National Research University Higher School of Economics

Artur R. Sarkisov

Constructing of an Optimal Portfolio on the Stock Market Using a Nonparametric Method

PhD Dissertation Summary
for the purpose of obtaining
Philosophy Doctor in Economics HSE

Academic Supervisor:
Elena A. Bujanova,
PhD in Physico-Mathematical Science

JEL: G10, G11, G17

Moscow – 2018

Problem description

Forming of an investment portfolio is one of the main subjects of modern finance theory. In economic literature there are a lot of researches devoted to forming optimal investment portfolio using technical and fundamental analysis of characteristics of issuers by parametric methods. However, often the application of parametric methods of analysis leads to ineffectiveness and bias of final estimates. This is due to the fact that in most cases the analyzed data series were either too short or had an unstable internal structure. Thus, at the moment, an important issue is the development of algorithms for handling with short time series and series with a large number of outliers. Due to these problems was caused the transition from methods based on distribution characteristics of variables to the nonparametric methods, which don't impose such requirements. Therefore, the relevance of dissertation research is caused by the need to improve approaches of the forming of investment portfolios on the basis of analysis of factors of technical and fundamental analysis.

Contribution

1. For the first time were developed algorithms of forming an investment portfolio on Russian stock market (MICEX) using three nonparametric methods: CART, ANN and kernel regression.
2. The effectiveness of nonparametric methods has been proved by comparing of returns of portfolio formed using nonparametric methods, market return and return of portfolio formed using parametric method. The analysis was conducted on time period of

stable market (jan.2016 – dec.2016) and on time of crisis (jan. 2008 – dec.2008)

3. The unique set of factors determining dynamic of MICEX index and ,hence, its return was found. This set of factors proved hypothesis of speculative nature of Russian stock market.
4. For the first time was conducted comparative analysis of effectiveness nonparametric method on Russian stock market

Brief literature review

In the area of constructing an optimal portfolio using nonparametric methods the following two areas of research should be highlighted. First: at early stages of development of this direction, researchers tried to show the effectiveness of the nonparametric methods, using only standard indicators of fundamental analysis. Second: in later studies attempts were made to obtain the optimal combination of factors for both technical and fundamental analysis, which would be most suitable for a specific market. Below is a list of the main works on the use of nonparametric methods in finance.

Breiman (1987) formulated the main advantages of nonparametric methods:

- Don't require preselection of parameters; there is no problem of misspecification
- Insensitive to monotonic transformation of independent variables
- Results are robust to sample outliers

Andriyashin, Hardly and Timofeev (2008), Breiman (1984), Chavarnakul and Enke (2009) described the theoretical basis for using the

method of classification trees to construct an optimal portfolio in the stock markets of developed countries.

The method of kernel regression was used Shiraishi & Taniguchi (2007) to estimate the variance of stock returns on the Tokyo Stock Exchange. The authors showed that by using of this nonparametric method it is possible to effectively predict the behavior of stocks and build an optimal portfolio. At the same time, the authors note that the effectiveness of a nonparametric method can be different depending on the data structure.

The so-called I (d) processes (Bierens, 1997) was used in research by Aue & Ming (2014), in which the authors built a model of optimal diversification in the selection of securities to the portfolio. The authors found that US stock markets (Dow Jones 30 and S & P 500) are pairwise cointegrated with the BRICS stock markets.

Krizanowski et al. (1992), Jang and Lai (1994), Freitas (2001), Alice and Wilson (2005), Vanstone (Vanstone et al., 2010) , Fernandez and Gomez (2012) used in their works the method of artificial neural networks to build the optimal portfolio. Authors of these researchers have developed a theoretical foundation for the portfolio selection using the ANN method and the main algorithms for determining the optimal size of a neural network. Not only projections of the cost of financial instruments were constructed using the ANN method, but also the effectiveness of this method was tested by fitting to existing data.

Agurar and West (2000), Avramov (2002), Barberis (2000), Brown (1976), Cremers (2002), Frost and Savarino (1986), Crawford and Wood (2016) used different variations of the Bayesian methods to construct effective trading strategies in stock markets. Moreover, in these studies, a

high degree of stability of the results was noted even though there were structural shifts in the data and short data series.

Objectives of the research

The main goal of the research is to evaluate effectiveness of nonparametric methods as an instrument for portfolio selection on Russian stock market and develop algorithms of portfolio formation and managing using nonparametric methods.

To achieve the goal the following objectives were formulated:

1. Order theoretical and empirical studies devoted to portfolio selection using parametric and nonparametric methods
2. Form a system of fundamental, technical and macroeconomic factors which will be used in portfolio selection by nonparametric methods
3. Develop an algorithm for constructing and rebalancing an investment portfolio using CART, ANN, and the method of kernel regression
4. Form investment portfolios using developed algorithms, and compare the obtained results with the market return, as well as with return of the portfolio formed using the parametric method. Make this check on the period of a stable market and during the crisis period
5. Identify the main determinants of the changes in the Russian stock index. Check selected factors on economic interpretation
6. Compare the effectiveness of nonparametric methods among themselves. Solve the problem of maximizing the investor's utility for different levels of risk aversion

7. Test the developed algorithms for false significance on the series of random walks

Methodology

In use of nonparametric methods an important step is selecting the correct input parameters (for example, for neural networks, the number of layers and neurons in each). Within the framework of this study, for each of the three methods (CART, ANN, and the method of kernel regression) optimal parameters of the architecture were selected, which allowed constructing effective algorithms for the formation of investment portfolios.

