ÚTIA

Akademie věd České republiky
Ústav teorie informace a automatizace

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

JIŘÍ KRTEK

COMPARING NEURAL NETWORK AND REGRESSION MODELS IN ASSET PRICING MODEL WITH HETEROGENEOUS BELIEFS

No. 2252 June 2009

ÚTIA AV ČR, P. O. Box 18, 182 08 Prague, Czech Republic
Fax: (+420)(2)86890378, E-mail: utia@utia.cas.cz

Comparing Neural Networks and regression models in Asset Pricing Model with Heterogeneous Beliefs

Jiri Krtek, Miloslav Vosvrda [1]

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

Department of Econometrics

Abstract:
The competition of four forecasting strategies in artificial market is studied in this paper. The environment of the market is modeled by adaptive belief system. Two neural networks was included in the quaternary of forecasting strategies. They were compared with rule of thumb and linear regression. BMV share was used as a risky asset which price strategies predicted. PX index, USD/CZK spot rate and Czech rep. GDP data was simulated in purpose to be inputs of two of models which probably caused instability of the market.

Keywords: neural networks, regression, adaptive belief system

JEL: C45, C53, G12

1 Introduction

If we are working with time series, we usually use linear regression or some ARIMA model. The more experienced of us could use non-linear regression, but this brings more difficulties.

[1]Supervisor

It needs more sense and experiences. The simple way how to bring non-linearity in the relation between regressors and explanatory variable is to choose suitable neural network. A small competition of four simple models for prediction of asset price is described in this paper. Two of the models predict the asset price at time t on the bases of the asset price at time $t-2$. Two other models predict asset price at time t on the bases of a value of PX index, USD/CZK spot rate and GDP of Czech Republic all at time $t-2$. Two of them are neural networks, 1 is linear regression, and the last one is simple "rule the thumb", which estimates asset price at time t by the value of asset price at time $t-2$. A simplified asset pricing with heterogeneous beliefs from [1] was chosen as a model in which mentioned forecasting strategies compete.

2 Adaptive Belief System

An adaptive belief system from Brock and Hommes, called Asset Pricing with Heterogenous Beliefs, was chosen to simulate market environment. We have population of agents who can choose from different forecasting strategies. They select one of them in each time step t. They base their decisions on the past performance of the mentioned strategies. They can invest in risk free asset with fixed rate of return r or in a risky asset. Let p_t be the price of risky asset at time t. Let W_{ht} denote wealth at time t of an agent who uses investment strategy h. The following recurrent formula holds for wealth of an agent of type h

$$W_{h,t+1} = (1+r)W_{ht} + (p_{t+1} - (1+r)p_t)z_{ht}, \tag{1}$$

where z_{ht} is his demand for risky asset at time t.

Agents are myopic mean-variance maximizers of wealth so they want to maximize

$$\max_{z_{ht}} E_{ht}W_{h,t+1} - \frac{1}{2}aV_{ht}W_{h,t+1}, \tag{2}$$

where E_{ht} is expectation of an agent of type h at time t, V_{ht} is conditional variance of an agent of type h at time t and a is the risk aversion parameter. From which we get

$$z_{ht} = \frac{E_{ht}[p_{t+1} - (1+r)p_t]}{aV_{ht}[p_{t+1} - (1+r)p_t]} = \frac{E_{ht}[p_{t+1} - (1+r)p_t]}{a\sigma^2}, \tag{3}$$

where the last equation holds when beliefs about conditional variance of excess return $p_{t+1} - (1+r)p_t$ are assumed to be constant for all strategies, i.e. $V_{ht}[p_{t+1} - (1+r)p_t] = \sigma^2$.

