FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

UNIVERZITY KOMENSKÉHO
V BRATISLAVE

D I P L O M O V Á P R Á C A

2004

SOŇA KILIANOVÁ

FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

UNIVERZITY KOMENSKÉHO
V BRATISLAVE

Ekonomická a finančná matematika

Analytical and numerical methods for stock index derivative pricing

Diplomová práca

Diplomant: Soňa Kilianová

Vedúci diplomovej práce: doc. RNDr. Daniel Ševčovič, CSc.

Bratislava 2004

Prehlasujem, že predkladanú diplomovú prácu

som vypracovala samostatne, len s pomocou konzultácií

a literatúry uvedenej v zozname.

Zároveň úprimne ďakujem svojmu diplomovému vedúcemu

doc. RNDr. Danielovi Ševčovičovi, CSc.

za všestrannú odbornú pomoc a ochotu prejavenú pri vedení práce.

V Bratislave 2. apríla 2004

Contents

Introduction **3**

1 Indices and their derivatives **4**
1.1 What indices are 4
1.2 Some of the more commonly-used indices 5
1.3 Current stock indices 8
1.4 The index derivatives 8
1.4.1 Index options 9

2 Stock index derivative pricing model **12**
2.1 The n-dimensional Black-Scholes equation 12
2.1.1 A stochastic equation for a derivative value 12
2.1.2 Construction of a riskless portfolio 14
2.1.3 The terminal conditions 16
2.2 Analytical solution to the Black-Scholes equation 16

3 Numerical methods **22**
3.1 Full space-time discretization 22
3.1.1 Discretization in time 22
3.1.2 Spatial discretization 23
3.1.3 The boundary conditions 24
3.2 The additive operator splitting technique 25
3.3 The order of the AOS approximation 27
3.3.1 The forming matrices 30
3.3.2 Error estimates 37
3.4 Computational parallelism 40

Conclusions **41**

References **42**

Introduction

In the last decades, we could observe rapidly expanding trading of financial derivative securities in financial markets. Mathematical modelling in finance as well as empirical analysis draw attention from researchers from a wide spectrum of disciplines, ranging from mathematics and statistics over operations research to economics.

In this paper, we focus on a special class of financial derivatives. We consider the pricing of European type of options whose values are determined by the prices of several underlying assets. Stock index options are a typical example of such a financial instrument. This paper deals with a theoretical framework which can be applied to index derivative pricing.

In the first chapter we explain the main characteristics of indices and introduce the mostly known examples such as the Dow Jones Industrial Index or the S&P 500 Index.

Next, we recall the generalized Black-Scholes equation derived in Kwok's book [6]. This equation governs the price of an option with multiple underlying assets where the asset prices all follow the lognormal distributions. We derive an analytical solution to the Black-Scholes equation consisting of solving an n-dimensional integral where n denotes the number of undelying assets.

Because of the practical uselessness of the analytical solution for indices with many underlyings (for example $n = 500$ when considering the S&P 500), several numerical attempts to solve this high-dimensional problem occurred worldwide. We can mention for example methods using Monte Carlo simulation ([10],[7]) or an algebraic approach ([8]).

In the third chapter, we introduce the additive operator splitting (AOS) method known mainly from image processing. Using this technique leads to the decomposition of the multi-dimensional problem into several one-dimensional problems, which can be solved very efficiently. We apply the AOS method to solve the Black-Scholes equation instead of solving a high-dimensional integral which is the analytical solution to this equation.

Finally, we pay attention to estimation of the error of the AOS scheme. We show that the use of this method generates (when choosing the appropriate spatial discretization step h and the time-step k) an error of the order $O(h^2)$ when comparing to a precise analytical solution to a parabolic differential equation.

1 Indices and their derivatives

In this chapter, we explain the basic terminology such as an index, a derivative and an index derivative. We introduce the main ways of index composing and deal with the mostly known indices in more detail. Next, we characterize the index options. The text is supplemented by tables and figures showing the real data on some indices and options on indices, in order to present briefly the current situation in the world market. Sources [9] and several internet sites were used[1].

1.1 What indices are

A portfolio encompassing all possible (or many) securities would be too broad to measure. Proxies such as stock indices have been developed to serve as indicators of the overall market's performance. In addition, specialized indices have been developed to measure the performance of more specific parts of the market, such as small companies.

It is important to realize that a stock price index by itself does not represent an average return to shareholders. By definition, a stock price index considers only the prices of the underlying stocks and not the dividends paid. Dividends can account for a large percentage of the total investment return.

One characteristic that varies among stock indices is how the stocks comprising the index are weighted in the average. Even if no explicit weighting is applied when calculating an average, there may be an implicit one. While a one dollar price change in one stock in a simple stock price index will have the same effect as a one dollar change in any other stock, a given percentage increase of a higher price stock influences the index more than a corresponding percentage increase of a lower price stock. For example, a 1% change in a $ 100 stock will change the index more than a 1% change in a $ 10 stock. For this reason, indices that are based on the simple summation of the stock prices are referred to as *price-weighted*.

As an example we can mention one of the mostly known indices at all, the American Dow Jones Industrial Average. The Japanese Nikkei 225 is constructed in a similar way.

In a price-weighted index, a change in the stock price of the largest company in the index would influence the average no more than an equal change in the stock

[1]
http://money.cnn.com/markets/world_markets.html
http://quote.cboe.com
http://www.neatideas.com/djia.htm
http://www.sec.gov/answers/indices.htm
http://www.theponytail.net/DOL/DOLnode8.htm
http://www.quickmba.com/finance/invest/indices.shtml

price of the smallest company in the index. However, the larger company's performance will have a greater impact on the economy. To consider the size of the company, a *market capitalization weighted* index (or *value-weighted* index) can be used, in which a company's impact on the index is proportional to the size of the company.

The majority of the main world indices is created in this way. We can name for example the Standard & Poor's 500 Index, NASDAQ Composite Index, DAX Index, Wilshire 5000 Index, London FTSE, MSCI indices, Czech PX 50 and Slovak SAX.

Some indices do not weight for market capitalization, but do adjust for price differences to remove the implicit price weighting. This unweighted method tracks the performance of an index in which equal dollar amounts are invested in the underlying stocks. Some consider an unweighted index to be a good indicator of a market's performance from the perspective of the investor who places an equal amount of money in each stock in his or her portfolio, regardless of its market capitalization. However, if every investor placed an equal amount of money in each investment, relatively few investors would own small-cap stocks, so an unweighted index would not reflect the portfolio performance of the average investor when all investors are considered.

1.2 Some of the more commonly-used indices

There are hundreds of indices that are designed to measure the broad market or a specific part of it.

The Dow Jones Industrial Average (DJIA) is a price-weighted index and is the most widely quoted stock index. It consists of 30 American ”blue chips” (stocks of great American companies) which are the best representatives of the American economy. The Index includes a wide range of companies from financial services companies over computer companies to retail companies, but does not include any transportation or utility companies, which are included in separate indices. The stocks included in the DJIA are not changed often.

The Index was created by Charles Dow in 1896 and is now the historically oldest stock index. The original index comprised only twelve stocks including General Electric, which as the only one can be still found in this index. In 1916, the DJIA was extended into twenty stocks, and this number increased to 30 in 1928.

The Dow Jones Industrial Average uses a divisor to adjust for events that result in no change in a company's value but that would otherwise influence the index. One such event is a stock split; another is the replacement of one company in the index by another. While this adjustment does not result in a change in the index value when a stock splits, because the index is price-weighted newly split stock will have a lower price and therefore less influence on the index.

Figure 1: The evolution of the DJIA Index from March 2000 to February 2001.
Source: BigCharts.com, February 16, 2001

The Dow Jones Industrial Average is calculated by the following formula:

$$DJIA = \frac{\sum_{i=1}^{30} S_i}{Divisor} \tag{1}$$

where S_i is the current market price of stock i.

The DJIA Index followed the evolution of the overall American stock market very precisely.

Standard & **Poor's 500 Index** is a capitalization-weighted index of 500 stocks intended to be a representative sample of leading companies in leading industries within the U.S. economy. Stocks in the Index are chosen for market size (large-capital), liquidity, and industry group representation. The S&P 500 Index includes industrial, utility, transportation, and financial stocks. It is widely used as a benchmark by institutional investors.

Except the American companies, the S&P 500 Index contains also some Canadian and only two European companies: Royal Dutch Petroleum and Unilever. More than 85 per cent of the them are traded in New York Stock Exchange (NYSE), the remainder mostly in NASDAQ and only a slight number in American Stock Exchange (AMEX).

Table 1: Currently mostly used stock indices, March 22, 2004 (Close of day)
Source: http://www.neatideas.com/djia.htm

Indicator	**Value**
DJIA	10064,75
S&P500	1095,40
NASDAQ Composite	1909,90
Russell 2000	568,99
Nikkei 225	11318,50
FTSE 100	4333,80
DAX	3729,23

The formula for computing the S&P 500 reads as:

$$SP = \frac{\sum_{i=1}^{500} S_i Q_i}{Divisor} \tag{2}$$

where S_i is the current market price of stock i, Q_i is the number of stocks emitted in the market.

Similarly as for DJIA, the divisor's function is to compensate the unfavorable influences of the corporate phenomena.

Deutscher Aktien Index (DAX) is composed of thirty German ”blue chips”. It belongs to the capital-weighted indices. It is calculated by the same rules as the S&P500.

A more detailed description of indices characteristics and their construction can be found in [9] or on internet web sites.

In summary, the index value is a weighted sum of current stock prices where weights are either equally distributed (DJIA) or depend on the number of emitted stocks (S&P 500, etc.). In general, the index value I can be expressed as

$$I = \sum_{i=1}^{n} w_i S_i$$

where w_i are weights corresponding to the index definition, S_i are the prices of n underlying stocks.

1.3 Current stock indices

At this place we present the values of some world stock indices. Figure 1 shows the time evolution of the daily values of the Dow Jones Industrial Average (DJIA) in the year 2000 and the beginning of 2001. We can observe its stochastic character when considering short time periods as well as its stability in 2000. The bottom part of the figure shows the volume traded. Table 1 shows the current values of the mostly known indices. Table 2 describes the situation on Asian exchanges. It shows the current values of the indices and the changes in index values to the given date. Table 3 deals with European indices.

Table 2: Current Asian stock indices.
Source: http://money.cnn.com/markets/world_markets.html

exchange	**index**	**level**	**change**	**%change**	**date**	**time (ET)**
Australia	All Ordinaries	3395.70	+10.50	0.31%	Mar 25	1:08:56
China	Shanghai Comp.	1820.94	-5.40	-0.30%	Mar 25	2:09:26
Hong Kong	Hang Seng	12520.21	-157.92	-1.25%	Mar 25	4:59:00
Indonesia	Composite	727.10	-11.69	-1.58%	Mar 25	4:03:45
Japan	Nikkei 225	11530.91	+165.92	1.46%	Mar 25	1:12:42
Japan	TOPIX	1160.00	+13.40	1.18%	Mar 25	4:05:33
Malaysia	Composite	890.14	-5.17	-0.58%	Mar 25	4:04:52
New Zealand	Top 40	2302.15	+11.97	0.52%	Mar 25	0:09:19
Philippines	PHS Composite	1393.12	+3.25	0.23%	Mar 24	23:38:20
Singapore	Straits Times Ind.	1830.40	-4.89	-0.27%	Mar 25	4:10:54
South Korea	KOSPI	853.38	-8.34	-0.97%	Mar 25	2:06:09
South Korea	KOSPI 200	112.16	-1.02	-0.90%	Mar 25	2:06:55
Taiwan	Weighted	6156.73	-56.83	-0.91%	Mar 25	1:09:49
Thailand	SET	664.66	-12.95	-1.91%	Mar 25	5:01:42

1.4 The index derivatives

A derivative is any financial instrument whose payoffs depend in a direct way on the value of an underlying variable at a time in the future. This underlying variable is also called the underlying asset, or just the underlying. Examples of underlying assets include financial asset, commodity, another derivative, index, interest rate, and many others.

