theory of stochastic processes, the expected value of $e^{-rt}X(t)$ is the same for any t, ie $E(e^{-rt}X(t)) = X(0)$. Therefore, if an investment strategy replicates an option with payoff $p(S(T))$, then $X(T) = p(S(T))$ and hence the price of the option at time $t = 0$ is

$$X(0) = E(e^{-rT}p(S(T)),$$

hence the option price can be found by computing numerically (or analytically) the expected value in this case.

On the other hand, we prove later that under the assumptions we made about the stock market behavior every option can be replicated and the replication strategy does not depend on μ. Thus we can find the correct price by taking $\mu = r - D$ in the market model 1.1 and evaluating the expected value of the discounted payoff. Moreover, it can be shown that even when exact replication is not possible, option prices can still be expressed as expected values of some random variables.

One way to compute an expected value of a stochastic variable numerically is to generate n values of the variable and compute the average of the result. This is called Monte-Carlo method.

Lemma 8 (MC error) *Assume that $Y_1, Y_2, \ldots$ is a sequence of iid random variables with $EY_i = a$ and $DY_i = \sigma^2 < \infty$. Denote $H_n = \frac{\sum_{i=1}^n Y_i}{n}$. Then, for sufficiently large values of n we have*

$$P(|H_n - a| \geq \varepsilon) \approx 2\,\Phi(-\frac{\varepsilon\sqrt{n}}{\sigma})$$

and hence with probability $1-\alpha$ *we have*

$$|H_n - a| \le \frac{-\Phi^{-1}(\frac{\alpha}{2})\sigma}{\sqrt{n}} \tag{1.4}$$

where Φ *is the cumulative distribution function of the standard normal distribution.*

Proof. The statement of the Central Limit Theorem is as follows: if Y_i are iid random variables with mean a and standard deviation σ, then

$$P(\frac{\sum_{i=1}^n Y_i - na}{\sigma\sqrt{n}} \le t) \underset{n\to\infty}{\to} \Phi(t)\ \forall t \in \mathbf{R},$$

where Φ is the cumulative distribution function of the standard normal distribution. Note that the condition $|H_n - a| \ge \varepsilon$ is satisfied if $H_n - a \le -\varepsilon$ or $H_n - a \ge \varepsilon$ and these two possibilities are mutually exclusive. So

$$P(|H_n - a| \ge \varepsilon) = P(H_n - a \le -\varepsilon) + P(H_n - a \ge \varepsilon).$$

Let us consider the first term on the right hand side. By multiplying both sides of the inequality

$$H_n - a \le -\varepsilon$$

by n and dividing by $\sigma\sqrt{n}$ we get an equivalent inequality

$$\frac{\sum_{i=1}^n Y_i - na}{\sigma\sqrt{n}} \le -\frac{\varepsilon\sqrt{n}}{\sigma}.$$

Therefore we get from the Central Limit Theorem that for large n values

$$P(H_n - a \le -\varepsilon) \approx \Phi(-\frac{\varepsilon\sqrt{n}}{\sigma}).$$

The second term requires a little more work. Since the complement of the event $H_n - a \ge \varepsilon$ is the event $H_n - a < \varepsilon$, we have

$$P(H_n - a \ge \varepsilon) = 1 - P(H_n - a < \varepsilon)).$$

Let us assume for simplicity, that Y_i are continuous random variables, then

$$P(H_n - a < \varepsilon) = P(H_n - a \le \varepsilon).$$

Using the same arguments as in the case of the first term, we get

$$P(H_n - a \le \varepsilon) \approx \Phi(\frac{\varepsilon\sqrt{n}}{\sigma}),$$

thus

$$P(H_n - a \ge \varepsilon) \approx 1 - \Phi(\frac{\varepsilon\sqrt{n}}{\sigma}).$$

Finally, the property $\Phi(x)+\Phi(-x)=1$ of the cumulative distribution function of the standard normal distribution gives us

$$P(H_n - a \geq \varepsilon) \approx \Phi(\frac{-\varepsilon\sqrt{n}}{\sigma}).$$

Therefore we we have shown that

$$P(|H_n - a| \geq \varepsilon) \approx 2\,\Phi(-\frac{\varepsilon\sqrt{n}}{\sigma}).$$

The derivation of the final error estimate of the theorem is an exercise for the reader.
□

As we see, the error behaves like $\frac{1}{\sqrt{n}}$, so the convergence of the method is quite slow.

We saw earlier (see formula (1.3)) that if the Black-Scholes model with a constant volatility σ holds then the stock price $S(T)$ corresponding to the trend $\mu = r - D$ is given by

$$S(T) = S(0)e^{(r-D-\frac{\sigma^2}{2})\cdot T+\sigma B(T)}.$$

Generating the prices is easy: since $B(T)$ is normally distributed with variance T, we can just generate values of a random variable distributed according to the standard normal distribution, multiply those values with $\sqrt{T}$ and use the results for $B(T)$ in the formula above. Thus, in this case we can use MC method to compute the prices of any European options: we just generate the values of stock prices, compute the average of the discounted pay-off values and estimate the error of the result by (1.4). If we want to compute with a given accuracy, then we just have to keep generating the values of stock prices $S(T)$ until the error estimate is less than the desired accuracy. Very often it is not possible to generate $S(T)$ values that correspond exactly to the stochastic differential equation; then it is necessary to use some approximation methods. One such method is the Euler method, where we divide the interval $[0,T]$ into m equal subintervals and use the approximations (int the case of Black-Scholes market model)

$$S_{i+1} = S_i(1 + (r-D)\,\Delta t + \sigma(S_i,t_i)\sqrt{\Delta t}X_i), i = 0,\ldots,m-1,$$

where S_i are approximations to $S(i\Delta t)$, $\Delta t = \frac{T}{m}$ and $X_i \sim N(0,1)$. Instead of $S(T)$ we use S_m, thus we use Monte-Carlo method to compute an approximate value of $\hat{V}_m$, where

$$\hat{V}_m = E[e^{-rT}p(S_m)].$$

It is known that if p is continuous and has bounded first derivative (ie it is Lipshitz continuous), then

$$|V - \hat{V}_m| = \frac{C}{m} + o(\frac{1}{m}),$$

where C is a constant that does not depend on m and $m \cdot o(\frac{1}{m}) \to 0$ as $m \to \infty$. Thus, if we use S_m instead of $S(T)$ and use Monte-Carlo method, then the total error is

$$|V - \bar{V}_{m,n}| \leq |V - \hat{V}_m| + |\hat{V}_m - \bar{V}_{m,n}| \leq \frac{C}{m} + o(\frac{1}{m}) + |\hat{V}_m - \bar{V}_{m,n}|,$$

where $\bar{V}_{m,n}$ is computed by generating n different final stock prices S_m. The last term is the error of the Monte-Carlo method and can be estimated easily. So, in order to compute the option price V with a given error ε, we should choose large enough m (so that the term $\frac{C}{m}$ is small enough, for example less than $\frac{\varepsilon}{2}$) and then use MC method with large enough n so that the MC error estimate is also small enough (less than $\frac{\varepsilon}{2}$). There is one trouble: we do not know C. There are several methods for determining approximately it's value:

1. Fix a value of n and choose several values of m: $m_1, m_2, m_3, \ldots, m_k$ (very often one chooses $m_{i+1} = 2\,m_i$). Then if the values of m are large enough (meaning that we can ignore the $o(\frac{1}{m})$ term), we have

$$V_{m_i,n} \approx V + \frac{C}{m_i} + \varepsilon_i,\ \ i = 1, 2, \ldots, k,$$

   where ε_i are independent and approximately correspond to the same normal distribution. So we have a linear regression model for determining the values of C and the true option price V. Unfortunately the 95% confidence interval for V is usually too wide for practical purposes, but we can use the largest absolute value of the limits of the 95% confidence interval of C as an estimate $\bar{C}$ for the true value of $|C|$.

2. We use a value of m_1, define $m_2 = 2m_1$ and compute $V_{m_1,n}$ and $V_{m_2,n}$. Their difference satisfies

$$V_{m_1,n} - V_{m_2,n} \approx \frac{C}{2m} + \varepsilon_1 - \varepsilon_2,$$

   where ε_1 and ε_2 are Monte-Carlo errors for computing $\hat{V}_{m_1}$ and $\hat{V}_{m_2}$, respectively. From here we get (how?) that with probability $1-\alpha$ the estimate

$$|C| \leq \bar{C} = 2m_1(|V_{m_1,n} - V_{m_2,n}| + e_1 + e_2),$$

   where e_i are the probability $1-\alpha$ MC error estimates of the corresponding computations.

Exercise 4 *Prove the previous error estimate for $|C|$.*

After we have estimated $|C|$, we can choose m large enough so that the term $\frac{C}{m}$ is sufficiently small (for example $\frac{1}{2}$ of the desired accuracy) and then choose n large enough so that the MC error is also sufficiently small.

1.5 Partial differential equation for European options

One way to price options is to derive a partial differential equation (PDE) for the price of the options and then solve the equations either explicitly or numerically.

1.5.1 Derivation of Black-Scholes PDE.

Consider an European option with the payoff $p(S(T))$. Our procedure is as follows:

1. we'll make an assumption about what variables the option price depends on;
2. assume that the option can be replicated by a self-financing investment strategy and derive a PDE for the option price;
3. we'll show that the assumption was justified by using a solution to the PDE for constructing a self-financing portfolio that replicates the option.

It is clear that the option price depends on time (or on how much is left until the expiration date) and on the current stock price. So the first thing to try is to assume that the option price is a function of those two variables, ie the price at time t is $v(S(t), t)$.

Assume that the function v is sufficiently smooth (meaning differentiable) for using Itō's lemma. Assume also that there exists a self-financing investment strategy that replicates the option, then the price of the option at any time should be equal to the value of the portfolio at that time, $v(S(t), t) = X(t)$. Let $\eta(t)$ be the number of shares at time t that determines (with the initial value X(0)) the self-financing strategy.

We know that (see 1.2)

$$dX(t) = (r\,X(t) - (r - D)\,\eta(t)S(t))\;dt + \eta(t)\,dS(t)$$

and according to Itō's formula we have

$$d(v(S(t), t) = \left(\frac{\partial v}{\partial t}(S(t), t) + S(t)^2 \frac{\sigma(S(t), t)^2}{2} \frac{\partial^2 v}{\partial s^2}(S(t), t)\right)\,dt + \frac{\partial v}{\partial s}(S(t), t)\,dS(t).$$

As, according to our assumptions we have $v(S(t), t) = X(t)$, the expressions for $dX(t)$ and $d(v(S(t), t)$ should also be equal. Thus, we should have

$$\eta(t) = \frac{\partial v}{\partial s}(S(t), t)$$

and

$$\frac{\partial v}{\partial t}(S(t), t) + \frac{S(t)^2 \sigma(S(t), t)}{2} \frac{\partial^2 v}{\partial s^2}(S(t), t) = r\,v(S(t), t) - (r - D)\,S(t) \frac{\partial v}{\partial s}(S(t), t).$$

The last equality is satisfied for all values of t and $S(t)$, if v is a function of two variables satisfying the partial differential equation

$$\frac{\partial v}{\partial t}(s, t) + \frac{s^2 \sigma^2(s, t)}{2} \frac{\partial^2 v}{\partial s^2}(s, t) + (r - D)s \frac{\partial v}{\partial s}(s, t) - r\,v(s, t) = 0.$$

Now we have derived a partial differential equations for the option price. It remains to show that we can indeed construct a replicating self-financing investment strategy for European options.

