Computational Finance Student Research Group

knfo.mimuw.edu.pl, knfo@mimuw.edu.pl

12 December 2018

Abstract

This document contains a brief description of the meetings of Computational Finance Student Research Group, which operates at the Faculty of Mathematics, Informatics and Mechanics – University of Warsaw.

Contents

1 Academic year 2017/18

1.1 20 October 2017 – Binomial Tree Model

Speakers: Michał Jaworski, Bartłomiej Polaczyk

The talk consisted of two following parts.

1. **Model CRR.** What is an option contract and how to price it in a binomial model. We examined the models that

   (a) have the **recombination property**, ie. the value after n steps depends only on the number of steps up (and down), not on their order,

   (b) have the **flat interest rate structure**, which means that the interest rate is constant in time,

   (c) have **good asymptotic properties**, that is, the distribution of the price of the underlying converges in law with diminishing time steps [1] to some limiting distribution.

   It was shown that the CRR (Cox-Ros-Rubinstein) model is the natural (with respect to the choice of parameters) binomial tree model satisfying the above criterions. In particular, it was shown that the limiting distribution is lognormal with parameters complying with the ones from the Black-Scholes model.

2. **Project proposals.** In the second part of the meeting, two projects for participants were proposed. Both of them were associated with the pricing of European and Asian options in the Binomial Tree and Black–Scholes model. More accurate description of the projects would be presented during the next meetings.

References

[1] Patric Billingsey (1999), *Convergence of Probability Measures, 2nd Edition*, John Wiley & Sons Inc.

[2] Marek Musiela, Marek Rutkowski (2005), *Martingale Methods in Financial Modelling*, Springer.

1.2 03 November 2017 – Stochastic Integral, Itô's Formula, Black–Scholes Model

Speakers: Michał Jaworski, Bartłomiej Polaczyk

The talk consisted of three following parts.

1. **Definition of Stochastic Integral.** We have presented some fundamental intuition standing behind the construction of stochastic integral, which is to be a random equivalent of Stielties integral, which means it should be approximated (in some sense) with appropriate sums, i.e.

$$\lim \sum X_{t_k}(M_{t_{k+1}-M_{t_k}}) = \int X_s \, dM_s.$$

[1] using Donsker Theorem, one can extrapolate this result to the uniform convergence of C.D.Fs on the whole time interval, see [1] section 8.

We have shown what conditions have to be met by processes X and M for the above limit to exists almost surely and in L^2 sense. On that basis we have defined stochastic differential equation and its (strong) solution.

2. **Itô's Formula.** By analogy to the fundamental theorem of calculus, saying that

$$df(x_t) = f'(x_t)\, dx_t \tag{1}$$

and using the informal interpretation

$$(dW_t)^2 = dt,$$

we have shown that in case of the functional of Wiener process of the form

$$dX_t = b\, dt + \sigma\, dW_t,$$

the expansion (1) should consist also of the elements of higher orders, i.e.

$$\begin{aligned} df(X_t) &= f'(X_t)\, dX_t + \frac{1}{2} f''(X_t)\,(dW_t)^2 \\ &= f'(X_t)\, dX_t + \frac{1}{2} f''(X_t)\, dt. \end{aligned}$$

The above result was extended onto the class of multivariate semimartingales.

3. **Black–Schoels Model.** At the last part of the talk Black–Scholes Model was presented. It assumes that market consists of:

- **risky asset** with price process given by the following SDE

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where $\mu \in \mathbb{R}, \sigma > 0, S_0 > 0,$

- **bank account** with value $B_t = e^{rt}$, where $r \geq 0$ is called *risk-free* interest rate.

It turned out that above mentioned stochastic differential equation meets the conditions of theorem of existence and uniqueness of the solution. Using Itô's Formula one can check that the solution has the following form

$$S_t = S_0 \exp\Big(\Big(\mu - \frac{\sigma^2}{2}\Big)t + \sigma W_t\Big).$$

Above process is called Geometric Brownian Motion (GBM). Next, we talked about the economic interpretation of μ, σ parameters and proved theorem that shows the form of the martingale measure (i.e. measure $\mathbb{P}^*$ such that discounted price process ($S^* = \frac{S}{B}$) is $\mathbb{P}^*$–martingale) under which price process has the following form

$$S_t = S_0 \exp\Big(\Big(r - \frac{\sigma^2}{2}\Big)t + \sigma W_t^*\Big),$$

where W^* is Brownian Motion with respect to measure $\mathbb{P}^*$. Above mentioned theorem is crucial, due to the fact that if the price of the financial derivative has to be determined, one has to calculate dicounted expected value of payoff with respect to the martingale measure $\mathbb{P}^*$.