Usually, derivatives are contracts to buy or sell the underlying asset at a future time, with the price, quantity and other specifications defined today. Contracts can

Table 3: Current European stock indices.
Source: http://money.cnn.com/markets/world_markets.html

exchange	index	level	change	%change	date	time (ET)
Amsterdam	AEX Index	334.53	+8.20	2.51%	Mar 25	11:40:00
Brussels	Bel 20	2346.82	+25.27	1.09%	Mar 25	12:03:00
Frankfurt	DAX	3811.92	+85.85	2.30%	Mar 25	11:45:00
London	FTSE	4373.60	+64.20	1.49%	Mar 25	11:36:00
Paris	CAC	3570.40	+51.95	1.48%	Mar 25	12:01:00
Switzerland	Market Index	5592.30	+88.30	1.60%	Mar 25	11:32:00

be binding for both parties or for one party only, with the other party reserving the option to exercise or not. If the underlying asset is not traded, for example if the underlying is an index, some kind of cash settlement has to take place. A cash settlement is a process by which the terms of an option contract are fulfilled through the payment or receipt in dollars of the amount by which the option is in-the-money as opposed to delivering or receiving the underlying stock. Derivatives are traded in organized exchanges as well as over the counter (OTC derivatives). Examples of derivatives include forwards, futures, options, caps, floors, swaps, and many others. By forming portfolios utilizing a variety of derivatives and underlying assets, one can substantially reduce her risk exposure, when an appropriate strategy is considered.

Derivative contracts provide an easy and straightforward way to both reduce risk (hedging), and to bear extra risk (speculating). As noted above, in any market conditions every security bears some risk. Using active derivative management involves isolating the factors that serve as the sources of risk.

1.4.1 Index options

An index call (put) option is a right to buy (sell) an index in the beforehand (at time $t = 0$) determined exercise price E, to the given expiration date $t = T$. The European option can be exercised at the expiration date only, unlike the American option which can be exercised at any time up to the date the option expires.

Each option (as well as each derivative) is characterized by a payoff function. If the index value I is compounded from several underlying stocks, $I = \sum_{i=1}^{n} w_i S_i$ where $w_i > 0$ are the index weights, then for the value of a Call option at expiry T we have

$$V_{Call}(I, T) = max(I - E, O)\,. \tag{3}$$

Table 4: Call options on Dow Jones Industrial index DJX-E (CBOE) from Mar 29,2004 at 11:16 ET. The price of DJX was 103.33, i.e. 10333 points.

Calls	Last Sale	Net	Bid	Ask	Vol	Open Int
04 Apr 96.00 (DJV DR-E)	7.10	+0.70	7.20	7.60	3	713
04 Apr 97.00 (DJV DS-E)	4.60	pc	6.20	6.60	0	180
04 Apr 98.00 (DJV DT-E)	5.20	+0.40	5.30	5.70	1	1318
04 Apr 99.00 (DJV DU-E)	3.90	pc	4.40	4.70	0	688
04 Apr 100.0 (DJV DV-E)	3.60	+0.80	3.50	3.70	79	13636
04 Apr 101.0 (DJV DW-E)	2.80	+0.80	2.65	2.90	183	3773
04 Apr 102.0 (DJV DX-E)	2.00	+0.60	1.90	2.10	278	4728
04 Apr 103.0 (DJV DY-E)	1.35	+0.45	1.25	1.40	89	11608
04 Apr 104.0 (DJV DZ-E)	0.85	+0.30	0.70	0.90	676	10440
04 Apr 105.0 (DJV DA-E)	0.40	+0.15	0.35	0.45	128	7936
04 Apr 106.0 (DJV DB-E)	0.15	–	0.15	0.20	156	24256
04 Apr 107.0 (DJV DC-E)	0.10	–	0	0.15	10	5920
04 Apr 108.0 (DJV DD-E)	0.05	–	0	0.15	50	4642
04 Apr 109.0 (DJV DE-E)	0.05	pc	0	0.15	0	700
04 Apr 110.0 (DJV DF-E)	0.15	pc	0	0.15	0	5013
04 May 96.00 (DJV ER-E)	7.60	pc	7.50	7.90	0	100
04 May 97.00 (DJV ES-E)	0	pc	6.60	7.00	0	0
04 May 98.00 (DJV ET-E)	0	pc	5.80	6.20	0	0
04 May 99.00 (DJV EU-E)	5.10	pc	5.00	5.40	0	20
04 May 100.0 (DJV EV-E)	3.40	pc	4.20	4.50	0	704
04 May 101.0 (DJV EW-E)	3.60	+0.40	3.40	3.70	3	206
04 May 102.0 (DJV EX-E)	2.35	pc	2.75	3.00	0	777
04 May 103.0 (DJV EY-E)	2.10	+0.15	2.15	2.40	1	2886
04 May 104.0 (DJV EZ-E)	1.35	pc	1.65	1.85	0	1776
04 May 105.0 (DJV EA-E)	1.30	+0.25	1.20	1.30	2	3474
04 May 106.0 (DJV EB-E)	0.80	+0.10	0.80	1.00	200	3340
04 May 107.0 (DJV EC-E)	0.60	+0.05	0.50	0.65	7	2234
04 May 108.0 (DJV ED-E)	0.40	+0.10	0.30	0.40	60	3282
04 May 109.0 (DJV EE-E)	0.20	pc	0.15	0.35	0	1140
04 May 110.0 (DJV EF-E)	0.15	pc	0.05	0.25	0	4319

For a Put option we have the payoff

$$V_{Put}(I,T) = max(E - I, O)\,. \tag{4}$$

We will talk about the payoff functions in more detail in Section 2.1.3.

One of the biggest option exchanges in the world is the Chicago Board Options Exchange (CBOE), founded in 1973. The European counterpart of CBOE is EUREX, founded in 1998.

Several rules are applied to index options. These are the general rules concerning the trading conditions. These rules are not same for all indices. The characteristics of the options on DJIA, S&P 500, or Dow 10 are the following ([9]):

- the basic element is a hundredth part of the index value,
- the option price is given in decimal numbers and one point equals 100 USD (or another currency),
- the minimal price change is 0.05 (5 USD) for those with value less than 3.0, and 0.1 (10 USD) for the others,
- the last trading day is one day (usually Thursday) before the third Friday in the month of expiration,
- the expiration date is Saturday after the third Friday in the month of expiration,
- the expiration cycle consists of the three nearest months, longer cycles are based on the quartal shift,
- index options are cash-settled, which means that, when exercising the option, the owner has not right to buy (sell) the index for the strike price, but only the financial settlement is applied, whereby the owner gains the money amount of the difference between the strike price and the current index value multiplied by 100,
- the opening index value of the day before expiration (usually Friday) is taken to be the current index value used for calculating the settlement rate.

Apart from the plain vanilla contracts which are American or European, a lot of other exotic options have appeared recently, mostly as OTC contracts. These include Asian options, digital options, lookback options, etc.

In Table 4 we present the values of Call options on Dow Jones Industrial index from March 29, 2004. The price of DJX was 103.33, i.e. 10333 points. The table comes from the internet site http://quote.cboe.com. The column "Last Sale" shows the price for which the last sale took place. The column "Net" means the net change in the index value. "Bid" means the price at which a buyer is willing to buy an option. "Ask" means the price at which a seller is offering to sell an option. "Vol" denotes the volume of transactions done. "Open Int" means Open Interest, i.e. the number of outstanding option contracts in the exchange market.

2 Stock index derivative pricing model

2.1 The n-dimensional Black-Scholes equation

Let us consider an European option price of which depends on prices of n risky assets and on the time to expiry τ. The goal of this section is to recall a derivation of a mathematical model describing the price evolution of this derivative.

The derivation process consists of two steps. At first we find a stochastic equation describing the evolution of the derivative value V in dependance on the time t and the prices $S_i, i = 1, ..., n$ of assets comprised in the index. Details of the derivation can be found in Kwok's book [6]. Then we will construct a self-financing portfolio comprising assets, options on these assets and riskless bonds. We shall extend the one-asset case considered in [12].

2.1.1 A stochastic equation for a derivative value

Now, our goal is to derive a stochastic differential equation for the index derivative value.

Let $S_i, i = 1, 2, ..., n$, be the price of the asset i and $V(S_1, S_2, ..., S_n, \tau)$ the value of the derivative on a given set of assets (where τ is the time to expiration). Assume that the asset prices behave according to lognormal diffusion processes:

$$\frac{dS_i}{S_i} = \mu_i dt + \sigma_i dZ_i, \qquad i = 1, 2, ..., n, \tag{5}$$

where μ_i and σ_i mean the expected rate of return and the volatility of asset i, dZ_i is the Wiener process' differential for stock i. Let ρ_{ij} indicate the correlation coefficient of dZ_i and dZ_j ,

$$E(dZ_i dZ_j) = \rho_{ij} dt, \qquad i, j = 1, 2, ..., n, i \neq j. \tag{6}$$

Each process Z_i can be considered as a linear combination of Wiener processes w_k with independent increments, $k = 1, 2, ..., n$. Thus

$$dZ_i = \sum_{k=1}^{n} \tilde{\sigma}_{ik} dw_k, \qquad i, j = 1, 2, ..., n,$$

with

$$E(dw_i dw_j) = \begin{cases} dt & i = j \\ 0 & i \neq j. \end{cases}$$

Then

$$E(dZ_i dZ_j) = \sum_{k=1}^{n}\sum_{l=1}^{n} \tilde{\sigma}_{ik}\tilde{\sigma}_{jl} E(dw_k dw_l) = \sum_{k=1}^{n} \tilde{\sigma}_{ik}\tilde{\sigma}_{jk} dt.$$

Therefore the correlation coefficients satisfy

$$\rho_{ij} = \sum_{k=1}^{n} \tilde{\sigma}_{ik}\tilde{\sigma}_{jk}, \qquad i,j = 1,2,...,n, i \neq j.$$

Lemma 2.1 (Itô lemma for functions with a vector argument) *Let* $f = f(\mathbf{x}, t) : \mathbb{R}^n \times \mathbb{R}$ *be a smooth function of a vector argument* $\mathbf{x} = (x_1, ..., x_n)^T$*. Let the variables* $x_i, i = 1,2,...,n$*, satisfy a system of stochastic differential equations:*

$$dx_i = \mu_i(\mathbf{x}, t)dt + \sum_{k=1}^{n} \bar{\sigma}_{ik}(\mathbf{x}, t) dw_k$$

where $\mathbf{w} = (w_1, w_2, ..., w_n)^T$ *is Wiener processes' vector whose components have increments independent of each other. In a vector form it reads as follows:*

$$d\mathbf{x} = \mu(\mathbf{x}, t)dt + K(\mathbf{x}, t)d\mathbf{w}$$

where $K(\mathbf{x}, t) = \{\bar{\sigma}_{ij}(\mathbf{x}, t)\}_{i,j=1,...,n}$ *is an* $n \times n$ *matrix. Then the first differential of* f *is given by*

$$df = \left(\frac{\partial f}{\partial t} + \frac{1}{2} Tr(K^T \nabla_x^2 f K)\right) dt + \nabla_x f d\mathbf{x}$$

where $\nabla_x f$ *is the gradient of the function* f *and* $Tr(K^T \nabla_x^2 f K) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} \sum_{k=1}^{n} \bar{\sigma}_{ik}\bar{\sigma}_{jk}$*.*

In our case we have $\mu_i(\mathbf{x}, t) = \mu_i S_i$, $\bar{\sigma}_{ik}(\mathbf{x}, t) = \sigma_i \tilde{\sigma}_{ik} S_i$, and according to Lemma 2.1

$$dV = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 V}{\partial S_i \partial S_j}\sigma_i\sigma_j S_i S_j \sum_{k=1}^{n}\tilde{\sigma}_{ik}\tilde{\sigma}_{jk}\right) dt + \sum_{i=1}^{n}\frac{\partial V}{\partial S_i} dS_i\,,$$

so the price of a derivative satisfies the stochastic differential equation

$$dV = \frac{\partial V}{\partial t}dt + \sum_{i=1}^{n}\frac{\partial V}{\partial S_i}dS_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j}dt\,. \tag{7}$$

2.1.2 Construction of a riskless portfolio

Our next step is construction of a riskless portfolio consisting of a certain number of several kinds of assets, options on these assets, and zero coupon riskless bonds. The case of a portfolio consisting of a certain number of one asset, options on this unique asset, and zero coupon riskless bonds, is developed in [12]. We will follow the three economic fundamentals of the classical Black-Scholes theory which are the zero increase in the investment, a self-financing portfolio, and a risk averse investor. The condition of zero increase in investment means that the portfolio elements are dynamically bought and sold so that no more investment is needed to maintain the zero riskiness of the portfolio. The condition of a self-financing portfolio means that a purchase or a sale of one portfolio element is compensated by a sale or a purchase of another one.