Theorem 9 *Let* $p : (0,\infty) \to [0,\infty)$ *be a locally integrable function,* r *the risk-free interest rate,* D *the rate of continuous dividend payment of the underlying stock and let* v *be the solution of the partial differential equation*

$$\frac{\partial v}{\partial t} + \frac{s^2\sigma^2(s,t)}{2}\frac{\partial^2 v}{\partial s^2} + (r-D)s\frac{\partial v}{\partial s} - rv = 0,\ 0 \le t < T,\ 0 < s < \infty \tag{1.5}$$

satisfying the final condition

$$v(s,T) = p(s), \quad 0 < s < \infty.$$

Assume that v *is twice differentiable in the region* $(0,\infty)\times[0,T)$ *and is bounded from below. Then the price of the European option with the exercise date* T *and payoff* $p(S(T))$ *at any time* $0 \le t \le T$ *is* $v(S(t),t)$ *and the option can be replicated with a self-financing investment strategy with the initial value* $X(0) = v(S(0),0)$ *and the stock holding* $\eta(t) = \frac{\partial v}{\partial s}(S(t),t)$*.*

Proof. Let X be the value of the portfolio corresponding to the self-financing investment strategy with the initial value $X(0) = v(S(0),0)$ and the stock holding of $\eta(t) = \frac{\partial v}{\partial s}(S(t),t)$. Then, according to Itō's Lemma we have

$$\begin{aligned} d(X(t) - v(S(t),t)) &= (rX(t) - r\eta(t)S(t) + D\eta(t)S(t))\ dt \\ &\quad \left(-\frac{\partial v}{\partial t}(S(t),t) - \frac{S^2(t)\sigma^2(S(t),t)}{2}\frac{\partial^2 v}{\partial s^2}(S(t),t)\right)\ dt \\ &= r(X(t) - v(S(t),t))\,dt. \end{aligned}$$

Thus the difference $X(t) - v(S(t),t)$ satisfies an ordinary linear homogeneous differential equation with the zero initial condition and hence $X(t) = v(S(t),t)\ \forall t \in [0,T]$. In particular, we have $X(T) = v(S(T),T) = p(S(T))$, so the investment strategy replicates the option. This proves the lemma. □

The equation (1.5) is called Black-Scholes equation.

1.5.2 Extra reading: An alternative approach to option pricing

There are many market models for which it is not possible to replicate all options by self-financing portfolios, then the previous procedure for deriving a partial differential equation for the option pricing function does not work. A popular alternative is as follows:

1. It is postulated that the option price can be expressed as an expected value, for example
$$V = E[e^{-rT}p(S(T))],$$
where $S(T)$ follows a suitable stochastic differential equation.

2. It is shown that the expected value can be computed as a value of a function that satisfies certain partial differential (or partial integro-differential) equation.

One result, that enables us to relate expected values with solutions of partial differential equations is Feynman-Kac theorem.

Theorem 10 *(Feynman-Kac) Assume that $X(\tau)$ is a process that satisfies*

$$dX(\tau) = \alpha(X(\tau), \tau)\, d\tau + \beta(X(\tau), \tau)\, dB(\tau),\ \ t \le \tau \le T$$

together with the initial condition $X(t) = x$. Let $q(t, x)$ and $p(x)$ be sufficiently well-behaved functions (so that the the expectations below exist). Then

$$E[\exp(-\int_t^T q(X(\tau), \tau)\, d\tau)p(X(T))] = v(x, t),$$

where v is the solution of the partial differential equation

$$\frac{\partial v}{\partial t} + \alpha(x, t)\frac{\partial v}{\partial x} + \frac{\beta(x, t)^2}{2}\frac{\partial^2 v}{\partial x^2} - q(x, t)v = 0$$

satisfying the final condition $v(x, T) = p(x)$.

Proof. Exercise for those who have taken the Martingales course. (Hint: show that if v satisfies the equation, then $\exp(-\int_t^s q(X(\tau), \tau)\, d\tau)v(X(s), s)$ is a martingale).□

The previous theorem has generalizations to multidimensional processes X and for different type of stochastic differential equations for X.

Using this result it is easy to show that if an option price is computed according to the assumption above in the case $dS(t) = S(t)((r - D)\, dt + \sigma(S(t), t)\, dB(t))$, then the option pricing function satisfies the Black-Scholes equation.

1.5.3 Extra reading: classification and properties of partial differential equations

Definitsioon 11 *A partial differential equation with respect to an unknown function u is* linear*, if all it's terms are products of some function (or constant) not depending on u, and u or some partial derivative of u.*

Black-Scholes equation is a linear PDE.

In the case of linear equations a linear combination of any number of solutions is also a solution.

Definitsioon 12 *The* order *of a PDE is the highest order of derivative of the unknown function appearing in PDE.*

The order of Black-Scholes equation is 2.

A non-complete classification of second order equations of two variables is as follows:

1. If for each independent variable there is a second order term that contains a derivatives with respect to that variable and if the highest order terms

$$a(x,t)\frac{\partial^2 u}{\partial t^2} + b(x,t)\frac{\partial^2 u}{\partial x \partial t} + c(x,t)\frac{\partial^2 u}{\partial x^2}$$

   are such that at a point (x,t) we have

$$b^2(x,t) - 4a(x,t)c(x,t) > 0,$$

   then the PDE is *hyperbolic* at that point. An example is the wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0.$$

   An example of a solution of the wave equation is $u(x,t) = \sin(x-t)$ or, more generally, $u(x,t) = f(x-t)$, where f is an arbitrary twice differentiable function.

2. If for each independent variable there is a second order term that contains a derivatives with respect to that variable and if the highest order terms

$$a(x,t)\frac{\partial^2 u}{\partial t^2} + b(x,t)\frac{\partial^2 u}{\partial x \partial t} + c(x,t)\frac{\partial^2 u}{\partial x^2}$$

   are such that at a point (x,t) we have

$$b^2(x,t) - 4a(x,t)c(x,t) < 0,$$

   then the PDE is *elliptic* at that point. An example is the Laplace equation

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} = 0.$$

3. If the equation is of the form

$$a(x,t)\frac{\partial u}{\partial t} + b(x,t)\frac{\partial^2 u}{\partial x^2} + \text{lower order terms},$$

   then at the points where $a(x,t) \neq 0,\ b(x,t) \neq 0$ the equation is a parabolic equation at that point. An example is the heat equation

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0.$$

   An example of a solution of the heat equation is $u(x,t) = e^{-t}\sin(x)$.

If an equation is of the same type at every point, then we say that the equation is hyperbolic, elliptic or parabolic. Black-Scholes equation is a parabolic equation.

Some properties of parabolic equations

1. Parabolic equations are well-posed in one direction of time only: if the value of solution at $t = t_0$ is given, then the value of corresponding solution can be found only for $t > t_0$ (in the case when the coefficients of $\frac{\partial u}{\partial t}$ and $\frac{\partial^2 u}{\partial x^2}$ have the same sign) or for $t < t_0$ (in the case when the coefficients of $\frac{\partial u}{\partial t}$ and $\frac{\partial^2 u}{\partial x^2}$ have different signs).

2. Parabolic equations are smoothing equation, the smoothness (differentiability) of solution does not depend on the smoothness of the given initial (or final) condition but on the smoothness of the coefficients only.

3. Parabolic equations have infinite propagation speed: if the value of the given initial (or final) condition is changed in a neighborhood of one point only, then the change has some effect of the solution at all other times at every point.

1.5.4 Special solutions of Black-Scholes equation

Often it is useful to try to find some solutions of a given partial differential equation that are of some special form. For example, we may try to find out if BS equation has solutions that are linear in s for all time moments. So we substitute our guess

$$v(s,t) = \phi_1(t) + s\phi_2(t)$$

into the BS equation and try to find if for some functions phi_1 and ϕ_2 the equation is valid for all $s > 0, t > 0$. Substitution gives us

$$\phi_1'(t) + s\phi_2'(t) + (r - D)s\phi_2(t) - r(\phi_1(t) + s\phi_2(t)) = 0.$$

This equation has to be 0 for all values of variable s, therefore the coefficient of s has to be 0, thus we get

$$\phi_2'(t) - D\phi_2(t) = 0.$$

Now, if ϕ_2 satisfies this equation, we get that the terms without s should be equal to 0 for all t, hence we get the condition

$$\phi_1'(t) - r\phi_1(t) = 0.$$

By solving those differential equations together with conditions $\phi_1(T) = c_1,\ phi_2(T) = c_2$ we get that for both constant and non-constant volatility case the functions of the form

$$v(s,t) = c_1 e^{-r\,(T-t)} + c_2 e^{-D\,(T-t)} s,\ c_1, c_2 \in \mathbf{R} \tag{1.6}$$

are solutions of the equation 1.5. One consequence of this is so called Put-Call parity.

Lemma 13 *(Put-Call parity) Let $P(S,t,T)$ and $C(S,t,T)$ denote the values of the European put and call options with the exercise price E and expiration time T at time t if the stock price is $S(t) = S$. Then*

$$C(S,t,T) = P(S,t,T) + e^{-D\,(T-t)}S - E\,e^{-r\,(T-t)}.$$

Exercise 5 *Prove the previous lemma by using the uniqueness of the solution of the final value problem of BS equation.*

Remark. *If $D = 0$, then Put-Call parity relation follows directly from an arbitrage argument even without the assumption of no transaction costs.*
The special solutions are important for constructing effective numerical methods.

1.5.5 Transformation of Black-Scholes equation to the heat equation

Using the change of variables $v(s,t) = u(x,t)$, where $x = \ln s$, we can transform the equation (1.5) to the form

$$\frac{\partial u}{\partial t}(x,t) + \alpha(x,t)\frac{\partial^2 u}{\partial x^2}(x,t) + \beta(x,t)\frac{\partial u}{\partial x}(x,t) - r\,u(x,t) = 0, \tag{1.7}$$

where

$$\alpha(x,t) = \frac{\sigma^2(e^x,t)}{2},\ \beta(x,t) = r - D - \frac{\sigma^2(e^x,t)}{2}.$$

The corresponding final condition for the function u is

$$u(x,T) = p(e^x),\ -\infty < x < \infty. \tag{1.8}$$

The equation (1.7) is a backward parabolic partial differential equation. It turns out that if σ is a constant, then we can further transform the equation to the standard heat equation.
First, note that by defining $u(x,t) = e^{-r\tau}\tilde{u}(x,\tau)$, where $\tau = T - t$ we get a usual parabolic PDE which does not have a term without derivatives:

$$\frac{\partial \tilde{u}}{\partial \tau}(x,\tau) = \alpha\frac{\partial^2 \tilde{u}}{\partial x^2}(x,\tau) + \beta\frac{\partial \tilde{u}}{\partial x}(x,\tau).$$

Now the change of variables

$$\tilde{u}(x,\tau) = w(y,\eta),\ \eta = \alpha\tau,\ y = x + \beta\tau$$

gives us the equation

$$\frac{\partial w}{\partial \eta}(y,\eta) = \frac{\partial^2 w}{\partial y^2}(y,\eta).$$

This is the heat equation. It is known that the solution of the heat equation has a representation

$$w(y,\eta) = \frac{1}{2\sqrt{\pi\eta}} \int_{-\infty}^{\infty} e^{-\frac{(y-\xi)^2}{4\eta}} w(\xi,0)\, d\xi.$$

Taking into account that

$$w(y,0) = u(y,T) = v(e^y,T),$$

$$v(s,t) = u(\ln s,t) = e^{-r(T-t)}\tilde{u}(\ln s, T-t) = e^{-r(T-t)} w(\ln s + (r - D - \frac{\sigma^2}{2})(T-t), \frac{\sigma^2}{2}(T-t))$$

we can now express the solution of the original Black-Scholes equation in an integral form:

$$v(s,t) = \frac{e^{-r(T-t)}}{\sqrt{2\pi(T-t)}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(\ln s + (r-D-\frac{\sigma^2}{2})(T-t)-\xi)^2}{2\sigma^2(T-t)}} p(e^{\xi})\, d\xi \tag{1.9}$$

Using this form it is possible to derive explicit formulas for several options.