References

[1] Marek Musiela, Marek Rutkowski (2005), *Martingale Methods in Financial Modelling*, Springer.

[2] Daniel Revuz, Marc Yor (2005), *Continuous Martingales and Brownian Motions*, Springer.

1.3 23 November 2017 – Binomial Option Pricing – CRR method, Blockchain & Bitcoin

Speakers: Dominik Klemba, Jan Kościałkowski, Przemysław Ryś

The talk consisted of two following parts.

1. **Project presentation – option pricing in CRR model.** At the meeting there was a presentation of the project, which was written in package R, concerning pricing (in the CRR Model) of the following types of Options:

   - **European**, in which payoff function has the form $X = (S_T - K)^+$. In this case analytical formulas were used.
   - **Asian,** in which payoff function has the from $X = (\frac{1}{n}\sum_{i=1}^{n} S_i - K)^+$. In this case Monte Carlo methods were used. Presented algorithm of pricing consisted of multiple (e.g. 10^5) simulation of paths of price process and calculation of the average value of payoffs of each simulation. Due to the fact that we could find the distribution of all possible outcomes of payoff function (because of model specification), it would be possible to determine the exact price of the Asian Option.

   Parameters of the model were selected so that *in the limit* (in our case – for *thick* division of the time interval $[0, T]$, where T – maturity time) the price of the European Option in CRR model converged to the price the same Option in the coresponding Black–Scholes Model (*23 October 2017 – Binomial Tree Model*). Furthermore, at the meeting, we also analyzed the time of getting the estimated value (by means of Monte Carlo methods) of the Asian Options. It was quite easy to observe that if the division of the interval $[0, T]$ is thick, then getting the price of an Option is time-consuming. The conclusion about the Monte Carlo methods was such that these methods are relatively easy to implement and they can give the preliminary insights of the values, but in case of more complicated problems they are too time–consuming.

   Code and more information about the project one can find in the GitHub repository – HERE. More about the application of Monte Carlo methods in Financial Engineering one can find in [1].

2. **Blockchain Technology & Bitcoin cryptocurrency.** The second part of the meeting was focused on the technological and economic aspects of Bitcoin cryptocurrency and Blockchain systems. The considerations' starting point was the strong rising trend of returns and volatility in Bitcoin quotations time series and the dynamic development of general cryptocurrency market. We presented and conducted the basic analysis of the BTC/PLN exchange rate since 2009, with particular emphasis on the data from the 2017 year. To illustrate the definite difference between cryptocurrencies and the 'standard' currencies, the basic statistics for BTC was compared with the respective ones for USD. The result was as the following:

- annualized average rate of return:

$$BTC : 670,99\%, USD : -16,02\%,$$

- annualized standard deviation of returns:

$$BTC : 83,28\%, USD : 9,22\%,$$

- Value at Risk for currency investment (settled in PLN) with one month (21 trading days) horizon ($VaR_{0.95}$) as a percentage of capital involved:

$$BTC : 38.17\%, USD : 4.54\%,$$

Value at Risk was calculated by the two methods - multivariate historical simulation method and analytical (variance-covariance) method with the confidence level 0.95%. The value presented above is a simple average of two considered methods results. The Bitcoin value increases in a very dynamic way but the volatility and risk are quite high too. However, the risk seems to be low compared to the returns - calculated IR coefficient is very high - equal to 8.057. The risk is really significant, but its high value is compensated by the enormously high returns.

In the further part, the Blockchain principle of operation has been presented. The main considerations concern the transaction authorizing process and the knowledge of *public-key cryptography*, crucial for understanding that. The speakers discuss the system security and the ways Bitcoin was attacked in the past. Another important issue concerned the essence of Bitcoin and contained arguments for classifying this as the currency or as the investment asset. The main reasons why Bitcoin cannot be used as the general currency in a country economy were presented. The limited liquidity, delays in transactions and the huge prices volatility make Bitcoin impossible to be used as the basic currency for trading, daily shopping and accumulating the capital by everyman. The main advantages of using Bitcoins is difficulty transaction tracking and safety from the state interference in people accounts.

The last part of the presentation concern the economic aspects of Blockchain. The low entry costs for system providers (miners) and the construction of automatic rewarding system guarantee size of network enough to provide safe transactions and always keep the fees at almost the lowest possible value. The fees are sometimes charged indirectly for instance by new coins emission and using it to pay off the miners.