The elements of our portfolio at time t are: Q_V pieces of options with price V, the number Q_{S_i} of assets i of price S_i $(i = 1, ..., n)$, and the value B of zero coupon riskless bonds. Then the assumption of the zero increase in investment can be expressed as

i) $\sum_{i=1}^{n} S_i Q_{S_i} + V Q_V + B = 0$,

and the condition of a self-financing portfolio can be written as

ii) $\sum_{i=1}^{n} S_i dQ_{S_i} + V dQ_V + \delta B = 0$

where dQ_{S_i}, dQ_V and δB mean the change in the number of assets, the number of options, and the change of the value of riskless bonds contained in the portfolio needed for self-financing the portfolio.

By differentiating $i)$ we obtain

$$\sum_{i=1}^{n} (S_i dQ_{S_i} + dS_i Q_{S_i}) + dV Q_V + V dQ_V + dB = 0.$$

The continuously compounded bonds follow the equation $dB = rBdt$. They are also used for the self-financing mechanism of the portfolio. Because of that the total change of the value of the bonds is $dB = rBdt + \delta B$. Using this knowledge and condition $ii)$ we realize that

$$\sum_{i=1}^{n} dS_i Q_{S_i} + dV Q_V + rBdt = 0$$

must hold. Proceeding by replacing dS_i and dV with (5) and (7), after some simple rearrangements we obtain

$$\tilde{\mu}dt + \sum_{i=1}^{n}(\sigma_i S_i Q_{S_i} + \frac{\partial V}{\partial S_i}\sigma_i S_i Q_V)dZ_i = 0$$

where

$$\tilde{\mu} = \sum_i \mu_i S_i Q_{S_i} + \frac{\partial V}{\partial t}Q_V + \sum_i \frac{\partial V}{\partial S_i}\mu_i S_i Q_V + \\ + \frac{1}{2}\sum_{i,j}\rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j}Q_V + rB\,. \tag{8}$$

In order to achieve a riskless portfolio we will eliminate all stochastic terms in the last equation by assuming

$$\sigma_i S_i Q_{S_i} + \frac{\partial V}{\partial S_i}\sigma_i S_i Q_V = 0, \qquad i = 1, 2, ..., n.$$

Then the relation for the number of assets i and the number of options reads

$$\frac{Q_{S_i}}{Q_V} = -\frac{\partial V}{\partial S_i}, \qquad i = 1, 2, ..., n. \tag{9}$$

With such a composition of the portfolio we can eliminate its risky part. Consequently, $\tilde{\mu} = 0$. After dividing (8) by nonzero Q_V and substituting $B = -\sum_i S_i Q_{S_i} - VQ_V$ from $i)$ we obtain

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + r\sum_{i=1}^{n} S_i \frac{\partial V}{\partial S_i} - rV = 0\,. \tag{10}$$

By replacing time t by time to expiry τ (i.e. $\tau = T - t$ where T is the expiration time) we have

$$\frac{\partial V}{\partial \tau} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + r\sum_{i=1}^{n} S_i \frac{\partial V}{\partial S_i} - rV\,, \tag{11}$$

$$0 < S_1, ..., S_n < \infty, \qquad \tau > 0.$$

Equation (11) is called *the generalized $n-$dimensional Black-Scholes partial differential equation* describing the evolution of the price V of a derivative in time. Notice that the equation does not depend on expected rates of return μ_i of the assets. For different types of derivatives different initial conditions at $\tau = 0$ have to be added to equation (11). The initial conditions to (11) are the final conditions at $t = T$ to (10), i.e. the terminal conditions (payoff functions) of the derivatives.

Based on the partial differential equations theory it is necessary to add also the boundary conditions to equation (11) resp. (10).

2.1.3 The terminal conditions

The terminal (payoff) conditions to the Black-Scholes equation (10) are determined by the type of a derivative and its specifics. These conditions are the functions of the derivative value in dependance of the index value at the expiry time T.

In the case of a Call option the payoff function read as:

$$V(\mathbf{S},T) = \max\left(\left(\sum_i w_i S_i\right) - E, 0\right) \tag{12}$$

where w_i are weights corresponding to the index definition, E is the exercise price of the index. The following idea is the reason for a Call option payoff function to have the form (12): if the current index value $\sum_i w_i S_i$ at time T exceeded the value E, the option premium would be the difference between the current value $\sum_i w_i S_i$ of the index and the exercise value E, i.e. $\sum_i w_i S_i - E$. On the other hand, if the current value does not exceed the exercise value E, we do not exercise the option, so it has a zero value.

We can follow a similar idea when deriving the payoff function for a Put option. If the current value $\sum_i w_i S_i$ of the index at time T is greater than the exercise value E, the option will not be exercised, thus its value is zero. On the other hand, if the current index value $\sum_i w_i S_i$ is less than the exercise value E, the value of the option at the expiry time $t = T$ is equal to $E - \sum_i w_i S_i$. Hence, the terminal condition for a Put option is the function

$$V(\mathbf{S},T) = \max\left(E - \left(\sum_i w_i S_i\right), 0\right) . \tag{13}$$

2.2 Analytical solution to the Black-Scholes equation

In this section we focus on the analytical solution to the Black-Scholes partial differential equation (11) with an arbitrary initial condition, as it is done in [6]. We shall look for the solution in the convolutionary form

$$V(\mathbf{S}, T-\tau) = e^{-r\tau} \int_{\mathbb{R}^n} V(\xi, T)\psi(\xi; \mathbf{S}, \tau) d\xi \tag{14}$$

where $\mathbf{S} = (S_1, S_2, ..., S_n)^T$ and $\psi(\xi; \mathbf{S}, \tau)$ is a function of an $n-$dimensional variable $\xi = (\xi_1, \xi_2, ..., \xi_n)^T$.

Seeking for the function $V(\mathbf{S}, \tau)$ of the form (14) and satisfying (11) is equivalent to the problem of finding a function ψ solving

$$\begin{cases} \frac{\partial \psi}{\partial \tau} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 \psi}{\partial S_i \partial S_j} + r \sum_{i=1}^n S_i \frac{\partial \psi}{\partial S_i} \\ \psi(\xi; \mathbf{S}, 0) = \delta(\xi - \mathbf{S}) \end{cases} \tag{15}$$

where

$$\delta(x) = \begin{cases} 0 & x \neq 0 \\ +\infty & x = 0 \end{cases}$$

represents the Dirac distribution.

We can easily prove this simply by substituting (14) into (11) and by recognizing that the function $V(\xi, T)$ in (14) is arbitrary.

We apply some transformations of variables on function ψ with effort to transform equation (15) to a one with a known solution. More precisely, $\frac{\partial\psi}{\partial\tau} = \frac{1}{2}\sum_i \frac{\partial^2\psi}{\partial x_i^2}$.

At first we adopt the following change of variables:

$$y_i = \frac{1}{\sigma_i}\left(r - \frac{\sigma_i^2}{2}\right)\tau + \frac{1}{\sigma_i}\ln S_i\,, \qquad i = 1, 2, ..., n, \tag{16}$$

and put $\Phi(\mathbf{y}, \tau) = \psi(\mathbf{S}, \tau)$ where $\mathbf{y} = (y_1, y_2, ..., y_n)^T$. For partial derivatives we obtain the relations

$$\frac{\partial\psi}{\partial\tau} = \sum_{i=1}^{n} \frac{\partial\Phi}{\partial y_i}\frac{\partial y_i}{\partial\tau} + \frac{\partial\Phi}{\partial\tau} = \frac{\partial\Phi}{\partial\tau} + \sum_{i=1}^{n}\frac{1}{\sigma_i}\left(r - \frac{\sigma_i^2}{2}\right)\frac{\partial\Phi}{\partial y_i}\,,$$

$$\frac{\partial\psi}{\partial S_i} = \frac{\partial\Phi}{\partial y_i}\frac{\partial y_i}{\partial S_i} = \frac{1}{\sigma_i S_i}\frac{\partial\Phi}{\partial y_i}, \qquad i = 1, 2, ..., n,$$

$$\frac{\partial^2\psi}{\partial S_i \partial S_j} = \frac{1}{\sigma_i\sigma_j S_i S_j}\frac{\partial^2\Phi}{\partial y_i \partial y_j}, \qquad i, j = 1, 2, ..., n,\ i \neq j,$$

$$\frac{\partial^2\psi}{\partial S_i^2} = -\frac{1}{\sigma_i S_i^2}\frac{\partial\Phi}{\partial y_i} + \frac{1}{(\sigma_i S_i)^2}\frac{\partial^2\Phi}{\partial y_i^2}, \qquad i = 1, 2, ..., n,$$

and then (15) becomes an equation with constant coefficients

$$\frac{\partial\Phi}{\partial\tau} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\frac{\partial^2\Phi}{\partial y_i \partial y_j}\,, \qquad -\infty < y_i < \infty, \tau > 0\,.\,, \tag{17}$$

This equation can be rewritten in the following matrix form:

$$\frac{\partial\Phi}{\partial\tau} = \frac{1}{2}\left(\ \frac{\partial}{\partial y_1},\ \ \frac{\partial}{\partial y_2},\ \ \cdots\ \ ,\frac{\partial}{\partial y_n}\ \right)\mathbf{R}\begin{pmatrix} \frac{\partial}{\partial y_1} \\ \frac{\partial}{\partial y_2} \\ \vdots \\ \frac{\partial}{\partial y_n} \end{pmatrix}\Phi \tag{18}$$

where $\mathbf{R}$ is an $n \times n$ symmetric matrix whose elements are $\mathbf{R}_{ij} = \rho_{ij}, i, j = 1, 2, ..., n, i \neq j$. The diagonal entries are taken to be 1.

The matrix $\mathbf{R}$ is always positive semidefinite. Henceforth, we will assume that $\mathbf{R}$ is positive definite. This is the case when stochastic processes $\{dZ_i, i = 1, ..., n\}$, are not perfectly correlated. It follows from the matrix theory that there exists an orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{Q}^T\mathbf{R}\mathbf{Q} = \mathbf{\Lambda}$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose entries are the eigenvalues of $\mathbf{R}$.