Exercise 6 *The transformations used above are not the only possible ones. Assume that the volatility σ is constant. Find a, b such that the function w defined by $u(x,t) = e^{ax+b\tau} w(\tau,x)$, where $\tau = T - t$, satisfies the partial differential equation*

$$\frac{\partial w}{\partial \tau}(\tau,x) = \frac{\sigma^2}{2}\frac{\partial^2 w}{\partial x^2}(\tau,x).$$

Exercise 7 *Consider the equation*

$$\frac{\partial v}{\partial t}(x,y,t) + x^2\frac{\partial^2 v}{\partial x^2}(x,y,t) + y^2\frac{\partial^2 v}{\partial y^2}(x,y,t) = 0.$$

Let us look for a solution of the form $v(x,y,t) = yw(z,t)$, where $z = \frac{x}{y}$. Find the partial differential equation that must be satisfied by the function $w(z,t)$. In the final equation only the variables z and t should be used!

1.6 Finite difference methods for Black-Scholes equation

A popular class of numerical methods for solving partial differential equations is finite difference methods, where approximate values of solutions at certain rectangular mesh points are found by replacing partial derivatives in the PDE by finite difference approximations (using only the values at the mesh points) and solving the resulting system of equations.

1.6.1 The idea of finite difference methods

Let u be the solution of the problem (1.7), (1.8). Since in a numerical computation we can find only finitely many numbers, we may try to compute a table of the approximate values of u. For this we fix the minimal and maximal values of x we are interested in (say x_{min} and x_{max}), the number m of subintervals of the time period $[0, T]$ (this is the number of time steps we use to get from 0 to T) and the number of subintervals n we use in the x direction.

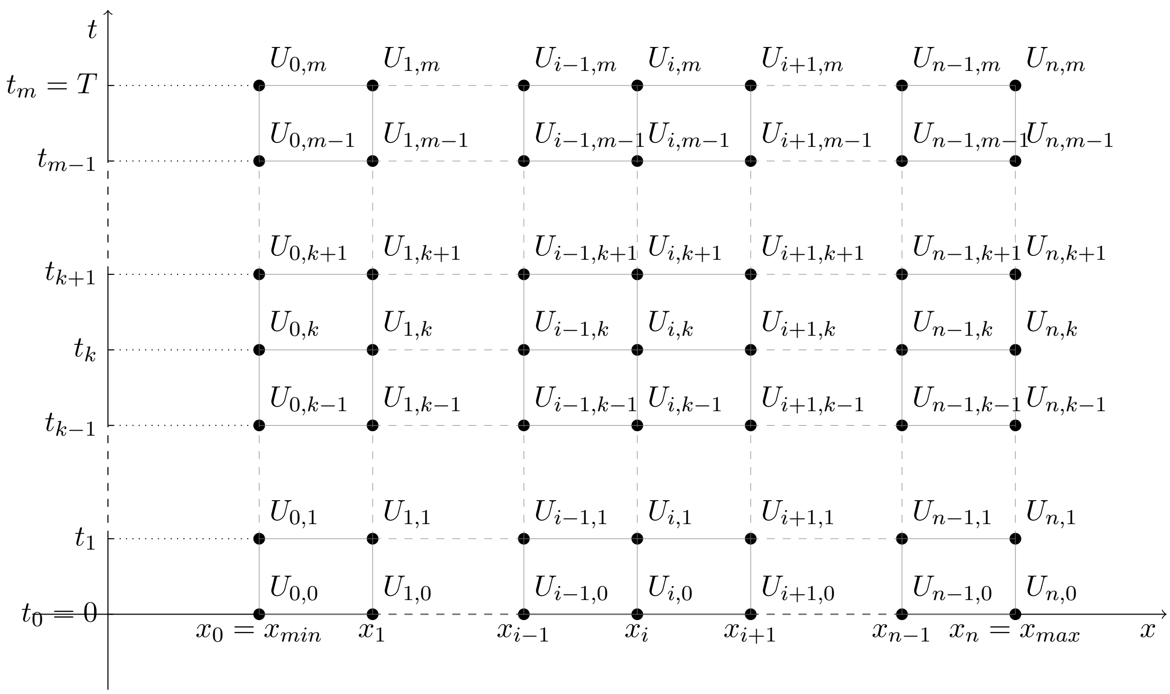
Denote

$$\Delta t = \frac{T}{m}, \quad \Delta x = \frac{x_{max} - x_{min}}{n}$$

and define

$$t_k = k\,\Delta t,\ k = 0, \ldots, m;\ x_i = x_{min} + i\,\Delta x,\ i = 0, \ldots, n.$$

Our aim is to find approximately the values $u_{ik} = u(x_i, t_k)$, ie we want to form a $(m+1) \times (n+1)$ table of approximate values. Let U_{ik} be the approximate values we write in the table. The notations are illustrated below.



The values of u at $t = T$ are given by the final condition (1.8). Therefore we have

$$U_{im} = p(e^{x_i}), \quad i = 0, 1, \ldots, n. \tag{1.10}$$

The values corresponding to $x = x_{min}$ and $x = x_{max}$ will be given by some boundary conditions discussed later.

In order to find the other values, we have to make use of the equation (1.7), where derivatives are replaced by numerical differentiation formulas. From the textbooks of numerical methods we can find the following approximate differentiation rules for a sufficiently smooth (meaning enough times continuously differentiable) function f:

$$f'(z) = \frac{f(z+h) - f(z)}{h} + O(h), \tag{1.11}$$

$$f'(z) = \frac{f(z) - f(z-h)}{h} + O(h), \tag{1.12}$$

$$f'(z) = \frac{f(z+h) - f(z-h)}{2h} + O(h^2), \tag{1.13}$$

$$f''(z) = \frac{f(z-h) - 2f(z) + f(z+h)}{h^2} + O(h^2), \tag{1.14}$$

where $O(h^q)$ denotes some function (which may be different in different formulas) that may depend on f, z satisfying the inequality $|O(h^q)| \leq const.h^q$ for all sufficiently small values of h. The first formula is called forward difference approximation, the second is backward difference approximation and the third is the central difference approximation of the derivative. The name of finite difference methods comes from replacing the derivatives (that are the limit of those formulas when h goes to 0) with approximations corresponding to some finite (small) values of h.

The same formulas can be used for approximating partial derivatives. For example, if we want to approximate $\frac{\partial u}{\partial x}(x,t)$, we consider the variable t fixed and use x in the role of z and Δx in the role of h in the above formulas, so applying for example (1.13) gives us

$$\frac{\partial u}{\partial x}(x,t) = \frac{u(x+\Delta x, t) - u(x - \Delta x, t)}{2\Delta x} + O(\Delta x^2).$$

We start by deriving an explicit finite difference method (meaning that the solution of the system of equations can be written out in an explicit form) for solving Black-Scholes PDE.

1.6.2 Explicit finite difference method

Let us derive a numerical scheme for the partial differential equation

$$\frac{\partial u}{\partial t}(x,t) = \alpha(x,\tau)\frac{\partial^2 u}{\partial x^2}(x,t) + \beta(x,t)\frac{\partial u}{\partial x}(x,t) + \gamma u(x,t),\ x \in \mathbf{R}, 0 \leq t < T \tag{1.15}$$

with the final condition

$$u(x,T) = u_0(x),\ x \in \mathbf{R}. \tag{1.16}$$

In the case of the transformed Black-Scholes equation (1.7) we have

$$\alpha(x,t) = \frac{\sigma^2(e^x,t)}{2},$$

$$\beta(x,t) = r - D - \frac{\sigma^2(e^x,t)}{2},$$

$$\gamma = -r$$

and

$$u_0(x) = p(e^x),$$

where p is the payoff function of the European option we are considering. Here we have assumed that the risk free interest rate r and the rate of continuously paid dividends D are constant but the volatility σ may depend on both time and the current stock price.

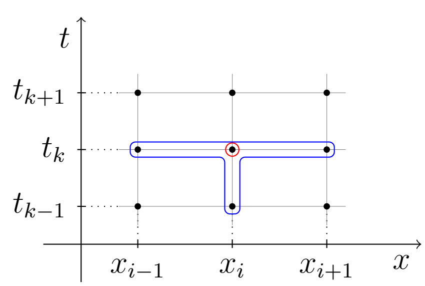
Derivation of the equations for the values U_{ik}.

After taking into account the final condition we still have $(n+1)\cdot m$ empty spaces in our table of approximate values. We get the values corresponding to $x = x_{min}$ and $x = x_{max}$) (or some additional equations corresponding to the values) from the boundary conditions. This means we have to use the PDE to derive $(n-1)\cdot m$ additional equations for the unknown values. The procedure for deriving those equations is the same for all finite difference methods. Namely we write down the equation (1.7) at $(n-1)\cdot m$ points and then approximate the derivatives at the chosen points by finite difference formulas using the values of the unknown function u only at the points that correspond to our table. We get different methods by choosing different points for writing out equations and by using different finite difference approximations for the derivatives.

In order to get an explicit method for backward parabolic equation we start by writing out the equation (1.7) at the points (x_i, t_k), $i = 1, \ldots, (n-1)$, $k = 1, \ldots, m$:

$$\frac{\partial u}{\partial t}(x_i, t_k) + \alpha(x_i, t_k)\frac{\partial^2 u}{\partial x^2}(x_i, t_k) + \beta(x_i, t_k)\frac{\partial u}{\partial x}(x_i, t_k) - r\, u(x_i, t_k) = 0.$$

To approximate the partial derivatives of u in the previous equation we use its values at the following grid points (surrounded by a green curve, the red circle denotes the point where we wrote down the equation).



Using the approximations (1.12), (1.13) and (1.14) for the time derivative, for the first derivative with respect to x and for the second derivative with respect to x, respectively, we get

$$\frac{\partial u}{\partial t}(x_i,t_k) = \frac{u(x_i,t_k)-u(x_i,t_{k-1})}{\Delta t} + O(\Delta t) = \frac{u_{ik}-u_{i,k-1}}{\Delta t} + O(\Delta t),$$
$$\frac{\partial u}{\partial x}(x_i,t_k) = \frac{u_{i+1,k}-u_{i-1,k}}{2\Delta x} + O(\Delta x^2),$$
$$\frac{\partial^2 u}{\partial x^2}(x_i,t_k) = \frac{u_{i-1,k}-2u_{ik}+u_{i+1,k}}{\Delta x^2} + O(\Delta x^2).$$

Thus the values u_{ki} of the exact solution u satisfy the relations

$$\frac{u_{i,k}-u_{i,k-1}}{\Delta t} + \alpha_{ik}\frac{u_{i-1,k}-2u_{ik}+u_{i+1,k}}{\Delta x^2} + \beta_{ik}\frac{u_{i+1,k}-u_{i-1,k}}{2\Delta x} - ru_{ik} + O(\Delta t + \Delta x^2) = 0, \tag{1.17}$$

where

$$\alpha_{ik} = \alpha(x_i,t_k),\ \ \beta_{ik} = \beta(x_i,t_k).$$

The idea of the finite difference methods is that throwing away the small error term $O(\Delta t + \Delta x^2)$ in (1.17) should cause only small errors in the results. Therefore we find the approximate values U_{ik} from the equations

$$\frac{U_{i,k}-U_{i,k-1}}{\Delta t} + \alpha_{ik}\frac{U_{i-1,k}-2U_{ik}+U_{i+1,k}}{\Delta x^2} + \beta_{ik}\frac{U_{i+1,k}-U_{i-1,k}}{2\Delta x} - r\,U_{ik} = 0. \tag{1.18}$$

The algorithm of the explicit finite difference method.