Suppose that this artificial market is closed, i.e. there is no outside supply of the risky asset. Let n_{ht} denote fraction of traders using at time t strategy h and H number of strategies. Then from 3 we get

$$\sum_{h=1}^{H} n_{ht}z_{ht} = \sum_{h=1}^{H} n_{ht}\frac{E_{ht}[p_{t+1} - (1+r)p_t]}{a\sigma^2} = 0, \tag{4}$$

from which we get formula for price of the risky asset at time t

$$\sum_{h=1}^{H} n_{ht} E_{ht}[p_{t+1}] = \sum_{h=1}^{H} n_{ht} f_{ht} = (1+r)p_t, \tag{5}$$

where $f_{ht} = E_{ht}[p_{t+1}]$ denotes forecasting strategy h. It may be a little confusing that the last known information for f_{ht} is from time $t-1$.

It was mentioned that agents select their strategy on the basis of its past performance. The fitness measure

$$U_{ht} = (p_t - (1+r)p_{t-1})z_{ht} + wU_{h,t-1} = (p_t - (1+r)p_{t-1}) \frac{E_{h,t-1}[p_t - (1+r)p_{t-1}]}{a\sigma^2} + wU_{h,t-1}, \tag{6}$$

is mathematical formularization of performance of strategy h at time t, $0 \leq w \leq 1$ is a memory parameter.

The fraction of trader type h at time $t+1$ is then given by

$$n_{h,t+1} = \frac{\exp(\beta U_{ht})}{Z_t},\ Z_t = \sum_{h=1}^{H} \exp(\beta U_{h,t-1}), \tag{7}$$

where parameter β is the intensity of choice.

This leads to co-evolution of trader types fractions and price of the risky asset. When we know fractions of trader types at time t, i.e. n_{ht}, we can compute p_t, from which we can compute U_{ht} and then $n_{h,t+1}$, and so on. So we can simulate market by this procedure.

3 Neural Networks

Neural networks are one of attempts at creating artificial intelligence. They are inspired by nature, especially by brain and his abilities to learn and generalize. We can recognize patterns or solve various optimization problems with them. They be used for approximation of various functions. They are used for forecasting in this paper.

Their main benefits are that they bring non-linearity in the problems, they are robust, they don't need any assumption and of course their generalization ability.

On the other side it is hard to choose their topology and to find their optimal parameters. Black-box property is also one of their drawbacks and don't forget on the possibility of overfitting them.

In this paper the feedforward neural networks with one hidden layer and one output are used.

We will describe active dynamics of perceptron and feedforward neural networks with one hidden layer and one output.

Let $\mathbf{x} = (x_0, x_1, \ldots, x_n)^T$, where $x_0 = 1$, be a vector consisting of a unit input and n inputs $x_1, \ldots, x_n$ of a perceptron. Let $\mathbf{w} = (w_0, \ldots, w_n)^T$, where w_0 is so called threshold or bias, denote a given vector of weights. Then the output $y = y(\mathbf{x}, \mathbf{w})$ of the perceptron is computed by following formula

$$y = f(\mathbf{x}^T.\mathbf{w}), \tag{8}$$

where f is some activation function, typically sigmoid function, i.e. $f(x) = \frac{1}{1+\exp\{-x\}}$.

The active dynamics of feedforward neural networks with one hidden layer of h perceptrons with sigmoid as activation function and one output is more complicated.

Suppose we have again n inputs $x_1, \ldots, x_n$ and let $\mathbf{x} = (1, x_1, \ldots, x_n)^T$. Let $\mathbf{w}_i = (w_{i0}, \ldots, w_{in})^T$, $i = 1, \ldots, h$, denote given vector of weights related to perceptron i from hidden layer and let $\mathbf{v} = (v_0, v_1, \ldots, v_h)^T$ denote given vector of weights related to neuron from output layer. We firstly compute outputs of perceptrons from the hidden layer u_i, $i = 1, \ldots, h$, and then use these outputs to compute the output $y = y(\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_h, \mathbf{v})$ of the network as

$$u_i = \quad f(\mathbf{x}^T.\mathbf{w}_i),\ i = 1, \ldots, h \tag{9}$$

$$y = \quad (1, u_1, \ldots, u_h).\mathbf{v}. \tag{10}$$

The adaptive dynamics or training of feedforward neural networks is quite difficult procedure. Suppose we have chosen some topology of neural network and wa want to find its optimal parameters, i.e. optimal vectors of weights. Let $M = \{(\mathbf{x}_k, y_k), k = 1, \ldots, N\}$ be training set which contains N input-output pairs. Our goal is to minimize error of the network, i.e.