Aiming to eliminate the matrix $\mathbf{R}$ from (18) it is required to apply such a transformation of the variable $\mathbf{y} = (y_1, y_2, ..., y_n)^T$ into a new variable $\mathbf{x} = (x_1, x_2, ...x_n)^T$ by which for the "old" function Φ and the new one (after transforming) $\tilde{\Phi}$ the relation

$$\left(\frac{\partial\tilde{\Phi}}{\partial x_1}, \quad \frac{\partial\tilde{\Phi}}{\partial x_2}, \quad \ldots \quad , \frac{\partial\tilde{\Phi}}{\partial x_n} \right)^T = \mathbf{\Lambda}^{\frac{1}{2}}\mathbf{Q}^T \left(\frac{\partial\Phi}{\partial y_1}, \quad \frac{\partial\Phi}{\partial y_2}, \quad \ldots \quad , \frac{\partial\Phi}{\partial y_n} \right)^T$$

holds, because then (18) changes into

$$\frac{\partial\tilde{\Phi}}{\partial\tau} = \frac{1}{2} \left(\frac{\partial}{\partial x_1}, \quad \frac{\partial}{\partial x_2}, \quad \ldots \quad , \frac{\partial}{\partial x_n} \right) \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \tilde{\Phi}\,.$$

Therefore we look for a transformation of $\mathbf{y} = (y_1, y_2, ..., y_n)^T$ into $\mathbf{x} = (x_1, x_2, ..., x_n)^T$ such that for the gradients of the original function Φ and the function $\tilde{\Phi}$ the following identity holds:

$$(\nabla_x\tilde{\Phi})^T = \mathbf{\Lambda}^{\frac{1}{2}}\mathbf{Q}^T(\nabla_y\Phi)^T \tag{19}$$

where $\tilde{\Phi}(\mathbf{x}(\mathbf{y})) = \Phi(\mathbf{y})$, and $\mathbf{x} = \mathbf{x}(\mathbf{y})$ is the required transformation. We have

$$(\nabla_y\Phi)^T = \left(\sum_{j=1}^n \frac{\partial\tilde{\Phi}}{\partial x_j}\frac{\partial x_j}{\partial y_1}, \quad \ldots \quad , \sum_{j=1}^n \frac{\partial\tilde{\Phi}}{\partial x_j}\frac{\partial x_j}{\partial y_n} \right)^T =$$

$$= \sum_{j=1}^n \frac{\partial\tilde{\Phi}}{\partial x_j} \left(\frac{\partial x_j}{\partial y_1}, \quad \ldots \quad , \frac{\partial x_j}{\partial y_n} \right)^T =$$

$$= \begin{pmatrix} \frac{\partial x_1}{\partial y_1}, & \frac{\partial x_2}{\partial y_1}, & \ldots & , \frac{\partial x_n}{\partial y_1} \\ \vdots & & & \vdots \\ \frac{\partial x_1}{\partial y_n}, & \frac{\partial x_2}{\partial y_n}, & \ldots & , \frac{\partial x_n}{\partial y_n} \end{pmatrix} \begin{pmatrix} \frac{\partial\tilde{\Phi}}{\partial x_1} \\ \vdots \\ \frac{\partial\tilde{\Phi}}{\partial x_n} \end{pmatrix} = \mathbf{J}^T(\nabla_x\tilde{\Phi})^T$$

where $\mathbf{J}$ denotes the Jacobi matrix of the transformation $\mathbf{x} = \mathbf{x}(\mathbf{y})$. Comparing with (19) we obtain

$$\mathbf{J} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{Q}^T,$$

which implies that the transformation $\mathbf{x} = \mathbf{x}(\mathbf{y})$ of variables has the form

$$\mathbf{x} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{Q}^T\mathbf{y}. \tag{20}$$

We can apply it because $\mathbf{R}$ is positive definite, i.e. all its eigenvalues are positive (nonzero). By applying it in (18) we obtain the following $n-$dimensional diffusion equation

$$\frac{\partial\tilde{\Phi}}{\partial\tau} = \frac{1}{2}\Delta\tilde{\Phi} \tag{21}$$

where Δ is the so called Laplace operator defined as $\Delta\tilde{\Phi} = \sum_{i=1}^{n} \frac{\partial^2\tilde{\Phi}}{\partial x_i^2}$.

Since the solution to (21) is generally known, we subsequently obtain a function satisfying (15) by returning to the original variables. In [13] it is shown that the fundamental solution to (21) is

$$\tilde{\Phi}(\mathbf{x}, \tau) = \frac{1}{(2\pi\tau)^{\frac{n}{2}}} \exp(-\frac{1}{2\tau}\mathbf{x}^T\mathbf{x}) \tag{22}$$

and it satisfies the initial condition

$$\tilde{\Phi}(\mathbf{x}, 0) = \delta(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \neq \mathbf{0} \\ +\infty & \mathbf{x} = \mathbf{0} \end{cases}$$

representing the Dirac distribution. By transforming $\mathbf{x}$ back to $\mathbf{y}$ we obtain the function

$$\Phi(\mathbf{y}, \tau) = \tilde{\Phi}(\mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{Q}^T\mathbf{y}, \tau) = \frac{1}{(2\pi\tau)^{\frac{n}{2}}} \exp(-\frac{1}{2\tau}\mathbf{y}^T\mathbf{R}^{-1}\mathbf{y}) \tag{23}$$

solving (18) with the initial condition

$$\Phi(\mathbf{y}, 0) = \sqrt{\det\mathbf{R}}\,\delta(\mathbf{y}).$$

Let us apply the following substitution:

$$w_i = \frac{\ln\xi_i - \ln S_i - (r - \frac{\sigma_i^2}{2}\tau)}{\sigma_i}, \qquad i = 1, 2, ..., n, \tag{24}$$

whose Jacobian is $(\prod_i \sigma_i \prod_i \xi_i)^{-1}$. Then

$$\psi(\xi; \mathbf{S}, \tau) = \frac{1}{(2\pi\tau)^{\frac{n}{2}} \sqrt{\det \mathbf{R}} \prod_i \sigma_i \prod_i \xi_i} \exp\left(-\frac{1}{2\tau} \mathbf{w}^T \mathbf{R}^{-1} \mathbf{w}\right) \tag{25}$$

satisfies (15) with the initial condition

$$\psi(\xi; \mathbf{S}, 0) = \begin{cases} 0 & \xi \neq \mathbf{S} \\ +\infty & \xi = \mathbf{S}. \end{cases}$$

But then the function

$$W(\mathbf{S}, \tau) = \int_{R^n} \psi(\xi; \mathbf{S}, \tau) W_0(\xi) d\xi \tag{26}$$

solves (15) with arbitrary initial condition $W(\mathbf{S}, 0) = W_0(\mathbf{S})$. Finally, the function

$$V(\mathbf{S}, T - \tau) = e^{-r(T-\tau)} \int_{R^n} \psi(\xi; \mathbf{S}, \tau) V_0(\xi) d\xi \tag{27}$$

is the desired solution to the Black-Scholes partial differential equation (11) with the initial condition $V(\mathbf{S}, T) = V(\mathbf{S}, T - 0) = V_0(\mathbf{S})$. We have found a solution in the form of the terminal function in convolution with the solution of parabolic equation (15).

Hence, if we want to calculate the value of a derivative on an index, we have to solve integral (27). For example, if the index is comprised of $n = 500$ stocks, we should solve a 500-dimensional integral. This is a very difficult problem. In addition, the function $V_0(\xi)$ appearing in the integral expression is in principle arbitrary and then (27) can become analytically unsolvable.

Therefore, we shall apply an another method for looking for a solution to the Black-Scholes partial differential equation (11). We will consider the following form of the solution:

$$V(\mathbf{S}, T - \tau) = e^{-r\tau} \psi(\mathbf{S}, \tau) = e^{-r\tau} \phi(\mathbf{y}, \tau) = e^{-r\tau} \tilde{\phi}(\mathbf{x}, \tau) \tag{28}$$

where $\tilde{\phi}$ is a solution to

$$\frac{\partial \tilde{\Phi}}{\partial \tau} = \frac{1}{2} \Delta_{\mathbf{x}} \tilde{\Phi} \,. \tag{29}$$

The following relations for the initial conditions hold:

$$\tilde{\phi}(\mathbf{x}, 0) = \phi(\mathbf{y}, 0) = \psi(\mathbf{S}, 0) = V(\mathbf{S}, T) = V_0(\mathbf{S}) \,.$$

The relations between the individual variables are known from (16) and (20):

$$\begin{aligned} &\mathbf{x} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{Q}^T\mathbf{y}\,, \\ &y_i = \frac{1}{\sigma_i}\left(r - \frac{\sigma_i^2}{2}\right)\tau + \frac{1}{\sigma_i}\ln S_i\,, \qquad i = 1, 2, ..., n\,. \end{aligned} \tag{30}$$

Then the function $V(\mathbf{S}, T-\tau)$ solves equation (11).

If we find a solution to (29), the function $V(\mathbf{S}, T-\tau)$ is easily to obtain by (28). Therefore, in the following chapter we will focus only on finding a solution to (29), hence

$$\left\{ \begin{array}{ll} \frac{\partial\tilde{\Phi}}{\partial\tau} = \frac{1}{2}\Delta_{\mathbf{x}}\tilde{\Phi} & \mathbf{x} \in \mathbb{R}^n, \tau \in [0, T] \\ \tilde{\Phi}(\mathbf{x}, 0) = \tilde{\Phi}^0(\mathbf{x}) & \mathbf{x} \in \mathbb{R}^n. \end{array} \right. \tag{31}$$

3 Numerical methods

In the previous chapter we derived explicit formula (27) for computing the value of a derivative whose underlying is not a unique asset but a set of assets. For the different types of derivatives and their specifications, the initial condition $V_0(\xi)$ (i.e. the payoff function of the derivative) varies. In general, $V_0(\xi)$ is an arbitrary function.

It would be very difficult and time expensive to solve the corresponding n-dimensional integral for a large dimension n (e.g. $n = 100$) given by (27), or even unsolvable. Several methods were applied worldwide to solve this problem. The most wide-spread is the use of Monte Carlo simulation, presented for example by Milstein $\&$ Shoenmakers ([10]) and Levy ([7]). Lo and Hui adopted an algebraic approach ([8]). Another approach is applied by Reisinger in [11].

Henceforth, let us concentrate only on looking for a solution to parabolic equation (29). For simplification of notation, we shall deal with function (32) defined below. The adjustment of our results for (32) to obtain results for (29) is straightforward.

We shall apply the additive operator splitting method to solve (32). First, we pay attention to the following section, in which we discretize the problem.

3.1 Full space-time discretization

3.1.1 Discretization in time

Have a look at the numerical solution to the general equation

$$\begin{cases} \frac{\partial u}{\partial \tau} - \Delta u = 0 & \mathbf{x} \in \mathbb{R}^n, \tau \in [0,T] \\ u(\mathbf{x},0) = u^0(\mathbf{x}) & \mathbf{x} \in \mathbb{R}^n. \end{cases} \tag{32}$$

First, we discretize (32) in time by dividing the time interval $[0,T]$ into m parts of equal length k. We denote by $u^j(\mathbf{x})$ a numerical approximation of $u(\mathbf{x},\tau)$ at time $\tau_j = jk$, i.e.

$$u^j(\mathbf{x}) \approx u(\mathbf{x}, jk), \qquad \mathbf{x} \in \mathbb{R}^n,\ j = 0, 1, ..., m\,.$$

After substituting the time derivative by the time difference

$$\frac{\partial u}{\partial \tau}(\mathbf{x}, jk) \approx \frac{u^j(\mathbf{x}) - u^{j-1}(\mathbf{x})}{k}\,,$$

the semidiscretization of (32) reads as:

$$\frac{u^j(\mathbf{x}) - u^{j-1}(\mathbf{x})}{k} - (Au^j)(\mathbf{x}) = 0 \tag{33}$$

where A represents the Laplace operator Δ, i.e. $Au = \Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}$. After rearranging (33) we obtain

$$u^j = (\mathbf{I} - k\mathbf{A})^{-1} u^{j-1}, \qquad j = 1, 2, ..., m. \tag{34}$$

3.1.2 Spatial discretization

Secondly, we discretize the spatial variable in each direction with the same number d of (internal) dividing points and the spatial step of size h. It means that the whole problem "lives" in an n-dimensional cube $\Omega = (-L, L)^n$ where L is large enough. Then for the spatial step we have $h = \frac{2L}{d}$. The procedure discussed in the following sections is not limited to such a spatial discretization but it could lead to a more complicated scheme with a more general spatial grid. We stick to this simplification for clarity of the idea.

We put the gridpoints into a sequence in this way: we place the origin of the co-ordinate system into one of the corners of the cube. The grid is already created (with the same number of dividing points in each direction). We number the dividing points on the individual axes, starting with number 1 and continuing "outwards". Each gridpoint has a characteristic consisting of n elements: in the i-th place there is the order number of the appropriate dividing point in direction i. Thereby we obtain the names of all gridpoints. In order to present the way of ordering we consider the following "low-dimensional" examples for $n = 2, d = 3$, and $n = 3, d = 3$ (see Figure 2).