Solving the equations (1.18) for $U_{i,k-1}$ we get

$$U_{i,k-1} = a_{ik}U_{i-1,k} + b_{ik}U_{ik} + c_{ik}U_{i+1,k},\ \ i=1,2,\ldots,n-1,\ \ k=1,\ldots,m, \tag{1.19}$$

where

$$a_{ik} = \frac{\Delta t}{\Delta x^2}\left(\alpha_{ik} - \frac{\beta_{ik}}{2}\Delta x\right),$$
$$b_{ik} = 1 - 2\frac{\Delta t}{\Delta x^2}\alpha_{ik} - r\Delta t,$$
$$c_{ik} = \frac{\Delta t}{\Delta x^2}\left(\alpha_{ik} + \frac{\beta_{ik}}{2}\Delta x\right).$$

The equations (1.19) are in a very convenient form: if we know the values corresponding to k-th column of the matrix U, then using those equations we can simply compute the values $U_{i,k-1},\ i=1,\ldots,n-1$. In order to be able to compute all values of the table we should additionally specify how the values of the zeroth and n-th rows should be computed. One way to do this is to specify some functions $\phi_1(t,x_{min})$ and $\phi_2(t,x_{max})$ and define $U_{0k} = \phi_1(t_k,x_{min}),\ U_{nk} = \phi_2(t_k,x_{max}),\ k=0,1,\ldots,m-1$.

Unfortunately we do not know what are the right functions ϕ_1 and ϕ_2 (ideally, they should be the values of the unknowns solution u at those boundaries) but we should try to specify some functions that are not too far from the right values. If the choice of the functions is not very good, then our approximate solution may have relatively large errors close to $x = x_{min}$ and $x = x_{max}$. The simplest reasonable choice is to require that the values of the approximate solution U remain constant at the boundaries, ie

$$U_{0k} = p(e^{x_{min}}),\ U_{nk} = p(e^{x_{max}}),\ k = 0, 1, \ldots, m-1. \tag{1.20}$$

Of cause this choice introduces some errors close to the boundary, therefore we should choose x_{min} and x_{max} so that the region of values of x we are interested in is quite far from both of those values (but between the values). We'll come back to the question of choosing good boundary conditions later.

To summarize, we have derived the following *explicit finite difference method* for solving the equation (1.7):

1. Fill according to (1.10) the m-th column of the table.
2. For each time step $k = m, m-1, \ldots, 1$
   (a) Fill according to (1.19) the $(k-1)$-th column, except the 0-th and the n-th value.
   (b) Compute according to (1.20) the 0-th and the n-th value of the column.

Stability of the explicit method

It turns out that the error of the approximate solution obtained by the explicit finite difference method does not always go to zero when m and n tend to infinity. Namely, if a certain relation between m and n values does not hold the difference between the exact values and the approximate solutions may grow by a factor that is bigger than one at each timestep, resulting in huge errors at the final time. If this happens, the method is called *instable*.

In order to understand better the phenomenon of instability, let us consider a situation where we have two different sets of the values $U_{ik},\ i = 0, \ldots, n$ and $\tilde{U}_{ik},\ i = 0, \ldots, n$ of the approximate solution at the k-th column. Then the values of the $(k-1)$-th column, computed according to the explicit finite difference method, satisfy the equations

$$U_{i,k-1} - \tilde{U}_{i,k-1} = a_{ik}(U_{i-1,k} - \tilde{U}_{i-1,k}) + b_{ik}(U_{ik} - \tilde{U}_{ik}) + c_{ik}(U_{i+1,k} - \tilde{U}_{i+1,k}),\ i = 1, 2, \ldots, n-1.$$

Let ε be the maximal difference of the values of U and $\tilde{U}$ at the k-th column, then

$$|U_{i,k-1} - \tilde{U}_{i,k-1}| \le (|a_{ik}| + |b_{ik}| + |c_{ik}|)\varepsilon,\ i = 1, 2, \ldots, n-1.$$

If all coefficients a_{ik}, b_{ik} and c_{ik} are non-negative then, taking into account the equality

$$a_{ik} + b_{ik} + c_{ik} = 1 - r\,\Delta t,$$

we get that the maximal error in the $(k+1)$-th row is bounded by $(1 - r\,\Delta t)\varepsilon$. That means that in this case the errors are not increasing and the method is *stable*. But if any of the coefficients is negative, then the sum of the absolute values of the coefficients may be larger than 1 and the errors in one time step may be multiplied by a factor that is larger than one in each of the subsequent time steps, resulting in huge errors at the final time. There are always some errors in numerical computations (the real numbers are not computed exactly, the approximate solution is not exactly equal to the theoretical one). Therefore, when implementing the method, it is important to choose m and n so that the coefficients are all positive since otherwise the answers may be totally inaccurate.

1.7 Basic implicit finite difference method. Crank-Nicolson method

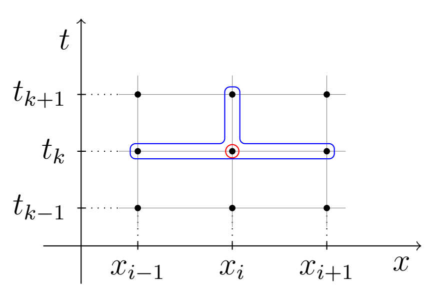
Explicit finite difference method is very convenient for implementation but it turned out to have a uncomfortable feature of being unstable if one does not choose the values of the discretization parameters m and n carefully. Next we consider some methods that are always stable but require the solution of a system of equations at each timestep.

1.7.1 Derivation of the basic implicit method

When deriving the basic implicit finite difference method we use the equation (1.7) at the points (x_i, t_k), $k = 0, 1, \ldots, m-1$, $i = 1, \ldots, n-1$ (in comparison with the explicit method, we use the points with $t = 0$ instad of $t = T$) and the forward difference approximation for the time derivative:

$$\frac{\partial u}{\partial t}(x_i, t_k) = \frac{u(x_i, t_{k+1}) - u(x_i, t_k)}{\Delta t} + O(\Delta t).$$

The derivatives with respect to x are approximated as before. This means that we use the following points (surrounded by the blue curve) for approsimating the partial derivatives of the equation at the point (x_i, t_k).



After substituting in the the approximations for the derivatives and throwing away the error terms we get

$$\frac{U_{i,k+1}-U_{i,k}}{\Delta t}+\alpha_{ik}\frac{U_{i-1,k}-2U_{ik}+U_{i+1,k}}{\Delta x^2}+\beta_{ik}\frac{U_{i+1,k}-U_{i-1,k}}{2\Delta x}-r\,U_{ik}=0,\ k=1,\ldots,m,\ i=1,\ldots,n-1. \tag{1.21}$$

After simplifications we get the following system of equations for finding the values U_{ik}:

$$a_{ik}U_{i-1,k}+b_{ik}U_{ik}+c_{ik}U_{i+1,k}=U_{i,k+1}, k=0,1,\ldots,m-1,\ i=1,\ldots,n-1, \tag{1.22}$$

where

$$\begin{aligned} a_{ik} &= -\frac{\Delta t}{\Delta x^2}\left(\alpha_{ik}-\frac{\beta_{ik}}{2}\Delta x\right), \\ b_{ik} &= 1+2\frac{\Delta t}{\Delta x^2}\alpha_{ik}+r\Delta t, \\ c_{ik} &= -\frac{\Delta t}{\Delta x^2}\left(\alpha_{ik}+\frac{\beta_{ik}}{2}\Delta x\right). \end{aligned}$$

In order to find the values U_{ik}, we have to fix suitable boundary conditions at $x=x_{min}$, $x=x_{max}$ and solve step-by-step the systems of equations for $U_{i,m-1},\ i=0\ldots,n$, $U_{i,m-2},\ i=0\ldots,n,\ \ldots,\ U_{i0},\ i=0\ldots,n$. The errors of the finite difference approximation is $O(\Delta t+\Delta x^2)$ (there is an additional error coming from specifying the boundary conditions).

1.7.2 The stability of the basic implicit method

We show that the basic implicit method is stable under quite general assumptions about the coefficients.

Lemma 14 *If* $b_{ik}\geq 0,\ a_{ik}\leq 0$ *and* $c_{ik}\leq 0,\ k=0,\ldots,m-1,\ i=1,\ldots,(n-1)$, *then the basic implicit method is stable.*

Proof. Suppose U_{ik} and $\tilde{U}_{ik}$ both satisfy the same boundary conditions and the equation (1.22)for some $k\in\{0,1,\ldots,m-1\}$ and that $|U_{i,k+1}-\tilde{U}_{i,k+1}|\leq\varepsilon\ \forall i$. Denote

$$E_i=U_{ik}-\tilde{U}_{ik},\ i=0,\ldots,n.$$

Denote by M the maximal value of $|E_i|,\ i=0,\ldots,n$. We want to show that $M\leq\varepsilon$; this shows the stability of the system. Since both U_{ik} and $\tilde{U}_{ik}$ satisfy (1.22), their difference also satisfies the system. We write the equation for the difference in the form

$$b_{ik}E_i=U_{i,k+1}-\tilde{U}_{i,k+1}-a_{ik}E_{i-1}-c_{ik}E_{i+1}.$$

By taking absolute values of both sides and using properties of the absolute value, we get

$$b_{ik}|E_i|\leq\varepsilon-a_{ik}|E_{i-1}|-c_{ik}|E_{i+1}|.$$

Here we used all of the assumptions of the lemma. We can make the right hand side larger, by replacing the absolute values of E_{i-1} and E_{i+1} with the maximal value M:

$$b_{ik}|E_i| \leq \varepsilon - a_{ik}M - c_{ik}M.$$

The last inequality holds for all $i = 1, \dots, n-1$. Choose the value of $i \in \{1, \dots, n-1\}$ such that $|E_i| = M$. In the case of that i we have

$$b_{ik}M \leq \varepsilon - a_{ik}M - c_{ik}M,$$

hence

$$(a_{ik} + b_{ik} + c_{ik})M \leq \varepsilon.$$

But $a_{ik} + b_{ik} + c_{ik} = 1 + r\Delta t$, hence we have shown that

$$M \leq \frac{\varepsilon}{1 + r\Delta t} < \varepsilon.$$

This proves the lemma. □

From the formulas of the coefficients it is easy to see, that the validity of the stability conditions does not depend on m and that under quite general assumptions about α and β the conditions hold for sufficiently large values of n.

1.7.3 Derivation of the Crank-Nicolson method

One problem with the numerical methods considered so far is that their accuracy with respect to time (first order accuracy, $O(\Delta t)$) is lower than with respect to the x variable (second order accuracy, $O(\Delta x^2)$). This means that if we want to reduce the error four times, we have to increase the value of m four times and the value of n two times, resulting in 8 times longer computation time. It would be much nice to have second order accuracy with respect to the t variable, too.

The low accuracy of the explicit and basic implicit methods with respect to time comes from the fact that both forward and backward difference (used for approximating the derivative with respect to t) have the first order accuracy at the points (x_i, t_k) where we wrote down our partial differential equation. But, taking into account that the central difference approximates a derivative with the second order accuracy, the finite difference approximation $\frac{\partial u}{\partial t} \approx \frac{u_{i,k+1} - u_{ik}}{\Delta t}$ is of the order $O(\Delta t^2)$ at the point $(x_i, t_k + \frac{1}{2}\Delta t)$. This gives the idea to try to get a better approximation of the partial differential equation by writing the equation out at those points before approximating the derivatives. But before going to the derivation of the method, we need one additional approximation formula.