$$\min_{\mathbf{w}} E(\mathbf{w}), \tag{11}$$

where $\mathbf{w}^T = (\mathbf{w}_1^T, \ldots, \mathbf{w}_h^T, \mathbf{v}^T)$ and

$$E(\mathbf{w}) = \sum_{k=1}^{N} (y(\mathbf{x}_k, \mathbf{w}) - y_k)^2. \tag{12}$$

The basic algorithm for minimizing error of the network and finding optimal parameters $(\mathbf{w})$ is backpropagation. It is omitted from this paper, see [2].

4 Application

It was already mentioned that the adaptive belief system was chosen to serve as a model with which we want to compare four different simple forecasting strategies.

We had BMV share daily data from Jan 1st, 2003 till Oct 3rd, 2008. The BMV share was chosen as the risky asset. We also had PX index daily data from Sep 7th, 1993 till Apr 3rd, 2009, USD/CZK spot rate daily data from Jan 1st, 1991 till Apr 3rd, 2009 and Czech Republic GDP quarterly data from 1st quarter 1996 till 4th quarter 2008. This data was chosen to serve as regressors in two of the models for forecasting.

The simplest model forecasts the price of the risky asset at time $t+1$ by the price at time $t-1$. The same input was also chosen to the second model which was neural network with one input, one output, one hidden layer with one perceptron with sigmoid as activation function. Third model was linear regression with price of BMV share at time $t+1$ as dependent variable and PX index, USD/CZK spot rate and GDP all from time $t-1$ as explanatory variables. Fourth model was again neural network with PX index, USD/CZK spot rate and GDP from time $t-1$ as inputs and price of BMV share at time $t+1$ as output. It had 1 hidden layer with 3 perceptrons with sigmoid as activation function. The second, third and fourth model were fitted from data. Neural networks was trained by Levenberg–Marquardt algorithm which is just modification of backpropagation. The RMSE of models are in Table 1.

Model	RMSE
Rule of thumb	0.824919
NN1i	0.82367
Linear regression	2.8179
NN3i	1.6 – 3.1

Table 1: RMSE of models; NN1i means Neural network with 1 input and NN3i means Neural network with 3 inputs.

The $ARMA(1,1)$ models was used for modeling data of PX index and USD/CZK spot rate. A process with linear trend and additive independent jumps with normal distribution and time to jump drawn from geometric distribution was chosen to model GDP data. These processes was utilized in simulation where inputs of the third and fourth mentioned model were drawn from these models.

The simplified asset pricing with heterogeneous beliefs was used to simulate the market, the mentioned four models were taken as the forecasting strategies. These models got the last known value of asset price or simulated values of PX index, USD/CZK spot rate and GDP from last time period as inputs. Then new price of risky asset was made from their forecasts.

The simulation was done for two different intensities of choice – for $\beta = 1$ and for $\beta = 30$. The moving of the price of the risky asset and of the fractions of trader types can be seen at Figure 1 and Figure 2 respectively for the first simulation and at Figures 3 and **??**. There is no steady state for both cases. The fractions of traders of type one and two are quite stable while the fractions of traders of type three and four change a lot and go against themselves. We can see that none of the strategies dominates. But another simulation showed dominance of first two strategies. It had 1000 iteration cycles, each with time horizon of the market equal to 1000. The average profits of the strategies were investigated in this simulation, i.e. profits of traders who never changed their strategies. The results are summarized in Table 2.

The data were analyzed mainly in R–project, the whole simulation was implemented in Mathematica.

