

On autoregressive moving-average models as a tool of virtual stock-exchange: experimental investigation

Jonas Mockus, Joana Katina, Igor Katin

Vilnius University, Institute of Mathematics and Informatics

Akademijos 4, LT-08663 Vilnius

E-mail: jmockus@gmail.com, joana@getweb.lt, igor@getweb.lt

Abstract. The objective of this work is to investigate experimentally the well-known autoregressive models as simplest algorithms simulating prediction processes of the stockholders using the historical stock rates only. The “virtual” stock exchange which applies these algorithms can help in testing various assumptions of investor behavior. To represent users that prefer linear utility functions, the autoregressive moving-average model ($ARMA-ABS(p, q)$), minimizing the absolute values of prediction errors is regarded, in addition to the traditional $ARMA(p, q)$ model which minimize the least square errors. The results of two hundred actual financial time series and a hundred of virtual ones are discussed in short.

Keywords: forecasting, autoregressive models, stock exchange, time series.

Introduction

The field of time series analysis and forecasting methods has significantly changed in the recent time due to the influence of new knowledge in non-linear dynamics. New methods such as artificial neural networks replaced traditional approaches which usually were appropriate for linear models only.

Nevertheless, there are still applications where estimations of linear processes, such as autoregressive and moving average models, are desirable. This includes the simulation of the stock exchange by some simple algorithms imitating actions of major stockholders based on their predictions of future stock rates and expected profits. In this setup we need prediction algorithms which are understandable for ordinary stockholders and not too difficult for calculations. Algorithms of autoregressive models satisfy these conditions. The common autoregressive model is $AR(p)$. This model reflects the opinions of risk-prone users because the squared prediction errors are minimized. To represent risk-neutral users we regard $AR-ABS(p)$ where absolute errors are minimized.

Applying these models we assume that each player predicts stock prices by $AR(p)$ or $AR-ABS(p)$ of order p . The scale parameters a of the model $AR(p)$ are estimated using the standard least squares algorithm for different p . The optimization of these parameters in $AR-ABS(p)$ is a piece-wise linear programming problem [12, 9] which can be formulated as the Linear Programming (LP) one.

Actual stockholders use their own ways of predicting. We regard the $AR(p)$ and $AR-ABS(p)$ models as the simplest initial approximations of their prediction pro-

cesses. The next approximations are $ARMA(p, q)$ and $ARMA-ABS(p, q)$ ¹ which corrects the past errors. The $AR-ABS(p)$ model was introduced in [12]. The $ARMA-ABS(p, q)$ model is a new element of this paper To optimize parameters b we have to solve multi-modal problem using global optimization methods [8, 10, 11, 3, 4]. Traditional methods [7, 2] solved similar problems using specific statistical techniques including sequential local optimization with no references to the inherent multimodality of errors as a function of MA parameters b .

1 Simulating stockholders predictions

The formal description of the $ARMA(p, q)$ models is well known [2, 7]. However, we shall describe these models in the form more convenient for discussions of this paper and supplement them by the new $ARMA-ABS(p, q)$ models representing linear utility functions.

Assume that a stockholder predicts next-day stock prices $Z(t+1)$ using the $AR(p)$ model [2]. Professional investors are trying to obtain additional information about the fundamentals of the stock and use advanced mathematical models. Thus the $AR(p)$ of order p model can be regarded as a simplest simulator of a nonprofessional player which is making investments based on the data observed during past p days.

The profit of the stockholder depends on the accuracy of prediction made at time s , $s = 1, \dots, t$, where t denotes the present time.

Assume that the stock rates changes following this simple relation

$$Z(s+1) = \sum_{k=1}^p a_k Z(s-k+1) + \epsilon_{s+1}. \quad (1)$$

This formula describes the common autoregressive model $AR(p)$ of order p . However, in the contest of this paper, relation (1) reflects opinions of stockholders which are making investment decisions based on the optimal next day predictions obtained using the past data. It means that we replace the standard assumptions of the autoregressive model by the single assumption that the relation (1) approximately represents opinions of some stockholders. In this paper, we shall compare the prediction models which minimize standard statistical prediction errors, such as MSE and MAE, with the models maximizing simulated profit.