References

[1] Paul Glasserman (2003), *Monte Carlo Methods in Financial Engineering*, Springer.

[2] Satoshi Nakamoto (2008), *Bitcoin: A Peer-to-Peer Electronic Cash System*, *https://bitcoin.org/bitcoin.pdf*.

[3] J. H. Witte (2016), *The Blockchain: A Gentle Four Page Introduction*, *https://arxiv.org/abs/1612.06244*.

[4] Jan Kościałkowski, *CRR Model – European and Asian Option pricing*, *https://github.com/KNFO-MIMUW/CRR-model-European-Asian-options*.

1.4 07 December 2017 – Poisson Process, group project – calculating VaR (Value at Risk) of the shares portfolio

Speakers: Jan Kościałkowski, Przemysław Ryś

The talk consisted of two following parts.

1. **Poisson Process – definition and properties.** The purpose of this short lecture was to define the Poisson process and elaborate on its properties. It began with a definition of a broader class of Levy processes, which means starting from 0, with independent and stationary increments and continuous in probability, and proceeded to show that by adding several assumptions one can get Wiener process or the very Poisson process. After defining the main object of interest, basic properties were displayed and proven:

   (a) trajectories are $\mathbb{P}$ – as. non-decreasing,

   (b) trajectories $\mathbb{P}$ – as. have values in $\mathbb{Z}_+$,

   (c) jumps are equal to 1 with probability 1,

   (d) $\lim_{t\to\infty} \frac{N_t}{t} = \lambda$,

   (e) times between jumps are independent random variables with exponential distribution.

   The next step was to justify the construction of the Poisson process as a supremum of a number of exponential variables fitting in between time 0 and t. A result stating that a process constructed this way satisfies a stronger version of increment independence was given without proof. Ultimately, compound Poisson process was defined, along with showing its application in insurance and how its trajectories depend on and vary from the Poisson point process. The last point was bringing up the Levy - Itô decomposition (characteristic triplet), which allows decomposing any Levy process into three independent processes of which one is a Wiener process with drift, the second a compound Poisson process and the third a pure jump martingale.

2. **The group project - estimating Value at Risk for the stock portfolio - multivariate historical simulation method.**

   The second part of the meeting was focused on the historical simulation method of estimating Value at Risk for the portfolio composed of stocks. The considered method is based on the time series of risk factors, calculated from the market data (e.g. returns of the stocks). The main part is calculating the vector of the risk factors predictions for the following period, after taking into account the current trend of the volatility and returns of the particular components of the portfolio. The forecasts of the risk factors are further used to generate the loss simulations for the considered portfolio in a specified period. Every single loss simulation comes from different, single observation of the risk factors in the historical period. The last step is Value at Risk calculation for the empirical distribution of losses.

   The considered method does not require any theoretical assumptions about the distribution of the risk factors but takes an assumption that the vectors of risk factors in different periods are independent and comes from the same (perhaps multidimensional) distribution. This assumption implies that the empirical distribution of the losses converges to the theoretical distribution, when the number of observations, using to estimate it converges to infinity, what means that the method for a big number of observations should give Value at Risk estimations very close to the real value. The method is based on the vectors of the different risk factors, appearing in the same time, so it contains some information about the statistical interdependence of these and does not require costly calculations of the variance-covariance matrix.

The theory was completed with the presentation of the script calculating Value at Risk for the exemplary stock portfolio in R and the visualization of its results. The historical simulation method's results have been compared with the ones from the variance-covariance method. The presented script will be developed as the group project, which the main purpose is to design and run the interactive web application. The target application will calculate Value at Risk for the portfolio predefined by a user by one of the analytics and simulations methods. The application will be helpful not only to calculate Value at Risk but also to compare the effects of the different methods and different portfolios.

References

[1] Ioannis Karatzas, Steven Shreve (1998), *Brownian Motion and Stochastic Calculus*, Springer.

1.5 21 December 2017 – Heston Stochastic Volatility Model

Speaker: Michał Jaworski

The aim of the meeting was to present the theory and implementation (code written in Python) of calibration of the Heston model based on the [1].