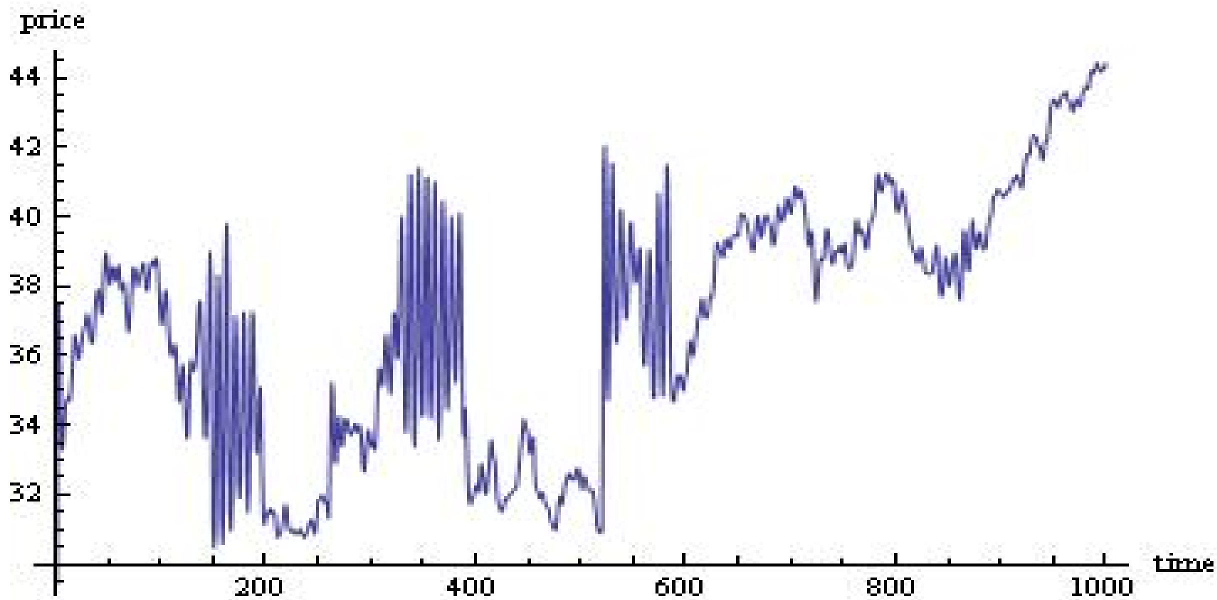

Figure 1: 1st simulation – plot of evolution of the price of risky asset; $\beta = 1$, $a = 3$, $\sigma^2 = 3$, $r = 1.01$, $w = 0$.