The alternative way of fitting $AR(p)$ parameters is the likelihood maximization which provides good mathematical results [13]. However, this approach appears more difficult for stockholders intuitive understanding and the mathematical advantages are not so important regarding the $AR(p)$ model just as a tool of the virtual stock exchange. We consider moving average model $MA(q)$, too, to simulate more sophisticated users which try to correct past errors, where

$$Z(s+1) = - \sum_{j=1}^q b_j \epsilon_{s-j+1} + \epsilon_{s+1}. \quad (2)$$

Minimizing the $MA(q)$ errors we have to minimize a polynomial function of degree t . We can see this by expanding the recurrent expression (2). Traditional methods of

¹ $ARMA-ABS(p, q)$ minimizes absolute errors of $ARMA(p, q)$ models where q is the number of moving average parameters b_j , $j = 1, \dots, q$.

parameter estimation do not consider this problem as multimodal [2]. However, some more recent authors apply global optimization techniques such as particle swarm optimization [10], evolutionary and genetic algorithms [11, 3]. The investigation of multimodality was not performed in these papers, so this work can be regarded as the first step in this direction.

We define the $ARMA(p, q)$ model by merging $AR(p)$ and $MA(q)$ into this single expression:

$$Z(s + 1) = \sum_{k=1}^p a_k Z(s - k + 1) - \sum_{j=1}^q b_j \epsilon_{s-j+1} + \epsilon_{s+1}. \tag{3}$$

To represent risk-neutral users, we may apply the $ARMA-ABS(p, q)$ model by minimizing the absolute errors instead of the squared ones.

1.1 $ARMA(p)$ and $ARMA-ABS(p, q)$ models

Denote by $Z(s)$ the stock price at time $s \leq t$. Denote by $a = (a_1, \dots, a_p)$ a vector of auto-regression (AR) parameters, and by $b = (b_1, \dots, b_q)$ a vector of moving-average (MA) parameters. Then the prediction error (residual) at moment $s \leq t$ is

$$\epsilon_s = Z(s) - \sum_{k=1}^p a_k Z(s - k) + \sum_{j=1}^q b_j \epsilon_{s-j}. \tag{4}$$

The optimal prediction parameters $a = a(b)$ as a function of b in the $ARMA(p, q)$ model, are defined by this condition

$$a_k = \arg \min_{a_k} \sum_{s=1}^t \epsilon_s^2. \tag{5}$$

In the $ARMA-ABS(p, q)$ model, we minimize absolute errors

$$a_k = \arg \min_{a_k} \sum_{s=1}^t |\epsilon_s|. \tag{6}$$

To solve non-linear optimization problem (5) we use the recursive quadratic programming [1]. The piece-wise linear problem (6) is posed in the following standard linear programming terms [9]:

$$\min_{v, u} \sum_{s=1}^t u_s, \tag{7}$$

$$u_s \geq \epsilon_s, \quad s = 1, \dots, t, \tag{8}$$

$$u_s \geq -\epsilon_s, \quad s = 1, \dots, t, \tag{9}$$

$$u_s \geq 0, \quad s = 1, \dots, t, \tag{10}$$

$$v_k^1, v_k^2 \geq 0, \quad k = 1, \dots, p, \tag{11}$$

$$\epsilon_s = Z(s) - \sum_{k=1}^p (v_k^1 - v_k^2) Z(s - k). \tag{12}$$

where v, u are the auxiliary variables and ϵ_s are residuals. The scale parameters a_k^i of the player i are defined by differences of auxiliary variables v .

Optimization of the b parameters is performed using the global optimization methods in both the models because the errors as a function of $b \in [-1, 1]$ may have up to t local minima where t is the length of time series.

2 Virtual stocks and historical data

Fig. 1 illustrates average profits of eight prediction models for two virtual ‘years’ equal to 720 virtual ‘days’.

Fig. 2 shows the average MSE of 100 samples of virtual stocks generated by USEGM.