• For the classification tree method was used the SE Rule for the training sample to determine the optimal tree size. By solving the optimization problem, the size of the test sample was obtained at level 35% of the total number of observations.

• For a ANN method were formed artificial neural networks with 3 hidden layers of 9 neurons in each. The method of constructing "learning with the teacher" with the backward propagation of the error was used.

• For the method of kernel regression was used the Nadaraya-Watson evaluation function was used, which showed better results than the Priestley-Kao and Gasser-Müller methods

It should be noted that despite the existence of procedures that allow selecting architecture parameters for nonparametric methods, the main selection method is the trial and error method. The effectiveness of input set of parameters could be studied only empirically on test samples.

For two methods (classification trees and artificial neural networks method) to form optimal trading strategy it was necessary to assign a certain class C_i from a finite set of classes $\mathrm{C} = \{1,2,\dots,J\}$. For financial assets it is convenient to form classes depending on their return (R_t): {buy asst}, {sell asset}, { hold asset}. The threshold value $\bar{R}_i$ was used for each asset as a criterion for assigning one of the three classes. $\bar{R}_i$ was based on the average value of the asset's return over certain period in the past:

$$\begin{bmatrix} R_t > \bar{R}, y_t = \{\text{buy asset}\} \\ R_t < -\bar{R}, y_t = \{\text{sell asset}\} \\ -\bar{R} \le R_t \le \bar{R}, y_t = \{\text{hold asset}\}, \end{bmatrix}$$

where threshold values $-\bar{R}$ и $\bar{R}$ were selected basing on the average return of the target market over analyzed period.

Stocks of the companies, which are included in MICEX calculation base, were analyzed in this research, because they are the most liquid instruments on Russian stock market. Weekly return of the stock was analyzed. It was calculated by following formula:

$$R_t = \frac{E_t(P_{t+1}) - P_t}{P_t}, \qquad (1)$$

where: R_t −return of share in period t,

P_t − price of the share in period t,

P_{t+1} − price of the share in period$t + 1$,

$E_t(P_{t+1})$ − expected from period t price of share in period $t + 1$.

Thus, the primary formation of portfolios constructed using the methods of classification trees and artificial neural networks consists

of following steps (example of the algorithm used for the test period 2016):

1. Input matrix X_i was formed for each company included in the calculation base of the MICEX index and include data for the period from January 2008 to December 2015
2. Class vector was formed for each observation $C_i = \{(\text{buy sell}), (\text{sell asset}), (\text{hold asset})\}$ depending on the ratio of weekly returns and index returns for the period from January 2008 to December 2015
3. Shares of companies were selected, which at the time of December 2015 were assigned a class {buy asset} based on the built classification trees and neural networks
4. Determine the weight of shares with class {buy asset} in portfolio by solving the next optimization task to maximize the Sharpe ratio of the portfolio:

$$\begin{cases} \max \frac{r_p - r_f}{\sigma_p} \\ r_{p=} \sum_{i=1}^{n} w_i r_i \\ \sigma_p = \sqrt{\sum_{i=1}^{n} w_i{}^2 \sigma_i{}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_i w_j k_{ij} \sigma_i \sigma_j} \\ \sum_{i=1}^{n} w_i = 1 \\ w_i \geq 0, \end{cases} \quad (2)$$

where: $r_i -$ return of share i,

$\sigma_i -$ st. dev. of share i,

$w_i -$ weight of share i in portfolio.

Risk free rate was determined as monthly average return of government bonds with tenor 3-5 years.

After the initial formation of portfolios they were rebalanced every 3 months. Classifications trees and neural networks were again estimated for each stock taking into account new data. During rebalancing there was a sale of securities of companies that were in the portfolio and were assigned the class {sell asset} in the process of rebalancing. Funds received from the sale of securities were allocated among assets with the class {buy asset}.

Weights of shares during rebalancing were determined by solving:

$$\begin{cases} \max \frac{r_p - r_f}{\sigma_p} \\ r_{p=}\ RHS + \sum_{i=1}^{v} \overline{w_i} r_i \\ \sigma_p = \sqrt{\sum_{i=1}^{n} {w_i}^2 {\sigma_i}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_i w_j k_{ij} \sigma_i \sigma_j} \\ \sum_{i=1}^{n} \overline{w_i} = 1 - w_{rhs} \\ w_i \geq 0, \end{cases} \quad (3),$$

where RHS assigned to shares with {hold asset} status. This share was calculated as sum of producs of weights of certain shares and their return.

During forming of a portfolio based on the Bayesian method of kernel regression there was no need to assign a particular class to dependent variables based on return. Rebalancing was conducted by solving the problem (2), because Bayesian method doesn't assume to select assets to hold. In this case, each time the portfolio was formed based on the expected return, estimated by the Bayesian method.