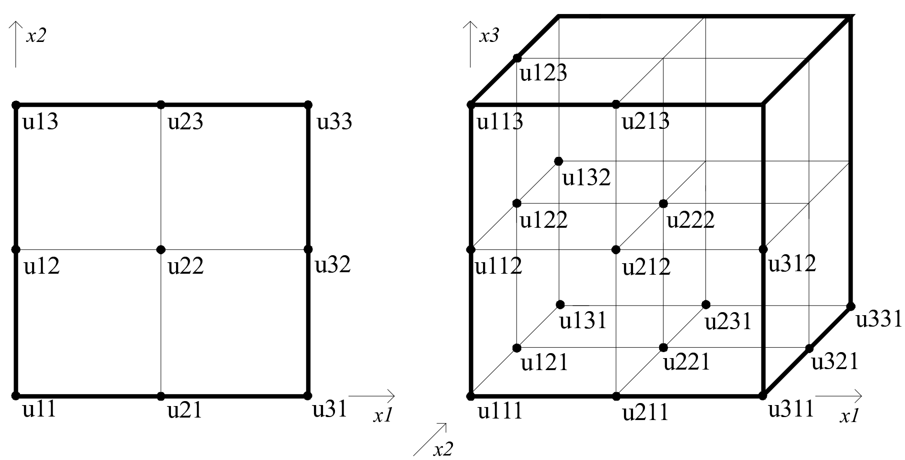

Figure 2: Spatial grid for $a)n = 2, d = 3$, $b)n = 3, d = 3$.

For $n = 2, d = 3$ the order of gridpoints is
$u_{11}, u_{21}, u_{31}, u_{12}, u_{22}, u_{32}, u_{13}, u_{23}, u_{33}$.

For $n = 3, d = 3$ their order is
$u_{111}, u_{211}, u_{311}, u_{121}, u_{221}, u_{321}, u_{131}, u_{231}, u_{331}$,
$u_{112}, u_{212}, u_{312}, u_{122}, u_{222}, u_{322}, u_{132}, u_{232}, u_{332}$,
$u_{113}, u_{213}, u_{313}, u_{123}, u_{223}, u_{323}, u_{133}, u_{233}, u_{333}$.

We use the central differences for the approximation of the second partial derivatives with respect to the space variables $x_i, i = 1, ..., n$. In the case of $n = 2$ with the spatial variable $\mathbf{x} = (x_1, x_2)$ the approximation in a gridpoint with characteristics $(p, q), p, q = 1, ..., d$ has the form

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \approx \frac{u_{p+1,q} + u_{p-1,q} + u_{p,q+1} + u_{p,q-1} - 4u_{p,q}}{h^2},$$

taking the boundary values $u_{0,.}, u_{.,0}, u_{d+1,.}, u_{.,d+1}$, equal to 0. Thus, the matrix representing the discretization of the Laplace operator for $n = 2, d = 3$ is then of the form

$$\mathbf{A} = \frac{1}{h^2}\left(\begin{array}{ccc|ccc|ccc}
-4 & 1 & 0 & 1 & & & & & \\
1 & -4 & 1 & & 1 & & & & \\
0 & 1 & -4 & & & 1 & & & \\
\hline
1 & & & -4 & 1 & 0 & 1 & & \\
& 1 & & 1 & -4 & 1 & & 1 & \\
& & 1 & 0 & 1 & -4 & & & 1 \\
\hline
& & & 1 & & 0 & -4 & 1 & \\
& & & & 1 & & 1 & -4 & 1 \\
& & & & & 1 & 0 & 1 & -4
\end{array}\right)$$

where h is the spatial step (same for all components of the spatial variable).

3.1.3 The boundary conditions

We cannot forget the boundary conditions to (32) in the domain $\Omega = (-L, L)^n$. For a single stock Call option (where the stock price S is considered to be in $(0, +\infty]$), the boundary conditions are ([12]):

$$V(0, t) = V_0(0) = 0,$$

$$\lim_{S \to +\infty} V(S, t) = V_0(S).$$

In the case of index Call option, the boundary conditions will have the form

$$\lim_{|\mathbf{S}|\to\mathbf{0}} V(\mathbf{0},t) = V_0(\mathbf{0}) = 0,$$

$$\lim_{|\mathbf{S}|\to+\infty} V(\mathbf{S},t) = V_0(\mathbf{S})$$

where $\mathbf{S} = (S_1, ..., S_n)^T$ is a vector of prices of the stocks comprised in the index.

In the case of equation (32) in the domain $\Omega = (-L, L)^n$, the boundary condition will be

$$u(\mathbf{x}, \tau) = u_0(\mathbf{x}) \tag{35}$$

for $\mathbf{x} \in \partial\Omega$, i.e. for some i with $|x_i| = L$.

3.2 The additive operator splitting technique

After the spatial discretization of (34), the operator $\mathbf{A}$ becomes a high-dimensional matrix. For example, for $n = 20$ and the number $d = 10$ of dividing points in the spatial variable (in each direction), this matrix is of type $d^n \times d^n = 10^{20} \times 10^{20}$. That's a huge number, but the discretization used isn't very fine. Moreover, for our purposes we need n to be much greater, e.g. $n = 100$ for S&P 100 or $n = 500$ for S&P 500. To solve the full-discretized problem (34) we had to solve a system of $N = d^n$ linear equations, which is for large n and d in general a very difficult task. A question arises, how to simplify solving such a high-dimensional problem.

In this section we apply the additive operator splitting method developed in [3], [16], [17]. This method is widely used in image processing. The application of the additive operator splitting technique to index derivative pricing resides in finding an approximate solution to (31) using the AOS scheme and subsequently in applying the appropriate transformations of the variable $\mathbf{x}$ into the variable $\mathbf{S}$ given in (30). We obtain the price of the derivative by using (28).

The main idea of the additive operator splitting method is the following: let us have operators (or matrices) $\mathbf{B_i}, i = 1, ..., n$. Replace the arithmetical mean by the harmonic one, i.e.

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{B}_i \qquad \longleftrightarrow \qquad n(\sum_{i=1}^{n}\mathbf{B}_i^{-1})^{-1}\,.$$

Then the inverse is approximated by

$$(\frac{1}{n}\sum_{i=1}^{n}\mathbf{B}_i)^{-1} \qquad \longleftrightarrow \qquad \frac{1}{n}\sum_{i=1}^{n}\mathbf{B}_i^{-1}\,.$$

Let $\mathbf{A} = \sum_{i=1}^{n} \mathbf{A_i}$ where $\mathbf{A_i}$ introduces the second partial derivative with respect to x_i. Then $\mathbf{I} - k\mathbf{A} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{B_i}$ where $\mathbf{B_i} = \mathbf{I} - kn\mathbf{A_i}$. The AOS approximation to $(\mathbf{I} - k\mathbf{A})^{-1}$ is

$$(\mathbf{I} - k\mathbf{A})^{-1} \qquad \longleftrightarrow \qquad \frac{1}{n}\sum_{i=1}^{n}(\mathbf{I} - kn\mathbf{A}_i)^{-1}.$$

Finally, the AOS-approximation of the time discretization (34) reads as:

$$\bar{u}^j = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{I} - kn\mathbf{A}_i)^{-1}\bar{u}^{j-1}, \qquad \bar{u}^0 = u^0, j = 1, ..., m. \tag{36}$$

The order of the introduced approximation will be discussed later. Now, notice that the solution to (36) can be written as

$$\bar{u}^j = \frac{1}{n}\sum_{i=1}^{n}\bar{v}_i^j$$

where

$$\bar{v}_i^j = (\mathbf{I} - kn\mathbf{A_i})^{-1}\bar{u}^{j-1}.$$

It means that

$$\frac{\bar{v}_i^j - \bar{u}^{j-1}}{k} - n\mathbf{A_i}\bar{v}_i^j = 0$$

and hence $\bar{v}_i^j$ is the approximation of the solution to the *one-dimensional* parabolic equation

$$\begin{cases} \frac{\partial v}{\partial \tau} - n\frac{\partial^2 v}{\partial x_i^2} = 0, & x_i \in \mathbb{R}, \tau \in (0, T), \\ v(\tilde{x}_1, ..., \tilde{x}_{i-1}, x_i, \tilde{x}_{i+1}, ..., \tilde{x}_n, 0) = & \\ \qquad = \bar{u}^{j-1}(\tilde{x}_1, ..., \tilde{x}_{i-1}, x_i, \tilde{x}_{i+1}, ..., \tilde{x}_n), & x_i \in \mathbb{R} \end{cases} \tag{37}$$

at time k, i.e.

$$\bar{v}_i^j(\mathbf{x}) \approx v(\mathbf{x}, jk).$$

However, the solution to (37) can be found in an explicit way

$$v(\mathbf{x}, \tau) = \int_{\mathbb{R}} G(x_i - \xi_i, \tau) v(x_1, ..., x_{i-1}, \xi_i, x_{i+1}, ..., x_n, 0) d\xi_i \,.$$

Hence, the approximation of v reads as

$$\bar{v}_i^j(\mathbf{x}) \approx \int_{\mathbb{R}} G(x_i - \xi_i, k)\bar{u}^{j-1}(x_1, ..., x_{i-1}, \xi_i, x_{i+1}, ..., x_n)d\xi_i \tag{38}$$

where $G(x, \tau)$ is the one-dimensional Green function for (37):

$$G(x, \tau) = \frac{1}{\sqrt{4\pi n\tau}} \exp\left(-\frac{x^2}{4n\tau}\right). \tag{39}$$

We can finally assemble the AOS approximation scheme for solving (32). It consists of solving the one-dimensional parabolic equations only and uses the AOS technique. We use the notation $\bar{\bar{u}}^j$ for a solution obtained by this method:

$$\bar{\bar{u}}^j = \frac{1}{n}\sum_{i=1}^{n} \bar{\bar{v}}_i^j, \qquad j = 1, 2, ..., m \tag{40}$$

where

$$\bar{\bar{v}}_i^j(\mathbf{x}) = \int_{\mathbb{R}} G(x_i - \xi_i, k)\bar{\bar{u}}^{j-1}(x_1, ..., x_{i-1}, \xi_i, x_{i+1}, ..., x_n)d\xi_i \tag{41}$$

and $\bar{\bar{u}}^{j-1} = \bar{u}^{j-1}$.

3.3 The order of the AOS approximation

We proceed by answering the question of the precision of the approximation scheme (40) derived in the previous section. We shall prove that

$$\|\bar{u}^j - u^j\| \leq const.k^{\nu_1}, \tag{42}$$

$$\|\bar{\bar{u}}^j - \bar{u}^j\| \leq const.k^{\nu_2} \tag{43}$$

where $\nu_1, \nu_2 \in \mathbb{N}$ and $\|.\|$ is a norm, what after putting together gives the estimate for the norm $\|\bar{\bar{u}}^j - u^j\|$ of the difference between the analytical solution of one-dimensional parabolic equations constructed by the AOS scheme and the precise solution to (32).

Now we shall aim to answer the first problem (42). Recall that

$$u^j = \mathbf{M}u^{j-1}, \qquad j = 1, ..., m,$$

$$\bar{u}^j = \overline{\mathbf{M}}\bar{u}^{j-1}, \qquad j = 1, ..., m,$$

where $\mathbf{M} = (\mathbf{I} - k\mathbf{A})^{-1}$ and $\overline{\mathbf{M}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{I} - kn\mathbf{A_i})^{-1}$. According to [3], the following theorem holds. We also recall the proof of this theorem because we need a precise form of the constant $C > 0$ appearing below.

At first, we recall the definition of the simultaneously diagonalizable matrices.

Definition 3.1 ([14]) *The set $\mathcal{F}$ of matrices $\mathbf{A}_1, ..., \mathbf{A}_n$ of the same size is simultaneously diagonalizable iff each of the matrices in $\mathcal{F}$ transforms to a diagonal one by a common similarity. Thus, matrices $\mathbf{A}_1, ..., \mathbf{A}_n$ are simultaneously diagonalizable iff there exists a matrix $\mathbf{W}$ (of the same size) with*

$$\mathbf{W^{-1}A_iW} = \mathbf{\Lambda_i}\,, \qquad i = 1, ..., n,$$

where $\mathbf{\Lambda_i}$ is a diagonal matrix.

Theorem 3.1 ([3]) *Let $n \in \mathbb{N}, k \geq 0$, and let $\mathbf{A_1}, ..., \mathbf{A_n} \in \mathbb{R}^{N\times N}$ be simultaneously diagonalizable matrices with eigenvalues in the left half plane. Then there exists a constant C with*

$$\|\mathbf{M} - \overline{\mathbf{M}}\| \leq Ck^2$$

where $\mathbf{M} = (\mathbf{I} - k\mathbf{A})^{-1}$ and $\overline{\mathbf{M}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{I} - kn\mathbf{A_i})^{-1}$.