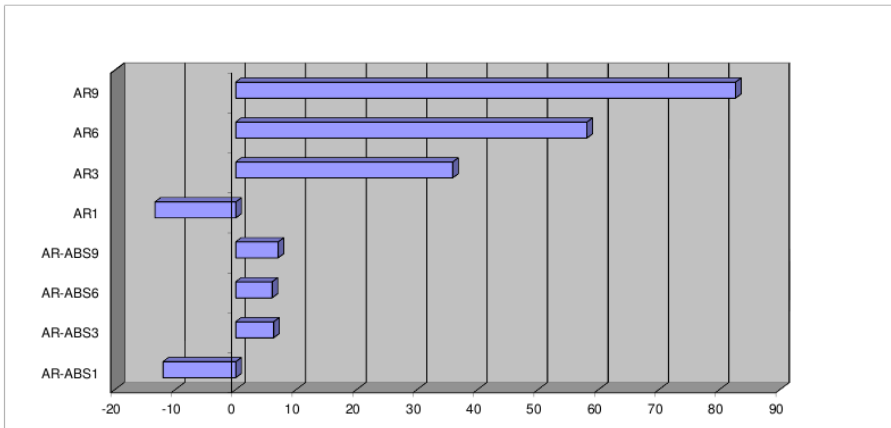


Fig. 1. Average final profits of eight prediction models, $a = 0.0$.

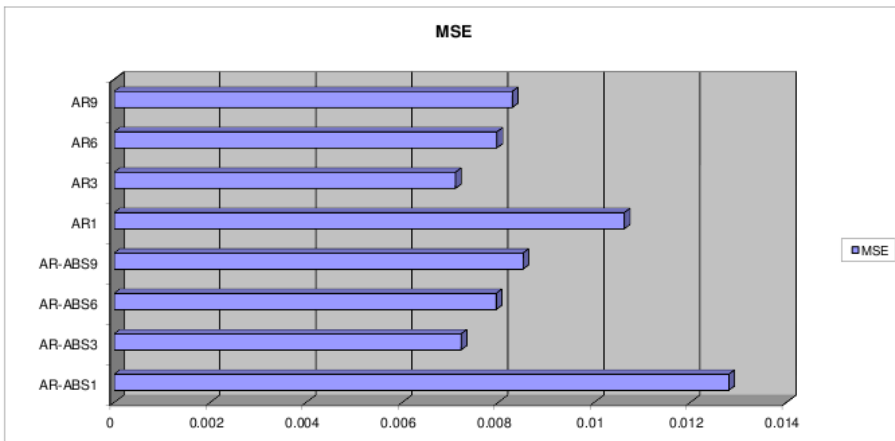


Fig. 2. MSE of virtual stock market, average of 100 samples by USEGM, $a = 0.0$.

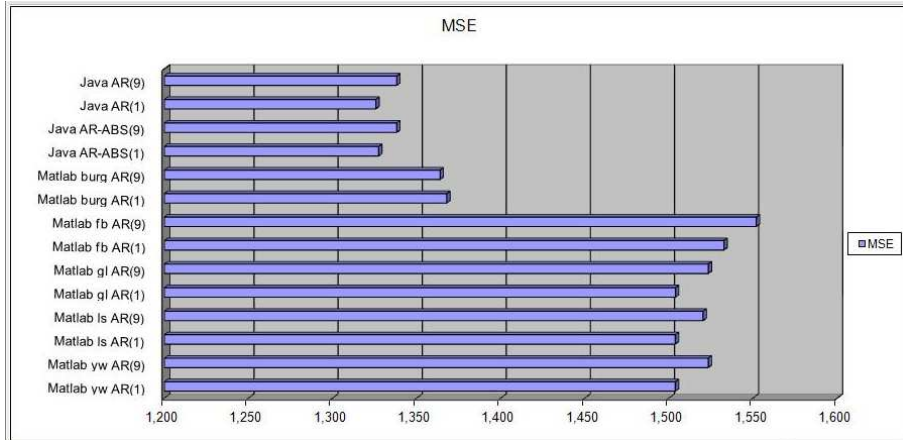


Fig. 3. MSE of real stock market, including Matlab results, average of 200 stocks.

In all the experiments, the maximal profits were provided using the maximal memory parameter $p = 9$. The minimal errors were achieved using lesser memory $p = 3$ with the only exception when minimal MAE and MSE were achieved using $p = 6$. This means that the prediction models which provide minimal errors do not necessarily show the maximal profit.

Now we shall test the prediction algorithms by historical data is obtained automatically from the Yahoo data base. Fig. 3 shows the average MSE of 200 stocks using $AR(p)$, $p = 1, 3, 6, 9$ including the corresponding results by the best Matlab version of the AR model.

The Matlab prediction errors are greater. A possible explanation is that Matlab AR model estimates an additional parameter defining the mean value of the time series. That provides lesser ‘learning’ error but not necessarily the minimal ‘testing’ error. The USEGM AR model ignores this parameter and predicts only the stock rate changes which are important making investment decisions.