Lemma 15 *Let f be a twice differentiable function. Then*

$$f(z) = \frac{f(z+h) + f(z-h)}{2} + O(h^2).$$

Proof. This formula like finite difference formulas can be proved by using Taylor's formula. Namely we have

$$f(z+h) = f(z) + f'(z)h + f''(\xi_1)\frac{h^2}{2},$$

where $\xi_1 \in [z, z+h]$ and

$$f(z-h) = f(z) - f'(z)h + f''(\xi_2)\frac{h^2}{2},$$

where $\xi_2 \in [z-h, z]$. By adding the formulas we get

$$f(z+h) + f(z-h) = 2f(z) + \left(f''(\xi_1) + f''(\xi_2)\right) h^2.$$

The last term is bounded by $const.\, h^2$ and thus, after solving for $f(z)$ we get the formula given in the lemma.□

Now we are ready to derive Crank-Nicolson method. Denote $t_{k+\frac{1}{2}} = t_k + \frac{\Delta t}{2}$. Let us use the following steps for deriving a finite difference method for our equation:

1. Write the equation (1.7) out at the points

   $$(x_i, t_{k+\frac{1}{2}}),\ k = 0, 1, \ldots, m-1,\ i = 1, \ldots, n-1)$$

   and use the approximation

   $$\frac{\partial u}{\partial t}(x_i, t_{k+\frac{1}{2}}) = \frac{u_{k+1,i} - u_{k,i}}{\Delta t} + O(\Delta t^2)$$

   for the time derivative.

2. Approximate terms containing u and it's partial derivatives with respect to x at $(x_i, t_{k+\frac{1}{2}})$ with the average values of those quantities at the points (x_i, t_{k+1}) and (x_i, t_k), ie use the approximations

   $$\begin{aligned}
   \gamma\, u(x_i, t_{k+\frac{1}{2}}) &= \frac{1}{2}(\gamma\, u(x_i, t_{k+1}) + \gamma\, u(x_i, t_k)) + O(\Delta t^2),\\
   \beta(x_i, t_{k+\frac{1}{2}})\frac{\partial u}{\partial x}(x_i, t_{k+\frac{1}{2}}) &= \frac{1}{2}(\beta(x_i, t_{k+1})\frac{\partial u}{\partial x}(x_i, t_{k+1}) + \beta(x_i, t_k)\frac{\partial u}{\partial x}(x_i, t_k)) + O(\Delta t^2),\\
   \alpha(x_i, t_{k+\frac{1}{2}})\frac{\partial^2 u}{\partial x^2}(x_i, t_{k+\frac{1}{2}}) &= \frac{1}{2}(\alpha(x_i, t_{k+1})\frac{\partial^2 u}{\partial x^2}(x_i, t_{k+1}) + \alpha(x_i, t_k)\frac{\partial^2 u}{\partial x^2}(x_i, t_k)) + O(\Delta t^2)
   \end{aligned}$$

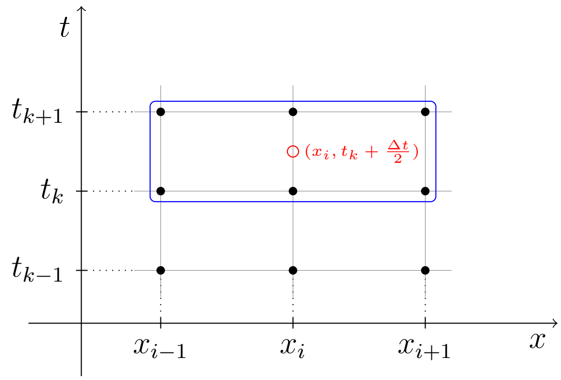
   and after that, replace the derivatives with the usual finite difference approximations.

3. Throw away the error terms and reorganize the equations so that the terms corresponding to $t = t_k$ are to the left of the equality and the terms corresponding to $t = t_{k+1}$ are on the right-hand-side of the equation.

After carrying through those steps we get a finite difference method of the form.

$$a_{ik}U_{i-1,k} + b_{ik}U_{ik} + c_{ik}U_{i+1,k} = d_{ik}U_{i-1,k+1} + e_{ik}U_{i,k+1} + f_{ik}U_{i+1,k+1}.$$

Schematically each equation involves the following grid points (red circle denotes the point where equation is written down, blue curve is around the grid points used in the corresponding discretized equation).



The error of the method (called Crank-Nicolson method) is of the order $O(\Delta t^2 + \Delta x^2)$

Exercise 8 *Verify, that the coefficients $a_{ik}, b_{ik}, \ldots, f_{ik}$ of the Crank-Nicolson method are given by*

$$\begin{aligned}
a_{ik} &= -\frac{\Delta t}{2\Delta x^2}\left(\alpha(x_i,t_k) - \frac{\beta(x_i,t_k)}{2}\Delta x\right),\\
b_{ik} &= 1 + \frac{\Delta t}{\Delta x^2}\alpha(x_i,t_k) + \frac{r\Delta t}{2},\\
c_{ik} &= -\frac{\Delta t}{2\Delta x^2}\left(\alpha(x_i,t_k) + \frac{\beta(x_i,t_k)}{2}\Delta x\right),\\
d_{ik} &= \frac{\Delta t}{2\Delta x^2}\left(\alpha(x_i,t_{k+1}) - \frac{\beta(x_i,t_{k+1})}{2}\Delta x\right),\\
e_{ik} &= 1 - \frac{\Delta t}{\Delta x^2}\alpha(x_i,t_{k+1}) - \frac{r\Delta t}{2},\\
f_{ik} &= \frac{\Delta t}{2\Delta x^2}\left(\alpha(x_i,t_{k+1}) + \frac{\beta(x_i,t_{k+1})}{2}\Delta x\right).
\end{aligned}$$

This method is used similarly to the basic implicit method: the values of U_{ik} are found step-by-step starting from $k = m - 1$, solving for each k a three-diagonal system of equations. It can be shown that the method is also unconditionally stable.

1.7.4 Using special solutions for boundary conditions

When constructing boundary conditions ϕ_1 and ϕ_2 in the case of option pricing it is often a good idea to use the fact, that the solution of the Black-Scholes equation satisfying the final condition $v(s,T) = c_1 s + c_2$ is

$$v(s,t) = c_1 e^{-r(T-t)} + c_2 e^{-D(T-t)} s.$$

Hence, if the payoff function is a linear function starting from some value $s = s_1$, then for large values of s the solution is practically equal to the special solution corresponding to this linear function. And if the payoff function is linear for $s < s_2$, than for small values of s the solution is practically equal to the special solution corresponding to that linear function. This gives us a possibility to define boundary values so that they are not very different from the actual option prices at those boundaries and hence to reduce significantly the error caused by introducing x_{min} and x_{max} in option pricing equations. For example, when finding the value of a call option price by solving the untransformed equation, we have that for large values of s the payoff is $p(s) = s - E$. Since for transformed equation $u(x,t) = v(e^x, t)$, a suitable boundary condition for $x = x_{max}$ is

$$\phi_2(t, x_{max}) = e^{-D(T-t)} e^{x_{max}} - E e^{-r(T-t)}.$$

1.7.5 Solving untransformed Black-Scholes equation

There are several problems with using the logarithmic transformation $x = \ln s$ before solving the Black-Scholes equation numerically. First, we get the values of the solution for stock prices that are unevenly spaced (at places $S_i = e^{x_i}$) and this is not very good if we want to form a table of option prices corresponding to evenly spaced intervals in the stock price; computing approximations for the derivatives is also more difficult. Second, this transformation makes our solution region doubly infinite while before the transformation we had a boundary at $S = 0$. This means that we have to introduce two artificial boundaries while for the untransformed equation it would be enough to specify only one artificial boundary $s = S_{max}$. Therefore it makes sense to try to solve the Black-Scholes equation without the logarithmic change of variables.

Recall that the option price satisfies the Black-Scholes partial differential equation

$$\frac{\partial v}{\partial t} + \alpha(s,t)\frac{\partial^2 v}{\partial s^2} + \beta(s)\frac{\partial v}{\partial s} - r\,u = 0,\ \ 0 < t \le T,\ \ s > 0 \tag{1.23}$$

together with the final condition

$$v(s,T) = p(s),\ \ s > 0.$$

Here

$$\alpha(s,t) = \frac{s^2 \sigma^2(s,t)}{2},$$
$$\beta(s) = (r - D)s.$$

Notice that the equation (1.23) is of the same form as (1.7), only instead of x we have s and the final condition and the values of the coefficients are computed differently.

This means that if we consider finite difference methods for finding the values of the function v at the points (s_i, t_k), where $s_i = i \cdot \frac{S_{max}}{n}), \; t_k = k \cdot \frac{T}{m},$, we can use the formulas for the coefficients we derived for (1.7) by changing x_i with s_i and Δx with $\Delta s = \frac{S_{max}}{n}$. Thus, we can use any of the methods derived so far without any additional effort for finding formulas for the coefficients.

Notice that if we take $s = 0$ in the equation (1.23) then we get an ordinary differential equation $\frac{\partial v}{\partial t}(0,t) = r\,v(0,t)$ which has the solution

$$v(0,t) = p(0)e^{-r(T-t)}.$$

Therefore we have an exact boundary condition at $s = 0$ so we have to specify only a condition for the artificial boundary $s = S_{max}$. The usual boundary conditions are:

- $v(S_{max}, t) = p(S_{max}), \; 0 \le t < T$. This condition can always be used but is relatively crude (meaning it introduces relatively large errors close to the boundary).
- If the payoff function p is linear from some point to infinity, $p(s) = k_1 + k_2 s, \; s \ge S_{max}$, then the boundary condition

  $$v(S_{max}, t) = k_1 e^{-r(T-t)} + k_2 e^{-D(T-t)} S_{max}$$

  corresponding to the special solution with the same final values gives usually much better results.

1.7.6 Computing the option prices with a given accuracy

In practical situations when one wants to compute option prices numerically, it is not enough just to get a value of the approximate solution. It is very important to know how to compute the value with a given accuracy. There are three sources of errors in using finite difference methods for computing option prices:

- the placement of artificial boundaries;
- the form of artificial boundary conditions used;
- the discretization error controlled by the parameters m and n.

Let us first consider estimating the discretization error.

Runge's error estimate

Usually a numerical procedure gives us the result with some error that we do not know:

$$result_1 = exact + error_1.$$

Very often we can rerun the numerical procedure with some other input parameters so that the error is (approximately) reduced by a certain factor $q > 1$: we get

$$result_2 = exact + error_2 \approx exact + \frac{error_1}{q}.$$

Then, by subtracting the second equation from the first one and reorganizing the terms, we get an estimate for the error of the second computation:

$$error_2 \approx \frac{result_1 - result_2}{q - 1}.$$

This is called Runge's error estimate.

This estimate can be used both in the case of solving the transformed equation or the untransformed equation. Let us discuss the usage in the case of solving the untransformed equation. In the case of finite difference methods we have considered so far, the (formal) error estimate is $O(\Delta t + \Delta s^2)$ (for the basic implicit method and for the explicit method) or $O(\Delta t^2 + \Delta s^2)$ in the case of Crank-Nicolson method. Assuming that the actual error behaves according to the estimate (the leading, meaning the most slowly decreasing term in the error expansion is shown in the estimate), the error is reduced four times, if we reduce Δs two times and Δt either four times (in the case of the basic implicit method or the explicit method) or two times (for Crank-Nicolson method). So, computing the numerical results first with $n = n_0,\ m = m_0$ (giving us $result_1$) and then with $n = 2n_0,\ m = 4m_0$ in the case of the basic implicit method or $m = 2m_0$ for Crank-Nicolson (giving us $result_2$), we can use the Runge's estimate with $q = 4$ to estimate the error at $S = S_0$. This means that if the computed option prices corresponding to $result_1$ are $U_{ik},\ k = 0, \ldots, m_0,\ i = 0, \ldots, n_0$ and the option prices corresponding to $result_2$ are $W_{ik},\ k = 0, \ldots, m_1,\ i = 0, \ldots, 2n_0$ (where $m_1 = 4m_0$ for basic implicit and $m_1 = 2m_0$ for Crank-Nicolson method), we estimate the discretization error as quantities

$$\frac{|U_{i_0,0} - W_{2i_0,0}|}{3},$$

where $i_0 = \frac{n_0 S_0}{s_{max}}$ for the first computation (so $s_{i_0} = S_0$).