Proof 3.1 Based on the assumption of the simultaneous diagonalizability of matrices $\mathbf{A_i}, i = 1, ..., n$, let us diagonalize these matrices by a matrix $\mathbf{W}$ consisting of the eigenvectors of any $\mathbf{A_i}$:

$$\mathbf{W^{-1}A_iW} = \mathbf{\Lambda_i} = \text{diag}(\lambda_{i,j}, 1 \leq j \leq N)$$

where $\mathbf{\Lambda_i}$ are the diagonal matrices composed from the corresponding eigenvalues. Hence,

$$\mathbf{W^{-1}}(\mathbf{I} - nk\mathbf{A_i})\mathbf{W} = \mathbf{I} - nk\mathbf{\Lambda_i},$$

$$\mathbf{W^{-1}}(\mathbf{I} - k\sum_{i=1}^{n}\mathbf{A_i})\mathbf{W} = (\mathbf{I} - k\sum_{i=1}^{n}\mathbf{\Lambda_i}),$$

and subsequently

$$\mathbf{W^{-1}}[(\mathbf{I} - k\sum_{i=1}^{n}\mathbf{A_i})^{-1} - \frac{1}{n}\sum_{i=1}^{n}(\mathbf{I} - nk\mathbf{A_i})^{-1}]\mathbf{W} =$$

$$= (\mathbf{I} - k\sum_{i=1}^{n}\mathbf{\Lambda_i})^{-1} - \frac{1}{n}\sum_{i=1}^{n}(\mathbf{I} - nk\mathbf{\Lambda_i})^{-1}$$

is a diagonal matrix where the l-th diagonal entry is given by

$$q_l = \Theta(k\sum_{i=1}^{n}\lambda_{i,l}) - \frac{1}{n}\sum_{i=1}^{n}\Theta(nk\lambda_{i,l}), \qquad \Theta(x) = \frac{1}{1-x}.$$

The Taylor expansion of the function Θ at $x_0 = 0$ reads as:

$$\Theta(x) = 1 + x + \frac{x^2}{(1-\xi)^3}, \qquad \xi = \xi(x) \in (0, x).$$

It yields

$$q_l = 1 + k\sum_{i=1}^{n}\lambda_{i,l} + \frac{k^2(\sum_{i=1}^{n}\lambda_{i,l})^2}{(1-\xi)^3} - \frac{1}{n}\sum_{i=1}^{n}\left(1 + nk\lambda_{i,l} + \frac{(nk\lambda_{i,l})^2}{(1-\xi_i)^3}\right) =$$

$$= k^2\left(\frac{(\sum_{i=1}^{n}\lambda_{i,l})^2}{(1-\xi)^3} - \frac{1}{n}\sum_{i=1}^{n}\frac{(n\lambda_{i,l})^2}{(1-\xi_i)^3}\right) =$$

$$=: k^2 g(\lambda_{1,l}, ..., \lambda_{n,l}).$$

It is assumed that we can find compact sets Q_i contained in the left complex half plane which enclose all eigenvalues of $\mathbf{A_i}$. Consequently, the continuous function g attains its maximum on $Q := Q_1 \times ... \times Q_n$,

$$\widehat{C} := \max\{|g(\mathbf{z})|; \mathbf{z} \in Q\}.$$

We thus have $|q_l| \le \widehat{C}k^2$ for $l = 1, ..., N$. The statement follows from

$$\|(\mathbf{I} - k\sum_{i=1}^{n}\mathbf{A_i})^{-1} - \frac{1}{n}\sum_{i=1}^{n}(\mathbf{I} - nk\mathbf{A_i})^{-1}\| \le$$

$$\le \|\mathbf{W}\|\;\|\mathbf{W}^{-1}\|\;\|(\mathbf{I} - k\sum_{i=1}^{n}\mathbf{\Lambda_i})^{-1} - \frac{1}{n}\sum_{i=1}^{n}(\mathbf{I} - nk\mathbf{\Lambda_i})^{-1}\| \le Ck^2$$

where

$$C = \widehat{C}\|\mathbf{W}\|\|\mathbf{W}^{-1}\|\,. \tag{44}$$

$\diamondsuit$

Later we will prove the validity of the assumptions in Theorem 3.1 for matrices $\mathbf{A_i}$ representing the second partial derivatives with respect to individual variables. Let us therefore assume that we can apply the above mentioned theorem.

We have

$$u^j - \bar{u}^j = \mathbf{M}u^{j-1} - \overline{\mathbf{M}}\bar{u}^{j-1} = (\mathbf{M} - \overline{\mathbf{M}})\bar{u}^{j-1} + \mathbf{M}(u^{j-1} - \bar{u}^{j-1})$$

and

$$\|u^j - \bar{u}^j\| \le \|\mathbf{M} - \overline{\mathbf{M}}\| \, \|\bar{u}^{j-1}\| + \|\mathbf{M}(u^{j-1} - \bar{u}^{j-1})\|$$

where $\|.\|$ is a yet unspecified norm. Introducing notation $\theta = max_{\lambda\in\sigma(\mathbf{M})}|\lambda|$ and referring to Theorem 3.2 presented below we will realize that

$$\|u^j - \bar{u}^j\| \le \|\mathbf{M} - \overline{\mathbf{M}}\| \, \|\bar{u}^{j-1}\| + \theta\|u^{j-1} - \bar{u}^{j-1}\|.$$

Next, supposing that

$$\|\bar{u}^j\| < \widetilde{C}, \qquad j = 0, 1, ..., m, \tag{45}$$

for some $\widetilde{C} \in \mathbb{R}$, we obtain

$$\|u^j - \bar{u}^j\| \le (1 + \theta + \theta^2 + ... + \theta^{j-1})Kk^2$$

with a constant $K = C\widetilde{C}$. Moreover, if $\theta < 1$ then

$$\|u^j - \bar{u}^j\| \le \frac{1}{1-\theta}Kk^2. \tag{46}$$

3.3.1 The forming matrices

For the validity of (46) we have to verify $\theta < 1$. Linear algebra theory defines the spectral radius of a matrix $\mathbf{B}$ as $r(\mathbf{B}) = \sup_{\lambda\in\sigma(\mathbf{B})} |\lambda|$ where $\sigma(\mathbf{B})$ is the spectrum of $\mathbf{B}$. By [4] (see also [5]) we have

$$r(\mathbf{B}) = \lim_{n\to\infty} \sqrt[n]{\|\mathbf{B}^n\|} = \sup_{n\to\infty} \sqrt[n]{\|\mathbf{B}^n\|}\,. \tag{47}$$

The next theorem deals with the real symmetric matrices, which as we will see are our case [2]. We recall the proof for the reader's convenience.

[2]For more general matrices, the Ljapunov theorem can be used. It shows that (for matrix $\mathbf{B}$ with spectral radius less than 1) there exists an equivalent norm $\|.\|_*$ such that $\|\mathbf{B}\|_* < 1$.

Theorem. *Let $\mathbf{B}$ be an $N \times N$ matrix with $r(\mathbf{B}) < 1$. Then there exists a norm $\|.\|_*$ on $\mathbb{R}^N$ such that*

$$\|\mathbf{B}\|_* = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Bx}\|_*}{\|\mathbf{x}\|_*} = r(\mathbf{B}) < 1.$$

The use of this theorem leads (using similar techniques as in our case) to error estimates too, although they are of worse quality.

Theorem 3.2 *Let* $\mathbf{B}$ *be a real symmetric* $N \times N$ *matrix. Let* $\theta = r(\mathbf{B})$*. Then*

$$\|\mathbf{Bu}\| \leq \theta \|\mathbf{u}\| \qquad \forall \mathbf{u} \in \mathbb{R}^N$$

where $\|.\|$ *is the Euclidean norm on* $\mathbb{R}^N$*.*

Proof 3.2 First, we consider the case when the eigenvalues of $\mathbf{B}$ are simple. Let $\mathbf{e}_1, ..., \mathbf{e}_N$ be the eigenvectors of $\mathbf{B}$. From the matrix theory we know that the eigenvectors corresponding to simple eigenvalues are orthogonal. Moreover, let $\|\mathbf{e}_i\| = 1, i = 1, ..., N$.

Let us express the vector $\mathbf{u} \in \mathbb{R}^N$ using the vectors of the orthonormal basis $\{\mathbf{e}_i, i = 1, ..., N\}$:

$$\mathbf{u} = \sum_{i=1}^{N} u_i \mathbf{e}_i, \qquad u_i \in \mathbb{R}.$$

Then

$$\mathbf{Bu} = \sum_{i=1}^{N} u_i \lambda_i \mathbf{e}_i, \qquad u_i \in \mathbb{R}$$

and exploiting the orthonormality of vectors $\mathbf{e}_i, i = 1, ..., N$ and the definition of θ, we have

$$\|\mathbf{Bu}\|^2 = \sum_{i=1}^{N} u_i^2 \lambda_i^2 \leq \theta^2 \sum_{i=1}^{N} u_i^2 = \theta^2 \|\mathbf{u}\|^2 .$$

Consequently,

$$\|\mathbf{Bu}\| \leq \theta \|\mathbf{u}\| .$$

In the case of multiple eigenvalues, the proof is very similar and the statement holds. ♢

Our goal is to verify whether there is $\theta = r(\mathbf{M}) < 1$ in our case. Let us investigate the eigenvalues of $\mathbf{M} = (\mathbf{I} - k\mathbf{A})^{-1}$ where $\mathbf{A} = \sum_{i=1}^{n} \mathbf{A_i}$. In the following, we change the notation of matrix $\mathbf{A}$ into $\mathbf{A_{(n)}}$ to express the dependance of $\mathbf{A}$ on n. In the previous section we discretized the spatial variable in each direction with the same number of dividing points d and the spatial step h, and we showed an example of the matrix representing the discretization of the Laplace operator for $n = 2, d = 3$. Generally, this matrix for arbitrary n, d will have the form

$$\mathbf{A_{(n)}} = \frac{1}{h^2}\left(\begin{array}{c|c|c|c|c|c} \mathbf{H} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{I} & \mathbf{H} & \mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{H} & \ddots & \mathbf{0} & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \ddots & \mathbf{H} & \mathbf{I} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{H} \end{array}\right)$$

where $\mathbf{H} = \mathbf{A_{(n-1)}} - 2\mathbf{I}$. The size of the matrix $\mathbf{A_{(n)}}$ is $N \times N, N = d^n$. In addition, $\mathbf{A_{(n)}} = \sum_{i=1}^{n} \mathbf{A_i}$. At the same time,

$$\mathbf{A_i} = \frac{1}{h^2}(-2\mathbf{I} + \mathbf{S_i})$$

where $\mathbf{I}$ is the identity matrix of the same type as $\mathbf{A_i}$, and $\mathbf{S_i}$ (of the dimension $N \times N$, too) is a special matrix of the form

$$\mathbf{S_i} = \left(\begin{array}{c|c|c|c} \mathbf{Z_i} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{Z_i} & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z_i} \end{array}\right)$$

where $\mathbf{Z_i}$ is a $d^i \times d^i$ matrix whose entries are equal to 1 on the $(d^{i-1}+1)$-th diagonal, direction "to the right" and "downwards"; otherwise they are equal to 0. Matrices $\mathbf{A_i}$ represent the second partial derivative with respect to x_i.