We started the talk with the recollection of the Black–Scholes model (*4 November 2017 – Black–Scholes Model*), in which we assume that at the market is one risky asset with the price process is given by the following SDE

$$dS_t = rS_t dt + \sigma S_t dW_t,$$

where W is Brownian Motion under the martingale measure $\mathbb{P}^*$, r is *risk-free* interest rate and σ is the volatility parameter. The advantage of the Black–Scholes model is that it is possible to derive the explicit formulas for the price of the European Options, which depends on the following parameters: K – strike price, T – time to maturity, S_0 – current spot price of the asset, r, σ. First two parameters are determined in the contract, next two are taken from the market. The only parameter that has to be estimated is σ. We have briefly discussed two types of volatility:

- **implied** – provided that at the market we can observe the price of the specific European The option then we can choose the volatility in our model such that the European Option with the same underlying, maturity date and strike price in the Black–Scholes model has the same price (calculated by means of Black–Scholes formulas) as this one observed at the market. We call it *implied volatility* because it is implied by the market.
- **historical** – which is based on the historical changes of stock prices. One can estimate the volatility of the stock prices by means of moving average.

Black–Scholes model is the most popular continuous model of the market due to its simplicity. As it was mentioned before, volatility is the only parameter to be determined. One of the most important features of this model is volatility smile, which is a phenomenon observed at the markets. Volatility smile can be displayed as the relation of the implied volatility and strike prices of the European Option (with fixed maturity). Implied volatility should be constant, because strike price should not influence the volatility of stock prices of underlying but in many market cases the more the Option is out–of–the–money or in–the–money, the higher the implied volatility is. What is more, it is observed that volatillity of underlying tends to change in time, if new information appears at the market.

In the view of mentioned imperfections of Black–Scholes model we discussed the Heston stochastic volatility model, in which stock price process is given by

$$dS_t = rS_t dt + \sqrt{v_t} S_t dW_t^1,$$

where

$$dv_t = \kappa(\bar{v} - v_t)dt + \sigma\sqrt{v_t} dW_t^2$$

and $d\langle W_t^1, W_t^2\rangle = \rho dt$. As we can see the variance is CIR process. In case of Heston model there are five unknown parameters: v_0 – initial variance, $\bar{v}$ – long-term variance, ρ – correlation of W^1 and W^2, κ – mean-reversion rate, σ – volatility of volatility. The main topic of the talk was calibration of Heston model, which consists of determining the parameters in such way that model best fit the reality. Based on the paper [1] we formulated problem of calibration as follows.

Assume that we observe N European call options with prices denoted by $C_{mkt}(K_i, T_i)$ for $i = 1, \ldots, N$. Let $C(\Theta, K_i, T_i)$ be the price of the European call option in the Heston model, where $\Theta = [v_0, \bar{v}, \rho, \kappa, \sigma]$. We formulated problem of calibration as

$$\min_{\Theta} \frac{1}{2} r^\top(\Theta) r(\Theta), \tag{2}$$

where $r(\Theta) = [C_{mkt}(K_1, T_1) - C(\Theta, K_1, T_1), \ldots, C_{mkt}(K_N, T_N) - C(\Theta, K_N, T_N)]$. It is non–linear least squares problem and in order to solve this, it is convenient to use Levenberg-Marquardt (LM) method, which interpolates between Gauss-Newton algorithm and gradient descent method.

So as to perform LM algorithm, in case of our problem, it is necessary to be able to effectively compute $C(\Theta, K, T)$ and $\nabla C(\Theta, K, T)$. Due to complexity of the model it is not possible to derive explicit formulas for price of the European Option and gradient with respect to Θ. Authors in the paper [1] proposed the forms of the $C(\Theta, K, T)$ and $\nabla C(\Theta, K, T)$, which consist of the integrals of functions, in which the characteristic function of the logarithm of the stock price is involved.[2] Therefore, the computation of price and gradient comes down to numerical approximation of the integrals.

In order to compute above mentioned integrals one can apply Gauss–Legendre quadrature, in which integral of function f over the interval $[a, b]$ is approximated as follows

$$\int_a^b f(x)dx \approx \frac{b-a}{2} \sum_{i=1}^{n} \omega_i f\Big(\frac{b-a}{2} x_i + \frac{b+a}{2}\Big),$$

where $(x_i)_{i=1}^n$ is the sequence of roots of n–th Legendre polynomial $P_n(x)$ and for all $i = 1, \ldots, n$ $\omega_i = \frac{2}{(1-x_i^2)[P_n'(x_i)]^2}$.

Summing up, if we observe market data i.e. sequence of market prices $C_{mkt}(K_i, T_i)$ then

[2] equations (6) and (22) in [1]

for the *reasonable*[3] choice of the initial vector of parameters Θ_0 we can perform LM algorithm in order to solve the non–linear least squares problem – (2). The outcome of the calibration is vector of parameters Θ^*. One can use Θ^*, discretization of SDEs and Monte Carlo methods to price more sophisticated financial derivatives (e.g. Asian Options) in calibrated Heston model.