Remark. Actually, the formal error estimates we have been using are correct only if the payoff function is sufficiently many times (at least two times) continuously differentiable. In the case of usual financial payoff functions (which have discontinuous derivatives) the error is not reduced by four but by a number between 2 and 4. Therefore, if we want to be more confident that the actual error is smaller than the estimated error, we should divide the difference of U and W by 2 or by 1 instead of three.

1.7.7 Computing with a given accuracy

There are known some theoretical estimates for the error caused by the artificial boundary conditions that allow one to choose the placement of the artificial boundary (for a given boundary condition) so that this component of error is less that a given number before starting numerical computations. Then one has to estimate only the discretization error when computing the option prices. Unfortunately those estimates are quite complicated, therefore we adopt a simple (although more time-consuming) approach in this course.

Suppose we want to find the current prices of an European option so that the error at the current stock price $S(0) = S_0$ was less that ε. When using a finite difference method for the untransformed equation, we have to fix only one artificial boundary; a good starting point is to take $s_{max} = 2 \cdot S_0$ (if σ is large or the time period is long, it may make sense to take larger value for s_{max}). In the case of solving the transformed equaiton, we also have to introduce another boundary for which we may take $s_{min} = \frac{S_0}{2}$ and define $x_{min} = \ln s_{min},\ x_{max} = \ln s_{max}$. Our procedure is as follows:

1. Fix a starting value n_0 and choose m_0 (in the case of the explicit method choose it from the stability constraint) and define $z = 1$, $\rho = 2$.

2. Solve the problem with a finite difference method (starting with $n = z \cdot n_0$ and $m = m_0$) in the domain $x_{min} = \ln S_0 \rho$, $x_{max} = \ln(\rho\, S_0)$ in the case of solving the transformed problem or in the domain $x_{min} = 0$, $x_{max} = \rho\, S_0$ in the case of solving the untransformed problem and estimate the error by Runge's method, until the (estimated) finite difference discretization error is less than $\frac{\varepsilon}{2}$.

3. multiply ρ by 2 and increase z by one in the case of the transformed equation or multiply it by two in the case of the untransformed equation and solve the problem with the same method (starting with $n = z \cdot n_0,\ m = m_0$ again until the (estimated) finite difference discretization error is less than $\frac{\varepsilon}{2}$. If the answer changes by more than $\frac{\varepsilon}{2}$ then repeat the step. Otherwise we have obtained the solution with the desired accuracy.

The reason for including the parameter z in the definition of the starting value of n is the desire to include the points s_i (or x_i in the case of the transformed equation) of the previous computation among the grid points after changing s_{min} and s_{max} values.

1.8 Pricing American options

American options, which give the holder the right to exercise the option at any time before the expiration time, are very popular and attractive for both buyers of the options and writers of the options. Buyers like the additional freedom compared to European options and the writers (sellers) like the possibility to earn extra money if the owner of the option does not choose the optimal time for exercising it.

1.8.1 An inequality for American options

It is clear that the price of an American option is never less than the price of the corresponding European option, so we have a lower bound on the option value. The following lemma allows us to obtain upper bounds.

Lemma 16 *If continuous and in the region $(s,t) \in (0,\infty) \times [0,T)$ two times continuously differentiable function $w(s,t)$ satisfies the inequalities*

$$\frac{\partial w}{\partial t} + \frac{s^2\sigma^2(s,t)}{2}\frac{\partial^2 w}{\partial s^2} + (r-D)s\frac{\partial w}{\partial s} - rw \leq 0,\ 0 \leq t < T,\ 0 < s < \infty$$

and

$$w(s,t) \geq p(s),$$

then the price $v(s,t)$ of the American option with the expiration date T and payoff function p satisfies the inequality $v(s,t) \leq w(s,t)$ $\forall (s,t) \in 0 \leq t < T,\ 0 < s < \infty$.

Proof. Fix $t_0 \in [0,T)$ and let $s_0 = S(t_0)$. Using the investment strategy $\eta(t) = \frac{\partial w}{\partial s}(S(t),t)$ with the initial wealth $X(t_0) = w(s_0,t_0)$ we get a portfolio which value $X(t)$ satisfies the inequality

$$\begin{aligned} d(X(t) - w(S(t),t)) = \Bigg(& r\,X(t) - r\,S(t)\frac{\partial w}{\partial s}(S(t),t) + D\,S(t)\frac{\partial w}{\partial s}(S(t),t) \\ & - \frac{\partial w}{\partial t}(S(t),t) - \frac{S(t)^2\sigma^2(S(t),t)}{2}\frac{\partial^2 w}{\partial s^2}(S(t),t)\Bigg)\,dt \\ \geq{} & r(X(t) - w(S(t),t))\,dt. \end{aligned}$$

Hence

$$d[e^{-rt}(X(t) - w(S(t),t))] \geq 0,$$

therefore also

$$\int_{t_0}^{t} d[e^{-r\tau}(X(\tau) - w(S(\tau),\tau))] = e^{-rt}(X(t) - w(S(t),t)) \geq 0\ \forall t \in [t_0,T].$$

This means that the value of the portfolio at any time t satisfies the inequality $X(t) \geq w(S(t),t) \geq p(S(t))$. Thus, at any time t_0, using the sum $w(S(t_0),t_0)$ we can form a self-financing portfolio which at any future time is at least as valuable as the option, therefore the option price can not be larger than $w(S(t_0),t_0)$.□

A corollary of the result is that the price of the European call option on a non-dividend paying stock is equal to the price of the American option with the same payoff function and the same exercise time.

It can be shown that the price of an American option at each point is either equal to the payoff function or satisfies the Black-Scholes differential equation.

Lemma 17 *The price of an American option with the payoff function p is a function of two variables t and s satisfying the following complementarity problem*

$$Lv(s,t) := \frac{\partial v}{\partial t} + \frac{s^2\sigma^2}{2}\frac{\partial^2 v}{\partial s^2} + (r-D)s\frac{\partial v}{\partial s} - rv \le 0 \quad 0 \le s < T,\ s \ge 0, \tag{1.24}$$

$$v(s,t) \ge p(s), \quad 0 \le s \le T,\ s \ge 0, \tag{1.25}$$

$$Lv(s,t)\,(v(s,t) - p(s)) = 0, \quad 0 \le s < T,\ s \ge 0. \tag{1.26}$$

1.8.2 Using finite difference methods for pricing American options

Let us consider American options with payoffs that depend only on the stock price at the time of exercising the option. Then the simplest possibility to compute the prices of American options is to modify a finite difference code of computing the values of European options so that at the end of each timestep we take the maximum of the computed values of U and the payoff function:

$$U_{ik} := \max(U_{ik}, p(s_i)),\ i = 1, \ldots, n.$$

This introduces additional error of the order $O(\Delta t)$, therefore modified Crank-Nicolson method does not converge faster than the basic implicit method.

1.9 Pricing Asian options

Asian options are options that depend on the the average stock price. There are two basic types of averages - arithmetic and geometric average, and the average can be compounded discretely (eg once a day) or continuously. We consider continuously compounded averages. Continuously compounded arithmetic average of the stock price over the period $[0, T]$ is

$$A = \frac{1}{T}\int_0^T S(t)\,dt,$$

the geometric average over the same period is

$$G = e^{\frac{1}{T}\int_0^T \ln(S(t))\,dt}.$$

Both of the averages give a value that is between the minimal and maximal stock prices over the period and it can be shown that the geometric average is never larger than the arithmetic average. There are four basic types of Asian options derived from the European put and call options. If the strike price E of European options is replaced in the payoff function by an average stock price, we get average strike put/call options. If the stock price itself is replaced by an average stock price, we get average price put/call options. Other types of payoff functions depending on the final stock price and the average stock price can be considered.

In order to derive partial differential equations for the prices of Asian options, we need a more general version of $It\bar{o}'s$ lemma.

Lemma 18 *Let $Y_1(t)$ and $Y_2(t)$ be two stochastic processes satisfying stochastic differential equations*

$$dY_i(t) = \alpha_i(Y_1(t), Y_2(t), t)\, dt + \beta_i(Y_1(t), Y_2(t), t)\, dB_i(t),\ \ i = 1, 2,$$

where B_1 and B_2 are brownian motions with correlation ρ (meaning that $(B_1(t_2) - B_1(t_1), B_2(t_2) - B_2(t_1)) \sim N(0, (t_2 - t_1)\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$ for all $t_1, t_2,\ t_1 < t_2$). Then

$$\begin{aligned} df(t, Y_1(t), Y_2(t)) = & \frac{\partial f}{\partial t}\, dt + \frac{\partial f}{\partial y_1}\, dY_1(t) + \frac{\partial f}{\partial y_2}\, dY_2(t) + \\ & \left(\frac{\beta_1^2}{2} \frac{\partial^2 f}{\partial {y_1}^2} + \rho \beta_1 \beta_2 \frac{\partial^2 f}{\partial y_1\, \partial y_2} + \frac{\beta_2^2}{2} \frac{\partial^2 f}{\partial {y_2}^2} \right) dt \end{aligned}$$

for all sufficiently smooth (having all needed partial derivatives) functions f.

Since arithmetic average can be expressed in terms of the integral of the stock price, let us introduce a new variable $I(t)$:

$$I(t) = \int_0^t S(\tau)\, d\tau.$$

Let us derive now the partial differential equation for the price of Asian options in the case of arithmetic average. The scheme is as before:

1. Make an assumption about what the option price depends on;

2. Assume that the option can be replicated by a self-financing portfolio;

3. use Ito's lemma for deriving the partial differential equation by requiring that the differential of the stock price and the differential of the value of the replicating portfolio to be equal.

It is clear that in the case of arithmetic average the option price depends on the current stock price, time and the integral of the stock price (since the average can be expressed in terms of the integral). Therefore assume that the Asian option price is a function of t, s and I, $v = v(t, s, I)$. In order to find differential of $v(t, S(t), I(t))$ we note that $I(t)$ satisfies

$$dI(t) = S(t)\, dt + 0\, dB_2(t),$$

so, according to Lemma 18, we have

$$\begin{aligned}
dv(S(t), I(t), t) &= \frac{\partial v}{\partial t}(S(t), I(t), t)\, dt + \frac{\partial v}{\partial s}(S(t), I(t), t)\, dS(t) + \frac{\partial v}{\partial I}(S(t), I(t), t)\, dI(t) \\
&\quad + \frac{S(t)^2\sigma^2}{2}\frac{\partial^2 v}{\partial s^2}(S(t), I(t), t)\, dt \\
&= \left(\frac{\partial v}{\partial t}(S(t), I(t), t) + S(t)\frac{\partial v}{\partial I}(S(t), I(t), t) + \frac{S(t)^2\sigma^2}{2}\frac{\partial^2 v}{\partial s^2}(S(t), I(t), t)\right) dt \\
&\quad + \frac{\partial v}{\partial s}(S(t), I(t), t)\, dS(t).
\end{aligned}$$

Recall, that the value of the self-financing portfolio corresponding to holding $\eta(t)$ stocks at any time t satisfies the equation

$$dX(t) = r(X(t) - \eta(t)S(t))\, dt + D\eta(t)S(t)\, dt + \eta(t)\, dS(t).$$

If the Asian option can be replicated, then the value of the corresponding self-financing portfolio is equal to the option price and the differentials of the option price and the value of the portfolio have to be the same. From the equality of $dv(t, S(t), I(t))$ and $dX(t)$ we get that $\eta(t) = \frac{\partial v}{\partial s}(t, S(t), I(t))$ and that the option price has to satisfy the differential equation

$$\frac{\partial v}{\partial t} + \frac{s^2\sigma^2}{2}\frac{\partial^2 v}{\partial s^2} + (r - D)s\frac{\partial v}{\partial s} + s\frac{\partial v}{\partial I} - rv = 0.$$

It is actually possible to reverse the derivation of the partial differential equation and to prove the following result.