For eigenvalues λ_{ij} of each matrix $\mathbf{A_i}, i = 1, ..., n$ we have

$$\frac{-4}{h^2} < \lambda_{ij} < 0, \qquad j = 1, ..., N. \tag{48}$$

It can be shown easily:

$$< \mathbf{A_i v}, \mathbf{v} >= \frac{1}{h^2} < (-2\mathbf{I} + \mathbf{S_i})\mathbf{v}, \mathbf{v} >= \frac{1}{h^2}[< -2\mathbf{Iv}, \mathbf{v} > + < \mathbf{S_i v}, \mathbf{v} >] =$$

$$= \frac{1}{h^2}[-2\sum_{j=1}^{N} v_j^2 + 2 \sum_{\substack{j=1,...,N \\ j \neq pd^i - q \\ p=1,...,d^{n-i} \\ q=0,1,...,d^{i-1}-1}} v_j v_{j+d^{i-1}}] =$$

$$= -\frac{1}{h^2}[\sum_{\substack{j=1,...,N \\ j \neq pd^i - q \\ p=1,...,d^{n-i} \\ q=0,1,...,d^{i-1}-1}} (v_j + v_{j+d^{i-1}})^2 + \sum_{\substack{j=pd^i+1,...,pd^i+d^{i-1} \\ p=0,1,...,\frac{d^{i-1}}{d^n-d^i}}} v_j^2] <$$

$$< 0 \qquad \forall \mathbf{v} \neq \mathbf{0}.$$

The reader can examine this procedure for, say, $n = 3, d = 3$. Now realize the following: if we define $\tilde{\mathbf{A}}_\mathbf{i} = \mathbf{A}_\mathbf{i} + \frac{4}{h^2}\mathbf{I}$, then for the spectrum we have

$$\sigma(\tilde{\mathbf{A}}_\mathbf{i}) = \sigma(\mathbf{A}_\mathbf{i}) + \frac{4}{h^2}. \tag{49}$$

In the same way as in the previous case we obtain

$$< \tilde{\mathbf{A}}_\mathbf{i}\mathbf{v}, \mathbf{v} >= \frac{1}{h^2} < (2\mathbf{I} + \mathbf{S}_\mathbf{i})\mathbf{v}, \mathbf{v} >= \frac{1}{h^2}[< 2\mathbf{I}\mathbf{v}, \mathbf{v} > + < \mathbf{S}_\mathbf{i}\mathbf{v}, \mathbf{v} >] =$$

$$= \frac{1}{h^2}[2\sum_{j=1}^{N} v_j^2 + 2 \sum_{\substack{j=1,...,N \\ j \neq pd^i - q \\ p=1,...,d^{n-i} \\ q=0,1,...,d^{i-1}-1}} v_j v_{j+d^{i-1}}] =$$

$$= \frac{1}{h^2}[\sum_{\substack{j=1,...,N \\ j \neq pd^i - q \\ p=1,...,d^{n-i} \\ q=0,1,...,d^{i-1}-1}} (v_j + v_{j+d^{i-1}})^2 + \sum_{\substack{j=pd^i+1,...,pd^i+d^{i-1} \\ p=0,1,...,\frac{d^{i-1}}{d^n-d^i}}} v_j^2] >$$

$$> 0 \qquad \forall \mathbf{v} \neq \mathbf{0}.$$

As a direct aftermath of this, all eigenvalues of $\tilde{\mathbf{A}}_\mathbf{i}$ are positive and thanks to (49) the first inequality in (48) is valid. At the same time, we confirmed that the assumption "eigenvalues belong to the left half plane" from Theorem 3.1 is fulfilled.

We look right now on the second assumption of this theorem, i.e. the simultaneous diagonalizability of matrices $\mathbf{A}_\mathbf{i}$.

Theorem 3.3 ([2], p.78-80) *Let* $\mathbf{A}, \mathbf{B}$ *be the real symmetric* $N \times N$ *matrices. The necessary and satisfying condition for the orthogonal matrix* $\mathbf{T}$ *with*

$$\mathbf{T}^T\mathbf{A}\mathbf{T} = diag(\lambda_1, ..., \lambda_N), \qquad \mathbf{T}^T\mathbf{B}\mathbf{T} = diag(\mu_1, ..., \mu_N)$$

to exist, where $\lambda_1, ..., \lambda_N$ *are the eigenvalues of* $\mathbf{A}$ *and* $\mu_1, ..., \mu_N$ *are the eigenvalues of* $\mathbf{B}$*, is the commutativity of matrices* $\mathbf{A}$ *a* $\mathbf{B}$*.*

We recall the proof for reader's convenience.
Proof 3.3 Suppose that the matrix $\mathbf{A}$ has simple eigenvalues. From

$$\mathbf{A}\mathbf{x}^i = \lambda_i \mathbf{x}^i$$

and from the commutativity of $\mathbf{A}, \mathbf{B}$ we have

$$\mathbf{A}(\mathbf{B}\mathbf{x}^i) = \mathbf{B}(\mathbf{A}\mathbf{x}^i) = \mathbf{B}(\lambda_i \mathbf{x}^i) = \lambda_i(\mathbf{B}\mathbf{x}^i). \tag{50}$$

As we can see, $\mathbf{B}\mathbf{x}^i$ is the eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ_i. Because of the simplicity of the eigenvalues, eigenvectors belonging to the same eigenvalue must be proportional. Hence,

$$\mathbf{B}\mathbf{x}^i = \mu_i \mathbf{x}^i, \qquad i = 1, ..., N,$$

where μ_i are scalars which have to be the eigenvalues of $\mathbf{B}$. It means that $\mathbf{A}, \mathbf{B}$ have the same eigenvectors $\mathbf{x}^1, ..., \mathbf{x}^N$. The matrix $\mathbf{T}$ can be taken as

$$\mathbf{T} = (\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N).$$

Now consider the case when λ is an eigenvalue of multiplicity s with the corresponding eigenvectors $\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^s$. Then (50) implies

$$\mathbf{B}\mathbf{x}^i = \sum_{j=1}^{s} c_{ij} \mathbf{x}^j, \qquad i = 1, ..., s\,,$$

for some $c_{ij} \in \mathbb{R}$. Notice that the matrix $\mathbf{C} = (c_{ij})$ is symmetric. It can be deduced from the orthonormality of vectors $\mathbf{x}^i$ and the symmetry of $\mathbf{B}$:

$$< \mathbf{x}^j, \mathbf{B}\mathbf{x}^i >= c_{ij} =< \mathbf{B}\mathbf{x}^j, \mathbf{x}^i >= c_{ji}.$$

Let us concentrate now on the question whether it is possible to construct such a linear combination of vectors $\mathbf{x}^i$ which would be the eigenvector of $\mathbf{B}$. Consider the linear combination $\sum_{i=1}^{s} a_i \mathbf{x}^i$:

$$\mathbf{B}(\sum_{i=1}^{s} a_i \mathbf{x}^i) = \sum_{i=1}^{s} a_i \mathbf{B}\mathbf{x}^i = \sum_{i=1}^{s} a_i (\sum_{j=1}^{s} c_{ij} \mathbf{x}^j) = \sum_{j=1}^{s} (\sum_{i=1}^{s} c_{ij} a_i) \mathbf{x}^j.$$

It can be inferred that by choosing a_i so that

$$\sum_{i=1}^{s} c_{ij} a_i = r_1 a_j, \qquad j = 1, ..., s, \tag{51}$$

we come to

$$\mathbf{B}(\sum_{i=1}^{s} a_i \mathbf{x}^i) = r_1(\sum_{i=1}^{s} a_i \mathbf{x}^i)$$

implying that r_1 is the eigenvalue of $\mathbf{B}$ and $\sum_{i=1}^{s} a_i \mathbf{x}^i$ is the appropriate eigenvector.

From (51) we have that r_1 is the eigenvalue of $\mathbf{C}$ and a_i are the components of the appropriate eigenvector. Hence, if $\mathbf{T_s}$ is the s-dimensional orthogonal map leading the matrix $\mathbf{C}$ to the diagonal form, then the vectors $\mathbf{z}^i$ obtained by

$$\begin{pmatrix} \mathbf{z}^1 \\ \mathbf{z}^2 \\ \vdots \\ \mathbf{z}^s \end{pmatrix} = \mathbf{T_s} \begin{pmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^s \end{pmatrix},$$

which are the images of the orthonormal set of vectors, appear to be the eigenvectors common to both matrices $\mathbf{A}, \mathbf{B}$. Carrying out the analogous mappings belonging to each multiplicity of the eigenvalue we can construct the matrix $\mathbf{T}$.

The necessity of the statement follows from the fact that two matrices for which

$$\mathbf{A} = \mathbf{T} diag(\lambda_1, ..., \lambda_N)\mathbf{T}^T, \qquad \mathbf{B} = \mathbf{T} diag(\mu_1, ..., \mu_N)\mathbf{T}^T,$$

commute if $\mathbf{T}$ is orthogonal. $\diamondsuit$

Both the theorem and the proof can be found in [2]. But we deal with not only two matrices, but with a set of n matrices. This situation is the subject of the following theorem.

Theorem 3.4 ([14], p.77) *The set $\mathcal{F}$ of normal matrices of the same size commutes if and only if it is simultaneously unitary diagonalizable (i.e. if each of the matrices in $\mathcal{F}$ transforms to a diagonal one by a common unitary similarity).*

Note: The (square) matrix $\mathbf{A}$ is called normal if $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$ where $\mathbf{A}^* = \bar{\mathbf{A}}^T$. The real symmetric matrices are a special case of normal matrices.

Hence, in order to prove the simultaneous diagonalizability of matrices $\mathbf{A}_i$ representing the discretized second partial derivatives we need to show the commutativity of the couples $\mathbf{A}_i, \mathbf{A}_j, i \neq j$. This is a consequence of the following statement.

Theorem 3.5 ([18]) *Let f be a positively definite quadratic form, let g a next arbitrary quadratic form on the Euclidean space V. Then there exists a base in V such that f and g become*

$$f(\mathbf{x}) = x_1^2 + ... + x_n^2,$$

$$g(\mathbf{x}) = \lambda_1 x_1^2 + ... + \lambda_n x_n^2.$$

Our matrices $\mathbf{A_i}, i = 1, ..., n$, are negatively definite because their eigenvalues are negative. Hence, the matrices $-\mathbf{A_i}, i = 1, ..., n$ are positively definite. In addition, they are real and symmetric, and therefore we know from the matrix theory that each of them is similar to a diagonal matrix by an orthogonal matrix $\mathbf{C}$ consisting of their orthonormal eigenvectors. This means that for $\mathbf{A_i}$ there exists $\mathbf{C} \in O(N)$ such that

$$\mathbf{C}^{-1}\mathbf{A_i}\mathbf{C} = \mathbf{C}^T\mathbf{A_i}\mathbf{C} = diag(\lambda_{i,1}, ..., \lambda_{i,N}) =: -\mathbf{D_i},$$

thus

$$\mathbf{A_i} = -\mathbf{C}\sqrt{\mathbf{D_i}}\sqrt{\mathbf{D_i}}\mathbf{C}^{-1}.$$

Then by (3.5)

$$\mathbf{A_j} = -\mathbf{C}\sqrt{\mathbf{D_i}}\mathbf{D_j}\sqrt{\mathbf{D_i}}\mathbf{C}^{-1}$$

where $\mathbf{D_j}$ is a diagonal matrix. Since the diagonal matrices commute, matrices $\mathbf{A_i}, \mathbf{A_j}$ commute too.

The commutativity of the mentioned matrices can be clear also intuitively by realizing the interchangeability of the order of deriving with respect to two different variables.

It means that all assumptions of Theorem 3.1 are fulfilled, so we really had the right to use it.

In addition, let us make the following remarks to the assumptions of Theorem 3.1:

i) eigenvalues in the left half plain for each matrix $\mathbf{A_i}, i = 1, ..., n$ ensure the regularity of matrices $(\mathbf{I} - nk\mathbf{A_i})$ and therefore also their invertibility,

ii) the simultaneous diagonalizability of $\mathbf{A_i}, i = 1, ..., n$ and eigenvalues in the left half plane ensure the regularity and invertibility of $(\mathbf{I} - k\sum_{i=1}^{n}\mathbf{A_i})$,

iii) the assumption of eigenvalues in the left half plane is restrictive to the choice of matrices $\mathbf{A_i}$, but allows an arbitrary k. On the other hand, for k small enough the statement holds for arbitrary matrices.

iv) Although the assumption of the eigenvalues belonging to the left half plane isn't important for the statement to be true, it is needed to prove the stability of the numerical scheme.