Results and implementation of calibration one can find in GitHub repository – HERE.

References

[1] Yiran Cui, Sebastian del Baño Rollin, Guido Germano (2017), *Full and fast calibration of the Heston stochastic volatility model*, European Journal of Operational Research 263 (2017) 625–638, *http://www0.cs.ucl.ac.uk/staff/G.Germano/papers/EurJOperRes_263_625_2017.pdf.*

[2] *Heston model – calibration and pricing*, *https://github.com/KNFO-MIMUW/Heston_model.*

[3] Marek Musiela, Marek Rutkowski (2005), *Martingale Methods in Financial Modelling*, Springer.

1.6 11 January 2018 – Construction of forward and yield curves

Speakers: Przemysław Dycha, Michał Jaworski

Description coming soon.

References

[1] Patrick S. Hagan, Graeme West, *Interpolation Methods for Curve Construction*, Applied Matematical Finance, Vol. 13, No. 2, 89–129, June 2006.

[2] *Hagan_West_interpolation*, *https://github.com/KNFO-MIMUW/Hagan_West_interpolation.*

1.7 08 February 2018 – Credit Suisse at MIMUW

Speaker: Nicolea Mera, PhD

The meeting consisted of a brief presentation of Credit Suisse Risk Department in Warsaw and one-hour lecture – *Market risk models and modeling challenges under the Fundamental Review of the Trading Book (FRTB).* Below you can find an abstract of the lecture.

An overview of changes to market risk modelling requirements and practice brought by regulation changes upcoming under the Fundamental Review of Trading Book will be presented. The ways in which FRTB proposes to address gaps in the existing Basel market risk regulation and capitalization framework will be discussed including how that will impact market risk model development, testing, evaluation and validation. Modelling and implementation challenges that will be faced by risk modelling practices globally will be discussed.

[3] one can find ranges for the parameters in Table 5 in [1]

1.8 27 February 2018 – Two Sigma Info Session

Speakers: Two Sigma representatives

Computational Finance Students Research Group alongside with QuantFin Foundation was responsible for organizing Two Sigma information session for students at MIMUW. Quant hedge fund Two Sigma is one of the world's leading systematic investment managers. At the event, there was a brief presentation of the company, which was conducted by managers from the Quantitative Research team and Q&A session.

1.9 15 March 2018 – K–Hartman–Watson distribution

Speaker: Maciej Wiśniewolski, PhD

The new general results for Brownian motion functionals are obtained by introducing the so-called K–transforms of the special function u. The transforms lead to the new distributions called here K-Hartman-Watson distributions. They are counterparts to the classical Hartman-Watson distributions, but with respect to the state-space variable. It is shown that the distribution of geometric Brownian motion e^{B_t} and its additive functional A_t are described by the composition of General Inverse Gaussian (GIG) distribution with K-Hartman-Watson distribution. As a consequence two distributional dependencies are presented: the first between the distribution of GIG distribution on $\mathbb{R}_+$ and the distribution of functional A_t, and the second one, between the distribution of GIG on a hyperplane and the distribution of the vector (A_t, e^{B_t}).

1.10 22 March 2018 – LIBOR market model

Speaker: Rafał Grądziel

Description coming soon.

1.11 12 April 2018 – Introduction to High-Frequency Trading & automated trading strategies

Speaker: Przemysław Ryś

At the meeting, we have discussed the main differences between intra-day and longer-term trading as well as high-frequency market data properties, which are crucial for trading purposes. The main characteristics of high-frequency data are periodic patterns occurrence, discrete price increases and unequally spaced quotations. Trading on the high-frequency data is far more complicated than operating on the traditional markets. Therefore, it requires the statistical and econometric tools to indicate time series properties and to make a high-performance algorithm for automated decision making and execution tasks in an efficient way. Afterward, the idea of the momentum and mean-reverting (contrarian) approach was explained and it became the starting point for the discussion about constructing trading strategies. The different approaches based on the technical analysis indicators were introduced. These strategies generate signals according to the information given by the calculated moving averages, moving risk measures and oscillators. The decision logic of strategies such as the two moving averages crossover or standard deviation based volatility breakout was explained and presented on the basis of exemplary data. Next, there was a discussion about the advantages and disadvantages of every specific method. The last part was focused on the practical aspects of selecting and testing strategies, with including the introduction to train cross-validation method and typical learning problems – data-snoop bias (overfitting) and selecting training and testing sample.

1.12 19 April 2018 – Implied Volatility, Volatility Smile

Speaker: Maciej Wiśniewolski, PhD

Description coming soon.