Theorem 19 *Under the assumption of the validity of the Black-Scholes market model the price v of an Asian option depending on continuously compounded arithmetic average satisfies the equation*

$$\frac{\partial v}{\partial t} + \frac{s^2\sigma^2}{2}\frac{\partial^2 v}{\partial s^2} + (r - D)s\frac{\partial v}{\partial s} + s\frac{\partial v}{\partial I} - rv = 0,\ 0 \le t < T,\ s, I \ge 0 \qquad (1.27)$$

and the final condition

$$v(s, I, T) = p(s, \frac{I}{T}),$$

where p is the payoff function (the owner receives at the final time the payoff $p(S(T), A)$, where A is the arithmetical average of the stock price over the period [0,T]). The price of the option at any time $t \in [0, T]$ is given by $v(S(t), \int_0^t S(\tau)\, d\tau, t)$ and the option is replicated with the self-financing strategy holding $\eta(t) = \frac{\partial v}{\partial s}(S(t), \int_0^t S(\tau)\, d\tau, t)$ stocks at any time $t \in [0, T]$.

Proof. Exercise for the reader.□

The derivation of the equation for the Asian option depending on the geometric average is similar (with I denoting the integral of logarithm of the stock price); the resulting equation has, as the only difference, the coefficient $\ln s$ in front of the term with $\frac{\partial v}{\partial I}$. Since the options depending on the geometric average are less common, we do not consider them further in the course. We'll end our discussion of Asian options depending on the geometric average with the remark that in the case of constant volatility there exist exact formulas for several types of options which depend on geometric average.

Exercise 9 *Derive (with explanations) the partial differential equation and the final condition for pricing Asian options depending on geometric average.*

As in the case of European options, it is useful to know some special solutions of the equation 1.27. Let us look for special solutions that are linear with respect to the variables s and I. So we look at solutions of the form

$$v(s, I, t) = \phi_1(t) + \phi_2(t)\, s + \phi_3(t) I.$$

Substituting this guess into the equation (1.27) we get

$$\phi_1'(t) + \phi_2'(t)s + \phi_3'(t)I + (r - D)s\phi_2(t) + s\phi_3(t) - r(\phi_1(t) + \phi_2(t)\, s + \phi_3(t)I) = 0.$$

Since this equality has to hold for all values of s and I the constant term, the coefficient of s and the coefficient of I have to be equal to 0. Thus we get a system of ordinary differential equations:

$$\begin{cases} \phi_1'(t) & = r\,\phi_1(t), \\ \phi_2'(t) & = D\phi_2(t) - \phi_3(t), \\ \phi_3'(t) & = r\,\phi_3(t). \end{cases}$$

Solving this system of equations we get in the case $r \neq D$ a family of solutions of the form

$$v(s, I, t) = C_1 e^{-r(T-t)} + e^{-D(T-t)} \left(C_2 - C_3 \frac{e^{-(r-D)(T-t)}}{r - D} \right) s + C_3\, e^{-r(T-t)} I$$

and in the case $r = D$ a family of solutions of the form

$$v(s, I, t) = C_1 e^{-r(T-t)} + e^{-r(T-t)} \left(C_2 + C_3(T - t) \right) s + C_3\, e^{-r(T-t)} I.$$

This special solutions allow us to obtain put-call parities for Asian options. Recall that there are two kinds of put/call options of Asian type: average strike options and average price options. Let us derive a put/call relationship for average strike options.

Since the payoff of the average strike call is $\max(s - A, 0)$ and the payoff of the average strike put is $\max(A - s, 0)$, the portfolio corresponding to buying one average strike call and selling (writing) one average strike put gives the holder at the final time the income $S(T) - A$, where A is the continuously computed arithmetic average of the stock price. This means that if $C(s, I, t)$ is the function giving the value of the call option and $P(s, I, t)$ corresponds to the put option, then

$$C(s, I, T) - P(s, I, T) = s - \frac{I}{T}.$$

Since the equation (1.27) is a linear equation and both C and P are solutions of this equation, the difference is also a solution of the equation satisfying a linear final condition. Since linear final conditions correspond to special solutions, the difference of the call and put value is given by the formula of the special solution with the constants $C_1 = 0,\ C_3 = -\frac{1}{T},\ C_2 = 1 - \frac{1}{T(r-D)}$ (if $r \neq 0$) or $C2 = 1$, if $r = D$. Thus, in the case $r \neq D$ we have

$$C(s, I, t) = P(s, I, t) + e^{-D(T-t)} \left(1 + \frac{e^{-(r-D)(T-t)} - 1}{T(r-D)} \right) s - \frac{e^{-r(T-t)}}{T} I$$

and in the case $r = D$ we have

$$C(s, I, t) = P(s, I, t) + e^{-r(T-t)} \left(1 - \frac{(T-t)}{T} \right) s - \frac{e^{-r(T-t)}}{T} I.$$

1.9.1 A finite difference method for pricing Asian options depending on arithmetic average

In order to solve the equation (1.27) numerically, we introduce artificial boundaries $I = I_{max}$ and $s = S_{max}$ and derive equations for determining approximate values of the option price at the points (s_i, I_j, t_k), where

$$t_k = k \cdot \frac{T}{m},\ \ s_i = i \cdot \frac{S_{max}}{n_s} \text{ and } I_j = j \cdot \frac{I_{max}}{n_I}$$

for some natural numbers m, n_s, n_I. Denote by V_{ijk} the approximate values of the option price at the point (s_i, I_j, t_k). In order to derive equations for determining the approximate values, let us use the PDE at the points

$$(s_i, I_j, t_k),\ i = 1, \ldots, (n_s - 1),\ j = 0, \ldots, (n_I - 1),\ k = 0, \ldots, (m-1).$$

Denote $\Delta s = \frac{S_{max}}{m}$, $\Delta t = \frac{T}{m}$ and $\Delta I = \frac{I_{max}}{n_I}$. Instead of the partial derivatives, we use the following formulas:

$$\frac{\partial v}{\partial t}(s_i, I_j, t_k) = \frac{v(s_i, I_j, t_{k+1}) - v(s_i, I_j, t_k)}{\Delta t} + O(\Delta t),$$
$$\frac{\partial v}{\partial I}(s_i, I_j, t_k) = \frac{v(s_i, I_{j+1}, t_k) - v(s_i, I_j, t_k)}{\Delta I} + O(\Delta I),$$
$$\frac{\partial v}{\partial s}(s_i, I_j, t_k) = \frac{v(s_{i+1}, I_j, t_k) - v(s_{i-1}, I_j, t_k)}{2\Delta s} + O(\Delta s^2),$$
$$\frac{\partial^2 v}{\partial s^2}(s_i, I_j, t_k) = \frac{v(s_{i+1}, I_j, t_k) - 2v(s_i, I_j, t_k) + v(s_{i-1}, I_j, t_k)}{\Delta s^2} + O(\Delta s^2).$$

After substituting those approximations to the PDE and throwing away the error terms, we get the following equations for the approximate values of the option prices:

$$a_{ik}V_{i-1,jk} + b_{ik}V_{ijk} + c_{ik}V_{i+1,jk} = \frac{s_i \Delta t}{\Delta I}V_{i,j+1,k} + V_{ij,k+1},$$

where $i = 1, 2, \ldots, n_s - 1,\ j = 0, 1, \ldots, n_I - 1,\ k = 0, \ldots, m-1$ and

$$\begin{aligned} a_{ik} &= \frac{\rho}{2}(-s_i^2\sigma^2(s_i, t_k) + (r-D)s_i\Delta s), \\ b_{ik} &= \left(1 + \rho s_i^2\sigma^2(s_i, t_k) + \frac{s_i\Delta t}{\Delta I} + r\Delta t\right), \\ c_{ik} &= -\frac{\rho}{2}(s_i^2\sigma^2(s_i, t_k) + (r-D)s_i\Delta s). \end{aligned}$$

The equations for the approximate values can be viewed as a three-diagonal system for finding the values of V_{ijk}, if the values corresponding to the level $t = t_{k+1}$ and the values $V_{i,j+1}^k,\ i = 0, \ldots, n_s$ have been found earlier.

The values of V_{ijm} can be found from the final condition:

$$V_{ijm} = p(s_i, \frac{I_j}{T}),\ i = 0, \ldots, n_s,\ j = 0, \ldots, n_I.$$

At the boundary $S = 0$ we have the exact boundary condition $v(0, I, t) = p(0, \frac{I}{T})e^{-r(T-t)}$, hence

$$V_{0jk} = p(0, \frac{I_j}{T})e^{-r(T-t_k)},\ k = 0, \ldots, (m-1).$$

In order to solve the system of equation for V_{ijk} we have to specify boundary conditions at the boundaries $I = I_{max}$ and $S = S_{max}$. There are a few possible boundary conditions one can specify at those boundaries:

- If the payoff function is piecewise linear, then we may take maximum of the corresponding special solution and 0 at the boundaries.
- Use the boundary condition that corresponds to S remaining constant at the boundary; in that case the value at the boundary corresponds to the function $e^{-r(T-t)}p(s, \frac{I+(T-t)s}{T})$.

In the case of average price call and put options it can be shown that we can choose $I_{max} = E \cdot T$ and the exact boundary condition at $I = I_{max}$ is (assuming $r \neq D$) 0 in the case of average price put and is given by

$$v(t, s, I_{max}) = \frac{e^{-D(T-t)}}{(r-D)T}(1 - e^{-(r-D)(T-t)})\, s$$

for average price call option.

The procedure for solving the option pricing equation is as follows.

1. Fill in the values for $V_{ijm},\ i = 0, \dots, n_s,\ j = 0, \dots, n_I$, using the payoff function.
2. For each $k = m-1, m-2, \dots, 0$
   (a) Apply the boundary condition at $I = I_{max}$ to define $V_{i,n_I,k},\ i = 0, \dots, n_s$.
   (b) For each $j = n_I - 1, \dots, 0$ solve the three-diagonal system for values $V_{ijk},\ i = 0, \dots, n_s$, using the boundary conditions at $s = 0$ and $s = S_{max}$ to get two additional equations .

If we need only the option price at $t = 0$, then it is not necessary to store the full matrix V of approximate option prices; we need only two levels, $Vold$ corresponding to $t = t_{k+1}$ that is known and $Vnew$ corresponding to the current level $t = t_k$. At the beginning $Vold$ is computed using the final condition and at the end of each timestep the values of $Vnew$ are copied to $Vold$.

Chapter 2

Options depending on two underlying stocks

Sometimes it is of interest to consider options whose payoff functions depend on more than one stock prices. A few examples of popular so called rainbow options:

- Dual Put option, where the owner has the right to sell one of the two underlying stocks at the exercise time T: $p(s_1, s_2) = \max(E_1 - s_1, E_2 - s_2, 0)$;
- Dual Strike option, where the owner has the right to buy one of the two underlying stocks at the exercise time T: $p(s_1, s_2) = \max(s_1 - E_1, s_2 - E_2, 0)$.

Let us assume that the two stocks satisfy the system of stochastic differential equations

$$dS_1(t) = S_1(t)(\mu_1\, dt + \sigma_1\, dB_1(t)),$$
$$dS_2(t) = S_2(t)(\mu_2\, dt + \sigma_2\, dB_2(t)),$$

where $\mu_1, \mu_2, \sigma_1, \sigma_2$ may depend on time and on both stock prices and $(B_1(t), B_2(t))$ is a two-dimensional Brownian motion with correlation ρ (the increments of B_1 and B_2 over any time interval Δt are jointly normal with standard deviations $\sqrt{\Delta t}$ and correlation ρ). For simplicity of presentation of main ideas, we are going to consider only the case of constant volatilities.

Let us consider an European option with payoff function $p(s_1, s_2)$ such that the payoff at time T is $p(S_1(T), S_2(T))$. We can use our standard procedure for deriving PDE for the option pricing function.