Now, return to our goal which is to estimate the eigenvalues $\mu_j, j = 1, ..., n$ of matrix $\mathbf{M} = (\mathbf{I} - k\mathbf{A})^{-1}$. By the Spectral mapping theorem ([4]) we come to a knowledge that

$$\mu_j = \frac{1}{1 - k\gamma_j}, \qquad j = 1, ..., N,$$

where γ_j denotes the eigenvalue of $\mathbf{A}$. We know that $\gamma_j = \sum_{i=1}^{n} \lambda_{ij}$ because the matrices $\mathbf{A_i}, i = 1, ..., n$ are simultaneously diagonalizable. At first we emphasize the validity of a strengthened version of (48). We know that $S^n = \{\mathbf{v} \in \mathbb{R}^n; \|\mathbf{v}\| = 1\}$ is a compact set in $\mathbb{R}^n$, so the function $\mathbf{v} \mapsto < \mathbf{Av}, \mathbf{v} >$ (which is continuous) attains its maximum in S^n. Let us label it with δ, and from (48) it follows that $\max_{\|\mathbf{v}\|=1} < \mathbf{Av}, \mathbf{v} >= \delta < 0$. Hence,

$$\frac{-4}{h^2} < \lambda_{ij} < \delta < 0\,, \qquad i = 1, ..., n,\ j = 1, ..., N. \tag{52}$$

Then, for all j, $\frac{-4n}{h^2} < \gamma_j < n\delta < 0$ and that yields

$$0 < \mu_j < \frac{1}{1 - kn\delta} < 1. \tag{53}$$

Hereby we proved that $\theta = r(\mathbf{M}) < 1$.

3.3.2 Error estimates

We have shown that $\theta = r(\mathbf{M}) < 1$ and hence the estimate (46) holds for a constant K. Now, we proceed with estimating the value of the constant K for which we have $K = C\widetilde{C}$. Let us recall that C is a constant appearing in Theorem 3.1, talking about the estimate of the norm $\|\mathbf{M} - \overline{\mathbf{M}}\|$. The constant $\widetilde{C}$ originates in the assumption that $\|\bar{u}^j\| < \widetilde{C}, j = 0, 1, ..., m$.

We pay our next attention to estimation of the constant C from Theorem 3.1. Recall that by (44) from Proof 3.1 the constant C originates from the constant $\widehat{C} = \max\{|g(\mathbf{z})|; \mathbf{z} \in Q\}$ where Q is a compact set in $\mathbb{R}^n$ containing all eigenvalues of matrices $\mathbf{A_i}$, and g is a function defined in the proof. By (52) we have $Q = (< -\frac{4}{h^2}, \delta >)^n$. Then for function g defined as

$$g(\mathbf{z}) = \frac{(\sum_{i=1}^{n} z_i)^2}{(1 - \xi)^3} - \frac{1}{n} \sum_{i=1}^{n} \frac{(nz_i)^2}{(1 - \xi_i)^3}$$

where $z_i \in < -\frac{4}{h^2}, \delta >$, $i = 1, ..., n$, $\xi \in (k\sum_{i=1}^{n} z_i, 0)$, $\xi_i \in (nkz_i, 0)$, $i = 1, ..., n$, the following estimate holds:

$$g(\mathbf{z}) \leq \frac{(n(-\frac{4}{h^2}))^2}{(1 - \xi)^3} - n \sum_{i=1}^{n} \frac{\delta^2}{(1 - \xi_i)^3} \leq$$

$$\leq n^2 \frac{16}{h^4} - n^2 \delta^2 \frac{1}{(1 - kn(-\frac{4}{h^2}))^3} \leq n^2 \frac{16}{h^4}.$$

Then for q_l from the proof of the theorem we have $q_l \leq n^2 \frac{16}{h^4} k^2$ for $l = 1, ..., N$. For the next purpose, let us denote the diagonal matrix from the proof whose entries are q_l as $\mathbf{Q}$.

Using the Euclidean norm on $\mathbb{R}^N$ defined as

$$\|\mathbf{u}\| = \sqrt{\frac{1}{N} \sum_{i=1}^{N} u_i^2} \tag{54}$$

for a vector $\mathbf{u} \in \mathbb{R}^N$, according to the result from the proof we have

$$\|\mathbf{M} - \overline{\mathbf{M}}\| \leq \|\mathbf{W}\| \|\mathbf{W}^{-1}\| \|\mathbf{Q}\| \leq n^2 \frac{16}{h^4} k^2 = Ck^2$$

because $\mathbf{W}$ is orthogonal[3] and therefore $\|\mathbf{W}\| = \|\mathbf{W}^{-1}\| = 1$. We have shown that the estimate of the constant C depends on the space-discretization step h and the number n of underlying assets, and has the value

$$C = n^2 \frac{16}{h^4} . \tag{55}$$

Next, notice that

$$\|\overline{M}\| \leq \|\overline{M} - M\| + \|M\| \leq Ck^2 + \theta$$

where $\theta = r(\mathbf{M})$. Moreover, by (53) we know that $\theta = \theta(k) = \frac{1}{1 - kn\delta}$. The Taylor expansion of $\theta = \theta(k)$ in $k = 0$ yields

$$\theta(k) \leq 1 - O(k) + O(k^2)$$

whereby

$$\|\overline{M}\| \leq Ck^2 + 1 - O(k) + O(k^2) = 1 - O(k) + O(k^2) < 1$$

for $k << 1$. It is obvious that

$$\|\bar{u}^j\| \leq \|\overline{M}\|^j \|\bar{u}^0\| \leq \|\bar{u}^0\| = \|u^0\| =: \widetilde{C}$$

because $\|\overline{M}\| < 1$. Hereby we verified the validity of the assumption (45). Hence, $\widetilde{C} := \|u^0\|$.

[3] The orthogonality of the matrix $\mathbf{W}$ follows from the fact that the matrices $\mathbf{A}_i$ representing the second partial derivatives with respect to x_i are real and symmetric.

Now we can proceed with estimating of the norm in (46) where the constant $K = C\widetilde{C}$:

$$\|u^j - \bar{u}^j\| \leq \frac{1}{1-\theta} K k^2 = \frac{1}{1-\theta} n^2 \frac{16}{h^4} \|u^0\| k^2 \leq$$
$$\leq \frac{1}{1 - \frac{1}{1-nk\delta}} n^2 \frac{16}{h^4} \|u^0\| k^2 =$$
$$= \frac{(1-nk\delta)}{-\delta} n \frac{16}{h^4} \|u^0\| k = H \frac{k}{h^4}$$

where H is a constant. Therefore, the estimate of (42) is

$$\|u^j - \bar{u}^j\| \leq O(\frac{k}{h^4})\,. \tag{56}$$

Next, the estimate (43) of the norm of the difference between the analytical solution to a one-dimensional parabolic equation and its numerical counterpart based on the implicit (in time) Euler scheme is by [15] as follows:

$$\|\bar{\bar{u}}^j - \bar{u}^j\| \leq O(k + h^2)\,. \tag{57}$$

Finally, we can join these two estimates together in order to obtain the error estimate of our method:

$$\|\bar{\bar{u}}^j - u^j\| \leq \|\bar{\bar{u}}^j - \bar{u}^j\| + \|\bar{u}^j - u^j\| \leq O(k + h^2) + O(\frac{k}{h^4})$$
$$= O(k + h^2 + \frac{k}{h^4})\,. \tag{58}$$

By choosing the ratio of k and h it is possible to control this error. For example, the choice $k = h^6$ implies

$$\|\bar{\bar{u}}^j - u^j\| \leq O(h^2)\,. \tag{59}$$

It is an interesting result because the time-step k is usually taken to be $k = h^2$ which is a well-known CLF condition.

It is worth to mention that the norm (54) corresponds in the continuous case to the norm

$$\|u\| = \sqrt{\int_{\Omega} |u(\mathbf{x})|^2 d\mathbf{x}} \tag{60}$$

where $\Omega = (-L, L)^n$.

We have shown that the error of the additive operator splitting method used for solving equation (32) in comparison to the precise analytical solution to this equation is in general of order $O(k + h^2 + \frac{k}{h^4})$. Especially for $k = h^6$, it is of order $O(h^2)$.

3.4 Computational parallelism

The advantage of the AOS scheme resides in the computational parallelism it offers. At each time τ_j, problem (40) is solved. For that purpose, at each time τ_j, we have to solve n one-dimensional integrals given by (41). This process can be parallelized. The individual one-dimensional integrals can be simultaneously solved and at the end of these n procedures (with the overall execution time equal to the duration of one procedure only) the output consists of n functions. The results are to be averaged in order to obtain the solution u^{j+1} at the next time-step $\tau_{j+1} = (j+1)k$ in accordance to (40). This is done subsequently for all time-steps $\tau_j, j = 0, ..., m-1$.

Conclusions

In this paper, we considered pricing of index options, i.e. options whose underlying is a set of assets. We have shown that instead of solving a high-dimensional integral which is the analytical solution to the Black-Sholes partial differential equation governing the price of an option, a better approach exists.

We applied the additive operator splitting technique to find an approximate solution of the mentioned equation. By using this method, the complicated multi-dimensional problem was replaced by several simple one-dimensional problems.

We have shown that by an appropriate choice of the time-discretization step k and the spatial step h we can control the order of the error caused by the AOS method, and achieve the order of $O(h^2)$ for $k = h^6$.

The advantage of the used method in comparison with other approaches resides in parallelism of computations it offers.

Computational simulations and tests based on real data are to be done to prove the practical usefulness of the introduced method. From the view of the needed time and space, this problem offers a challenge for a whole next paper. Therefore, let this paper be considered just as a conceptual framework useful for multi-asset derivative pricing.

References

[1] J. Babušíková, M. Slodička, J. Weisz: *Numerické metódy*, Comenius University, Bratislava, 2000

[2] R. Bellman: *Vvedenie v teoriyu matric*, Nauka, Moskva, 1976

[3] B. Fischer, J. Modersitzki: *Fast diffusion registration*, http://www.math.mu-luebeck.de/workers/modersitzki/publications, 2001

[4] V. Hutson, J. S. Pym: *Applications of Functional Analysis and Operator Theory*, University of Sheffield, Academic Press, 1980

[5] L. V. Kantorovich, G. P. Akilov: *Funkcional'nyj analiz*, Nauka, Moskva, Glavnaya redakciya fiziko-matematicheskoj literatury, 1984

[6] Y. K. Kwok: *Mathematical Models of Financial Derivatives*, Springer-Verlag Singapore Pte. Ltd., 1998.

[7] G. Levy: *Multi-asset derivative pricing using quasi-random numbers and Monte Carlo simulation*, NAG Ltd UK, 2002

[8] C. F. Lo, C. H. Hui: *Pricing multi-asset financial derivatives with time-dependent parameters - Lie algebraic approach*, The Chinese University of Hong Kong, Hong Kong

[9] M. Lopušan: Diploma thesis, Comenius University, Bratislava, year of completion 2005

[10] G. N. Milstein, J. G. M. Shoenmakers: *Monte Carlo construction of hedging strategies against multi-asset European claims*, Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany; Ural State University, Ekaterinburg, Russia

[11] Ch. Reisinger: *Pricing Options on Big Baskets*, www.math.ethz.ch/ amatache/workshopslides/ reisinger.pdf

[12] D. Ševčovič: *Analytické a numerické metódy oceňovania finančných derivátov*, Comenius University, Bratislava, 2001

[13] D. Ševčovič: *Parciálne diferenciálne rovnice*, Comenius University, Bratislava, 1999

[14] F. Šik: *Lineární algebra zaměřená na numerickou analýzu*, Masaryk University, Brno, 1998

[15] E. Vitásek: *Numerické metódy*, SNTL, Prague, 1987

[16] J. Weickert: *Applications of Nonlinear Diffusion in Image Processing and Computer Vision*, Acta Mathematica Universitatis Comenianae, Vol. LXX (1), pp. 33-50, 2001

[17] J. Weickert, B. M. ter Haar Romeny, M. A. Viergever: *Efficient and Reliable Schemes for Nonlinear Diffusion Filtering*, IEEE, http://www.mia.uni-saarland.de/weickert/ publications.html, 1998

[18] P. Zlatoš: *Lineárna algebra*, Sylaby prednášky pre 1.ročník odboru Informatika na FI MU v Brne