1.13 17 May 2018 – Goldman Sachs - Quant Workshop

Speakers: Michał Jaworski, Bartłomiej Polaczyk, quantitative analysts from Goldman Sachs

The meeting was organized in cooperation with Goldman Sachs Quantitative team and consisted of two parts. It began with a lecture given by KNFO representatives – M. Jaworski and B. Polaczyk. The lecture started with a description of Black-Scholes model formulation, solution and its extensions, then it moved to the subject of volatility surface modelling and finally an overview of Monte Carlo methods for finance was given with control variates and importance sampling methods described thoroughly.

After the lecture followed the workshop. The students were divided into small groups and were given programming tasks consisting in implementing various stochastic models and variance reduction methods presented during the lecture. During the worksop, Goldman Sachs and KNFO representatives were mentoring the groups.

References

[1] John C. Hull (2014), *Options, Futures, and Other Derivatives*, Pearson 9th edition.

[2] Marek Musiela, Marek Rutkowski (2005), *Martingale Methods in Financial Modelling*, Springer–Verlag, 2005.

[3] Paul Glasserman (2003), *Monte Carlo Methods in Financial Engineering*, Springer–Verlag.

1.14 24 May 2018 – Pricing of Catastrophe derivatives

Speaker: Jan Kościałkowski

Catastrophic derivatives are present in the market since the 1990s. They help insurers and reinsurers to hedge against large and difficult-to-predict claims. The lecture began with a concise introduction to random measure theory which developed into a definition of the Cox process as an integer-valued random measure. Basing on it we constructed a modified Black-Scholes market which served to obtain a practically applicable formula for the price of a CatEPut option, a typical example of a catastrophic instrument emitted by Aon. The goal was to expose links between stochastic theory of finance and general insurance mathematics. Such crossovers highlight the need to have a firm grasp of basic applied mathematics concepts to be able to face real-life applications of stochastic processes.

References

[1] Daryl J. Daley, David Vere-Jones, *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods, Second Edition*, Springer, Nowy Jork 2003.

[2] Takahiko Fujita, Naoyuki Ishimura, Daichi Tanaka, *An Arbitrage Approach to the Pricing of Catastrophe Options Involving the Cox Process*, Hitotsubashi Journal of Economics 49, 67-74, 2008.

[3] John F.C. Kingman, *Poisson Processes*, Clarendon Press, Oxford 1993.

1.15 29 May 2018 – QFRG & Labyrinth HF at MIMUW

Speakers: Krzysztof Kość; Paweł Sakowski, Ph.D.; Robert Ślepaczuk, PhD

The meeting with Quantitative Research Group WNE UW and Labyrinth HF Project took place on 29th of May 2018 as part of co-operation between the Computational Finance Students Research Group and QuantFin Foundation. Speakers presented a new start-up project of Labyrinth HF and shared results of the newest research on algorithmic strategies on the cryptocurrency market, focused on contrarian and momentum effects. Additionally, team present out-of-sample performance of strategies, based on that research.

Academic paper abstract: We report the results of investigation of the momentum and contrarian effects on cryptocurrency markets. The investigated investment strategies involve 100 (amongst over 1200 present as of date Nov 2017) cryptocurrencies with the largest market cap and average 14-day daily volume exceeding a given threshold value. Investment portfolios are constructed using different assumptions regarding the portfolio reallocation period, width of the ranking window, the number of cryptocurrencies in the portfolio, and the percent transaction costs. The performance is benchmarked against: (1) equally weighted and (2) market-cap weighted investments in all of the ranked assets, as well as against the buy and hold strategies based on (3) S&P500 index, and (4) BTCUSD price. Our results show a clear and significant dominance of the short-term contrarian effect over both momentum effect and the benchmark portfolios. The information ratio coefficient for the contrarian strategies often exceeds two-digit values depending on the assumed reallocation period and the width of the ranking window. Additionally, we observe a very significant diversification potential for all cryptocurrency portfolios with relation to the S&P500 index.

References

[1] Krzysztof Kość, Paweł Sakowski, Robert Ślepaczuk, *Momentum and contrarian effects on the cryptocurrency market*, `https://www.wne.uw.edu.pl/files/2615/2405/6484/WNE_WP268.pdf`

2 Academic year 2018/19

2.1 08 November 2018 – Description of projects

Below one can find description of the projects presented at the meeting. We assume that participants of our Group will work on the implementation of projects. The results will be presented at the Group's profile at GitHub: `https://github.com/KNFO-MIMUW`.