- Step 1: assume that there is function $v(s_1, s_2, t)$ such that the option price at time t is $v(S_1(t), S_2(t), t)$.
- Step 2: Assume that the option can be replicated by a self-financing portfolio. Now the portfolio consists of holdings of two stocks and bank (or money market) account and it's change in time is described by

$$dX(t) = (X(t) - \eta_1(t)S_1(t) - \eta_2(t)S_2(t))r\, dt + \eta_1(t)S_1(t)D_1\, dt + \eta_2(t)S_2(t)D_2\, dt + \eta_1(t)\, dS_1(t) + \eta_2\, dS_2(t.)$$

By the no Arbitrage condition then we have that $X(t) = v(S_1(t), S_2(t), t)$ for all times $t \le T$.

- Step 3: By Itô's formula we have

$$d\,v(S_1(t), S_2(t), t) = \frac{\partial v}{\partial t}(\cdot)\,dt + \frac{\partial v}{\partial s_1}(\cdot)\,dS_1(t) + \frac{\partial v}{\partial s_2}(\cdot)\,dS_2(t) + \\ \frac{S_1(t)^2\sigma_1^2}{2}\frac{\partial^2 v}{\partial s_1^2}(\cdot)\,dt + \rho\,S_1(t)S_2(t)\sigma_1\sigma_2\frac{\partial^2 v}{\partial s_1 \partial s_2}(\cdot)\,dt + \frac{S_2(t)^2\sigma_2^2}{2}\frac{\partial^2 v}{\partial s_2^2}(\cdot)\,dt.$$

  Comparing this to the equation for $dX(t)$ we see, that in order to have equality we have to take $\eta_1(t) = \frac{\partial v}{\partial s_1}(S_1(t), S_2(t), t)$ and $\eta_2(t) = \frac{\partial v}{\partial s_2}(S_1(t), S_2(t), t)$. The equality of the remaining terms (when we take into account that $S_1(t)$ and $S_2(t)$ can take any positive values) leads to a partial differential equation for v.

Based on previous argument we have shown that the price v of European option with a payoff function $p(s_1, s_2)$ should satisfy a two-dimensional Black-Scholes Equation

$$\frac{\partial v}{\partial t} + \frac{\sigma_1^2 s_1^2}{2}\frac{\partial^2 v}{\partial s_1^2} + \rho\sigma_1\sigma_2 s_1 s_2 \frac{\partial^2 v}{\partial s_1 \partial s_2} + \frac{\sigma_2^2 s_2^2}{2}\frac{\partial^2 v}{\partial s_2^2} \\ + (r - D_1)s_1\frac{\partial v}{\partial s_1} + (r - D_2)s_2\frac{\partial v}{\partial s_2} - rv = 0,\ \ s_1, s_2 \ge 0$$

with the final condition

$$v(s_1, s_2, T) = p(s_1, s_2),\ \ s_1, s_2 \ge 0.$$

We are going to derive only the simplest (explicit) numerical method for pricing European options depending on two underlying stocks. But to have an efficient explicit method, it is a good idea to transforms the variables so that the equation simplifies and in the case of constant volatilities has constant coefficients.

Let us use the change of variables $v(s_1, s_2, t) = u(\frac{\ln(s_1)}{\sigma_1}, -\rho\frac{\ln(s_1)}{\sigma_1} + \frac{\ln(s_2)}{\sigma_2}, t)$, where u is a function of x_1, x_2 and t. In order to derive an equation for u, we take derivatives and substitute the results into PDE for v. In the following formulas we use notations

$x_1 = \frac{\ln(s_1)}{\sigma_1}$, $x_2 = -\rho\frac{\ln(s_1)}{\sigma_1} + \frac{\ln(s_2)}{\sigma_2}$ in the arguments of u:

$$
\begin{aligned}
\frac{\partial v}{\partial t}(s_1, s_2, t) &= \frac{\partial u}{\partial t}(x_1, x_2, t), \\
\frac{\partial v}{\partial s_1}(s_1, s_2, t) &= \frac{\partial u}{\partial x_1}(x_1, x_2, t)\frac{1}{\sigma_1\, s_1} + \frac{\partial u}{\partial x_2}(x_1, x_2, t)\frac{-\rho}{\sigma_1\, s_1}, \\
\frac{\partial v}{\partial s_2}(s_1, s_2, t) &= \frac{\partial u}{\partial x_2}(x_1, x_2, t) \cdot \frac{1}{\sigma_2\, s_2}, \\
\frac{\partial^2 v}{\partial s_1^2}(s_1, s_2, t) &= \frac{\partial^2 u}{\partial x_1^2}(x_1, x_2, t) \cdot \frac{1}{\sigma_1^2\, s_1^2} - \frac{\partial u}{\partial x_1}(x_1, x_2, t) \cdot \frac{1}{\sigma_1\, s_1^2} + \\
&\quad \frac{\partial^2 u}{\partial x_1 \partial x_2}(x_1, x_2, t) \cdot \frac{-2\rho}{\sigma_1^2\, s_1^2} + \\
&\quad \frac{\partial^2 u}{\partial x_2^2}(x_1, x_2, t) \cdot \frac{\rho^2}{\sigma_1^2\, s_1^2} + \frac{\partial u}{\partial x_2}(x_1, x_2, t) \cdot \frac{\rho}{\sigma_1\, s_1^2} \\
\frac{\partial^2 v}{\partial s_1 \partial s_2}(s_1, s_2, t) &= \frac{\partial^2 u}{\partial x_1 \partial x_2}(x_1, x_2, t)\frac{1}{\sigma_1\, s_1 \sigma_2 s_2} + \frac{\partial^2 u}{\partial x_2^2}(x_1, x_2, t) \cdot \frac{-\rho}{\sigma_1\, s_1 \sigma_2 s_2} \\
\frac{\partial^2 v}{\partial s_2^2}(s_1, s_2, t) &= \frac{\partial^2 u}{\partial x_2^2}(x_1, x_2, t) \cdot \frac{1}{\sigma_2^2\, s_2^2} + \frac{\partial u}{\partial x_2}(x_1, x_2, t) \cdot \frac{-1}{\sigma_2\, s_2^2}
\end{aligned}
$$

So the equation is transformed to the equation

$$
\begin{aligned}
\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial^2 u}{\partial x_1^2} + \frac{1-\rho^2}{2}\frac{\partial^2 u}{\partial x_2^2} + \frac{r - D_1 - \frac{1}{2}\sigma_1^2}{\sigma_1}\frac{\partial u}{\partial x_1} \\
+ \left(\frac{r - D_2 - \frac{1}{2}\sigma_2^2}{\sigma_2} - \rho\frac{r - D_1 - \frac{1}{2}\sigma_1^2}{\sigma_1}\right)\frac{\partial u}{\partial x_2} - ru = 0.
\end{aligned}
$$

From the final condition for v it follows that u satisfies the condition

$$u(x_1, x_2, T) = p(e^{\sigma_1 x_1}, e^{\sigma_2(x_2 + \rho x_1)}).$$

In order to solve the transformed equation numerically we introduce artificial boundaries $x_{1,min}, x_{1,max}, x_{2,min}, x_{2,max}$ and look for approximate solution at the points $(x_{1,i}, x_{2,j}, t_k)$, where

$$t_k = k\,\Delta t,\;\; x_{1,i} = x_{1,min} + i\,\delta x_1,\;\; x_{2,j} = x_{2,min} + j\,\delta x_2$$

and

$$\Delta t = \frac{T}{m},\;\; \delta x_1 = \frac{x_{1,max} - x_{1,min}}{n_1},\;\; \delta x_2 = \frac{x_{2,max} - x_{2,min}}{n_2}$$

for some natural numbers m, n_1, n_2.

We use the usual notation $U_{i,j,k} \approx u(x_{1,i}, x_{2,j}, t_k)$. Using the PDE at $(x_{1,i}, x_{2,j}, t_k)$ for $i = 1, \ldots, n_1 - 1,\; j = 1, \ldots, n_2 - 1$ and $k = 1, \ldots, m$ we have to use the backward difference formula for the time derivative (since we are using the equation at final time, where the final condition is given), derivatives with respect to x_1 and x_2 can be approximated by central difference formulas.

Replacing the derivatives with suitable finite difference approximations and solving for $U_{i,j,k-1}$ we get an explicit finite difference method

$$U_{i,j,k-1} = a_{-1,0}U_{i-1,j,k} + a_{0,0}U_{i,j,k} + a_{1,0}U_{i+1,j,k} + a_{0,1}U_{i,j+1,k} + a_{0,-1}U_{i,j-1,k}, \tag{2.1}$$

where

$$\begin{aligned}
a_{-1,0} &= \frac{\Delta t}{2}\left(\frac{1}{\Delta x_1^2} - \frac{r - D_1 - \frac{1}{2}\sigma_1^2}{\sigma_1 \Delta x_1}\right),\\
a_{1,0} &= \frac{\Delta t}{2}\left(\frac{1}{\Delta x_1^2} + \frac{r - D_1 - \frac{1}{2}\sigma_1^2}{\sigma_1 \Delta x_1}\right),\\
a_{0,-1} &= \frac{\Delta t}{2}\left(\frac{1-\rho^2}{\Delta x_2^2} - \frac{r - D_2 - \frac{1}{2}\sigma_2^2}{\sigma_2 \Delta x_2} + \rho\frac{r - D_1 - \frac{1}{2}\sigma_1^2}{\sigma_1 \Delta x_2}\right),\\
a_{0,1} &= \frac{\Delta t}{2}\left(\frac{1-\rho^2}{\Delta x_2^2} + \frac{r - D_2 - \frac{1}{2}\sigma_2^2}{\sigma_2 \Delta x_2} - \rho\frac{r - D_1 - \frac{1}{2}\sigma_1^2}{\sigma_1 \Delta x_2}\right),\\
a_{0,0} &= 1 - r\Delta t - \frac{\Delta t}{\Delta x_1^2} - \frac{\Delta t(1-\rho^2)}{\Delta x_2^2}.
\end{aligned}$$

This method is stable in the case of sufficiently large values of n_1 and n_2, if the coefficient Δt (or the number of timesteps m) is chosen so that the coefficient $a_{0,0}$ is nonnegative. In order to find the option prices with the explicit method one has to specify also boundary conditions at the artificial boundaries. The simplest choice is to fix the value at the boundaries to be equal to the value of the payoff function at the corresponding points.

Algorithm of using the method is as follows:

1. Fill in the values coming from the final condition: $U_{ijm} = p((e^{\sigma_1 x_{1,i}}, e^{\sigma_2(x_{2,j}+\rho x_{,i})}))$
2. For each $k = m, m-1, \ldots, 1$
   (a) Use boundary conditions to compute $U_{i,0,k-1},\ U_{i,n_2,k-1}$ for $i = 0, \ldots, n_1$ and $U_{0,j,k-1},\ U_{n_1,j,k-1}$ for $j = 0, \ldots, n_2 - 1$
   (b) Compute the values $U_{i,j,k-1}$ according to formulas (2.1) for $i = 1, \ldots, n_1 - 1$, $j = 1, \ldots, n_2 - 1$

The convergence rate of the method (for sufficiently nice final condition) is $O(\Delta t + \Delta x_1^2 + \Delta x_2^2)$ if the stability condition is satisfied.

Bibliography

[1] Black F, Scholes M, "The pricing of options and corporate liabilities," *Journal of Political Economy 81* (1973), 637-659.

[2] Glassermann, P. Monte Carlo Methods in Financial Engineering. Springer 2004.

[3] Strang G, Fix G. An Analysis of The Finite Element Method. Prentice Hall 1973.

[4] Wilmott P, Howison S, Dewynne. J, The Mathematics of Financial Derivatives. A Student Introduction. Cambridge, 1995.

[5] Wilmott P. Paul Wilmott on quantitative finance (Volumes 1,2). Jon Wiley and Sons, 2000.