3 Concluding remarks

While autoregressive and moving average models may be too simplistic for practical forecasting, it can serve as a useful tool for studies of market behavior by presenting an easy way for simulating different scenarios of stockholders strategies. For example, simulations can explain stock market reaction to deliberately set strategies of a major stockholder, such as manipulation of asset prices, designed to lower their value.

We have investigated the autoregressive model with 200 actual financial time series and 100 “virtual” stock rates generated by the USEGM model [5]. The unexpected, and practically the most important, result is that the prediction models which provide minimal errors do not necessarily show the maximal profit. Therefore, to estimate the expected profitability better, we have to simulate the financial markets using and improving the corresponding market models, similar to USEGM. The algorithm and software is on the website [6] and can be tested by a browser with Java support.

References

- [1] M.C. Biggs. Constrained minimisation using recursive quadratic programming: some alternative subproblem formulations. In *Towards Global Optimisation*. North Holland, Amsterdam, 1975.
- [2] J.H. Cochrane. *Time Series for Macroeconomics and Finance*. University of Chicago, Graduate School of Business, Chicago, U.S., 2005.
- [3] C. Gaetan. Subset (ARMA) model identification using genetic algorithms. *J. Time Ser. Anal.*, **21**:561–570, 2002.
- [4] C. Hongfang and S. Rui. ARMA model parameter optimized estimate method. In *Cryptography, and Network Security, Data Mining and Knowledge Discovery, E-Commerce and Its Applications, and Embedded Systems, IACIS International Symposium*, pp. 22–26, Los Alamitos, CA, USA, 2010. IEEE Computer Society. Available from Internet: <http://doi.ieeecomputersociety.org/10.1109/CDEE.2010.13>.
- [5] I. Katin. The Java software implementing the market simulation model USEGM. Available from Internet: <http://optimum2.mii.lt/igor/index.html>, 2012.
- [6] J. Katina. The Java software implementing the prediction models $AR(p)$ and $AR-ABS(p)$. Available from Internet: <http://optimum2.mii.lt/joana/launch.html>, 2012.
- [7] G. Melard. Algorithm AS 197: A fast algorithm for the exact likelihood of autoregressive-moving average models. *Appl. Stat.*, **33**:104–114, 1984.
- [8] J. Mockus. *A Set of Examples of Global and Discrete Optimization: Application of Bayesian Heuristic Approach*. Kluwer Academic Publishers, 2000. ISBN 0-7923-6359-0.
- [9] J. Mockus. On simulation of the Nash equilibrium in the stock exchange context. *Informatica*, **23**:77–104, 2012.
- [10] S. Rolf, J. Spravey and W. Urferz. Using particle swarm optimization and AIC criteria. In *Proceedings of the IEEE/IAFE*. IEEE, 1995.
- [11] M.S. Voss and X. Feng. Model identification and parameter estimation of ARMA models by means of evolutionary algorithms. In *15th Triennial World Congress, Barcelona, Spain*. IFAC, 2002.
- [12] H. Wagner. Linear programming techniques for regression analysis. *J. Am. Stat. Assoc.*, **54**:206–212, 1959.
- [13] I. Yao and P.J. Brockwell. Gaussian maximum likelihood estimation for ARMA models II: Spatial processes. *Bernoulli*, **12**:403–429, 2006.

REZIUMĖ

Autoregresiniai slenkančio vidurkio modeliai kaip virtualios finansų biržos įrankis: eksperimentinis tyrimas

J. Mockus, J. Katina, I. Katin

Darbo tikslas yra eksperimentiškai ištirti tradicinius autoregresinius modelius kaip paprasčiausius algoritmus, imituojančius akcininkų, naudojančių vien tik istorinius duomenis, prognozių procesus. Šie modeliai naudojami virtualioje akcijų biržoje tikrinant įvairias prielaidas apie akcininkų elgesį. Įvertinant neutralius rizikai akcininkus, tradicinis AR modelis, minimizuojantis kvadratinės paklaidas, papildomas modeliu $AR-ABS$, minimizuojančiu absoliutines paklaidas. Trumpai aptariami 200 realiųjų akcijų ir 100 virtualių akcijų tyrimo rezultatai, minimizuojantys vidutines paklaidas.

Raktiniai žodžiai: prognozavimas, autoregresiniai modeliai, akcijų birža, laiko eilutės.