2.1.1 Time Series prediction

The aim of the project is to implement and test different methods of the stock prices time series prediction. We assume that each participant of the project will implement different method. Built models should predict stock price of the specific company one day ahead - models will be tested and compared with each other on the real data.

Plan of the project

1. We select specific company from S&P 500 and set of time series, which may have influence on the prices of selected company (for instance - if you decide to predict the prices of Exxon Mobil Corp, then we probably should look at the prices of oil). Based on the selected set of time series we conduct basic analysis of the time series.

2. We conduct basic analysis of the time series - we check trend, seasonality, heteroscedasticity, stationarity, etc. What is more, one can compute statistics and plot the results.

3. We build predictive models by means of different *classic* time series prediction and machine learning techniques. Methods to consider:

   - SVM - support vector machine
   - LSTM - Long Short-Term Memory
   - ARIMA, ARMA, SARIMA, etc
   - GARCH
   - Exponential smoothing, Holt-Winters
   - Kalman filtering
   - MLP - Multilayer perceptron
   - CNN - Convolutional Neural Network
   - other, which can be found in the publications, books, etc

Project will be written in Python. Repository of the project can be found here:
`https://github.com/KNFO-MIMUW/Stock_Prices_Prediction`.

2.1.2 Sentiment Analysis

The aim of the project is to compare the space of stock prices time series of companies from S&P 500 with the space of sentiment time series observed on Twitter towards mentioned companies, where by sentiment time series we understand the fraction of *positive* posts about the specific company from S&P 500. Hypothesis is such that, if we perform clustering on both spaces, we should observe the same results, namely clusters of stock prices time series should contain the same companies as clusters of sentiment time series. It may imply that the structures of dependencies of dynamic of prices and sentiment are similar. The main challenge is to determine, what the *similarity* of two time series is. We are going to test different measures and clustering techniques of time series. In order to make time series more comparable, we are going to work on the *returns* from stock prices and sentiment. Project will be written in Python.

Plan of the project

1. We construct yearly (2017) time series of sentiment towards companies based on the posts on Twitter. We will use package `TextBlob` in order to determine the character of the selected statement towards specific company (whether it is *positive* or *negative*). To have an easy access to posts from Twitter one can use package `twitterscrapper`.

2. Clustering of the stock prices time series and sentiment time series. In order to do this step, several different approaches are possible, for example:

   - one can try to compute the *distance* between time series by means of simple correlation or, for instance, DTW (dynamic time warping). Once we compute the *distance*, we are able to use classical clustering techniques and check which time series are close to each other. We can also build a graph, in which vertices will be represented as companies and edges as the distance between two specific companies and than use centrality measures from the network theory to identify the most important vertices (companies).
   - the other way of solving this exercise is to look at the time series of the specific company as the observation in 252-dimensional (we are looking at working days) space and try to implement dimensional reduction techniques such as SOM (self-organising map) - and then based on the reduced dimension use clustering methods.

3. Visualization of results - presentation of graphs, maps in elegant (possibly interactive) way.

Repository of the project can be found here:
`https://github.com/KNFO-MIMUW/Sentiment_analysis`

2.1.3 Pricing of derivatives

The aim of the project is to build an application, which allows user to price and determine the sensitivities of different financial derivatives. We plan to use different stochastic, local volatility models and numerical methods to solve our problem. Project will be written in statistical package `R` and in order to build an application we will use package `Shiny`.

Plan of the project

1. We build an application framework in `Shiny`.

2. We select specific underlying (or set of underlyings) - it is important for us to have access to volatility surface, because implemented pricing models will be calibrated to volatility surface.

3. We implement different models that will be calibrated to data choosen in the previous step. Models to consider:

   - Heston model
   - CEV (constant elasticity of variance) model
   - Bates model
   - other models from the class of local and stochastic volatility models.

In majority of cases we cannot represent price or sensitivities by means of explicit formulas, hence there is a need to use numerical methods to PDE's, SDE's and Monte Carlo methods (we will try to implement not only basic MC but also advanced variance reduction methods).

4. Visualize results in the application

Repository of the project can be found here:
`https://github.com/KNFO-MIMUW/Derivative_pricing`.

2.2 20 November 2018 – QuantDay conference

"QuantDay" was a conference organised by Computational Finance Student Research Group MIMUW and QuantFin Foundation. The aim of the conference was to show application of mathematical methods such as stochastic processes, numerical methods, machine learning into finance. At the event there were six lectures conducted by academic staff and business professionals. There were around 150 participants of the conference. Below one can find descriptions of talks. Photos from the event can be founded at the website:
`knfo.mimuw.edu.pl/index.php/quant-day/`.