The description of the variables used to form the input vector is below:

- **Macroeconomic:** GDP, GDP growth rate, inflation, inflation growth rate, bi-currnecy Basket value, trade balance, Brent's oil price
- **Fundamental variables:** current ratio, quick ratio, cash ratio, long-term debt-to-equity, total debt-to-equity, debt ratio, financial

leverage, net profit margin, return on equity, return on assets, return on common equity , changes of EPS to share price, Book Value per Share / Share Price, FCFE, changes of FCFE, company capitalization

- **Technical variables:** MA with different windows, Bollinger bands, momentum, Relative Strength Index (RSI), Moving Avereage Convergence/Divergence (MACD), Stochastic, trade volume and Bid – ask-spread

Using the optimal for each method input variable vector selected from the initial set of variables investment portfolios were constructed based on each of the three nonparametric methods. The constructed portfolios showed higher profitability than market and portfolios constructed using parametric method throughout the entire investment horizon of 2016. The constructed portfolios were well diversified: only within one rebalancing of kernel regression portfolio was formed, in which the weight of one share was 21%. Based on the results of this analysis the highest profitability were shown by portfolios constructed using the method of kernel regression. In addition, nonparametric methods showed their high efficiency on 2008 data forecasting. At the same time, the constructed portfolios showed a greater concentration than on 2016 data, but they were still diversified.

Based on the results of the analysis of the most significant factors in the construction of asset valuation based on nonparametric methods, it was stated that in the construction of portfolios by all three methods, the most significant variables for both periods (2016 and 2008) were: momentum, oil price and value bid-ask spread. The obtained results confirm the hypothesis that the Russian stock market is speculative. Investors analyze the movement of oil prices as a proxy for the general state of the Russian economy; try to invest in shares of companies that have already been

leaders in terms of past growth rates; ensure that these shares are liquid with a minimum Bid-ask spread.

In addition to solving the standard problem of maximizing income, which was formulated regardless to the preferences of the potential investor, the task of maximizing utility for different values of the parameter of marginal risk aversion was solved. Based on the results of solving the maximization problem, it was stated that use nonparametric methods allows to obtain profitability not lower than the market rate for investors with risk aversion factors up to **16.01** for the method of kernel regression, **15.9** for the method of artificial neural networks , **15.4 -** for method of classification trees. The standard risk aversion factor varies from 10-12 (Janecek, 2004), hence, use nonparametric methods allows to obtain a yield above the market for a large number of potential investors. This fact confirms the effectiveness of nonparametric methods as a tool for forming investment portfolios on the stock market.

At the last stage, the stability of the results was checked on random walk series. The following hypothesis was tested: main results of research were random and it is possible to obtain high profitability on random walk series using nonparametric methods, which means that results didn't based on economic relationships. For testing this hypothesis random walk data series were created and algorithms created in this research were testem on this data series. 50 different random walk time series of stock prices were created, "market index" were recalculated basing on this series and stock perfomances. According to the results of testing, no cases were revealed when use of non-parametric methods allows to obtain yield higher than the modeled "market". This fact indicates that the results of research are not accidental and are based on economic interrelations between the variables.

Main findings

1. The necessity of using nonparametric methods as a tool for constructing an investment portfolio was proved. The efficiency of this type of methods was demonstrated by comparing the profitability of portfolios formed using nonparametric methods with the return of the market portfolio, and also the portfolio constructed using the parametric method.
2. It was shown that the main determinants of the Russian stock market are Momentum, Bid-Ask Spread, and the price of oil (Brent). This result is stable, because it was confirmed in the analysis of all three nonparametric methods over both periods: 2008 and 2016
3. It was proved that the portfolio formed using the method of kernel regression had showed the highest return over time horizons January 2016 - December 2016 and January 2008-December 2008. The portfolio formed using the method of kernel regression also showed a strictly higher yield than portfolios constructed using other methods during solving the utility maximization problem for the investors with different risk aversion factors.

4. It was shown that during testing of the algorithms on random walk series no positive results of the return was obtained , which proved statistical significance of results.

List of author's original article

1. Sarksiov A.R., Golodova J.G. Constructing of Investment Portfolio of Commercial Bank: Accounting of the Indicators of the Issuers

on the Stock Market// Journal of Finance and Credit. 2012. № 35. pages 24-29. - 0.4 quire (personal author's contribution – 0.3 quire)

2. Sarksiov A.R., Bujanova E.A. Constructing of Optimal Portfolio on Russian Stock Market Using Nonparametric Method – Classification and Regression Tree // Journal of Corporate Finance Research. 2016. № 1. pages 46-58. – 1 quire (personal author's contribution – 0.6 quire)
3. Sarksiov A.R., Bujanova E.A. Constructing of Optimal Portfolio on Russian Stock Market Using Nonparametric Method – Artificial Neural Network // Journal of Corporate Finance Research. 2017. № 3. pages 100-110. – 1 quire (personal author's contribution – 0.6 quire)