"Why you should NOT invest in BTC mining?"

Speaker: Grzegorz Zakrzewski, member of Quantitative Finance Research Group

The presentation summarized main findings of the paper "Why you should not invest in mining endeavour? The efficiency of BTC mining under current market conditions." The detailed profitability model of BTC mining was presented together with extensive sensitivity analysis. The presentation outlined assumptions of the mining industry and the consequences for the BTC equilibrium state. The efficiency of BTC mining under current market conditions was presented. After thorough analysis of initial assumptions concerning:

- the price of mining machine and its effective amortization period,
- difficulty and hashrate of BTC network,
- BTC transaction fees, and
- energy costs,

it was proved that currently BTC mining was far from positive profitability, except for some rare cases.

"Challenges of stochastic modelling of option values and warranties built in the insurance products"

Speaker: Marek Wielgosz, Actuarial Manager in Aviva ASEC

Many of currently sold insurance products are connected to quotations of financial instruments. The talk covered the process of valuation of liabilities, which among others consists of the construction of mathematical models, calibration of models to market data and scenarios simulations. What is more, at the lecture there was a presentation of antithetic variates method - one of the variance reduction techniques of Monte Carlo methods. Theory of variance reduction is important from the business view, due to the fact that it contributes to shortening the time of reporting and improving the quality of results obtained by means of simulation.

"Application of Machine Learning in banking industry"

Speaker: Robert Małysz, Associate Partner at EY, Data & Analytics leader

The talk presented the overview of Machine Learning methods, which are used in order to solve problems in the banking industry. Two following areas of the banking industry were considered:

- Business Area, in which there is a freedom to choose any methods. Examples are as follows: clustering algorithms allow segmentation of clients, which could have a positive influence on the effectiveness of marketing campaign; another example is using NLP (Natural Language Processing), which allows accelerating the document processing, which may contribute to making faster business decisions.
- Risk Area, in which due to Regulators only selected models are allowed. One can use ML methods, for example, to achieve more accurate results in risk models, which may result in making better decisions to grant a loan.

"How can we test algorithmic trading strategies?"

Speaker: Robert Ślepaczuk, co-founder Labyrinth HF

Presentation contained the basic principles of testing of algorithmic investment strategies, especially from the point of view of traps that may appear during this process. Emphasis was placed on the detailed elements of the testing process, focusing on both the process of selecting of appropriate investment concepts, databases, real testing environment replication, as well as optimization techniques, the process of building closed automatic investment applications and final real-time tests. An additional advantage was the practical discussion of the investment strategy included in the Short Volatility group.

"Introduction to CVA and key modelling concepts"

Speaker: Rafał Muchorski, associate at Goldman Sachs, Credit Quantitative Analysis

The talk presented an introduction to CVA (Credit Valuation Adjustment) and key modeling concepts. Highlighting the role of credit risk in the 2008 global financial crisis, presenting a few real-life examples of the issues that arise due to counterparty creditworthiness and its consequences. Present the concept of CVA as default risk hedge, starting from the basic idea and building up to some math formulas reflecting it. Provide a simple example of calculating CVA for a standard IRS swap; pricing first with no default assumption and then including default risk and demonstrating its impact on pricing. Then presenting some counterparty risk mitigants.

"Factor Modelling: from classic approach to machine learning"

Speakers: Rafał Grądziel & Karol Partyka, analysts at Goldman Sachs, Model Risk Management

The talk covered the industry-wide applications of factor models. Dimensionality reduction is a useful tool in portfolio construction, risk monitoring, and portfolio management. Starting with the application of factor decomposition to mean-variance optimization, we discussed the merits of statistical and fundamental factor models and how they can be used to compute Value-at-Risk. The discussion also covered pitfalls that can arise in practice, as well as the potential of machine learning to improve the factor modeling.

3 Proposed talks in academic year 2018/19

- Calculation and interpretation of the Greeks by means of different methods (e.g. Monte Carlo methods, Malliavin Calculus), Hedging
- Financial Derivatives on cryptocurrency markets
- Bond Markets Models
- Pricing of derivatives in models with jumps and sentimental variable
- Loss Distribution Approach (LDA) for operational risk
- Multifractal analysis of market data
- Multicurve Models
- Copulas
- Calibration of Short Rate Models by means of Neural Networks
- Probability of Default in the context of Credit Risk, where intensity is given by the specific Stochastic Differential Equation
- Discussion over mistakes in thought process based on the book – Thinking, Fast and Slow – Daniel Kahneman