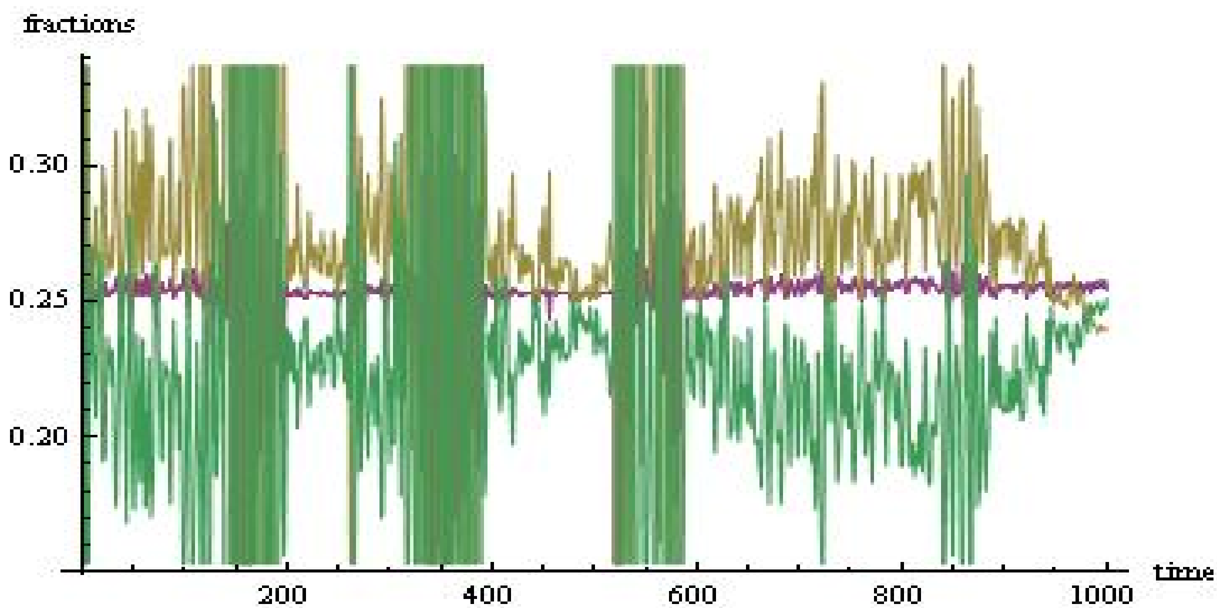

Figure 2: 1st simulation – plot of fractions of trader types; $\beta = 1$, $a = 3$, $\sigma^2 = 3$, $r = 1.01$, $w = 0$; Rule of thumb – blue, NN1i – violet, Linear regression – beige, NN3i – green (blue line and violet line almost merge).

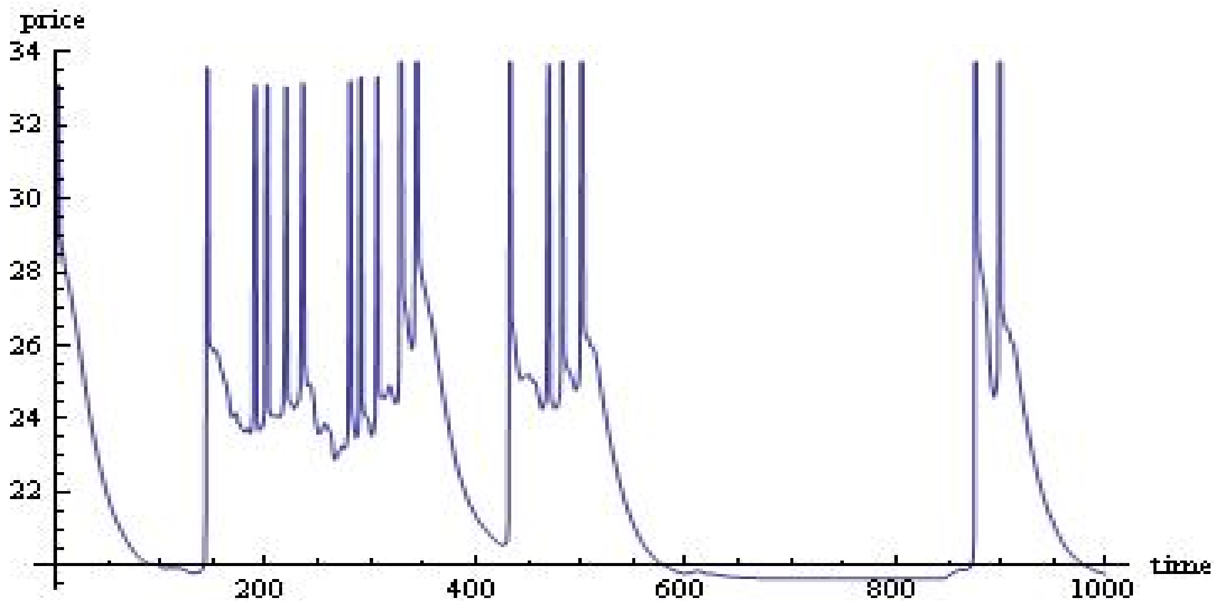

Figure 3: 2nd simulation – plot of evolution of the price of risky asset; $\beta = 30$, $a = 3$, $\sigma^2 = 3$, $r = 1.01$, $w = 0$.

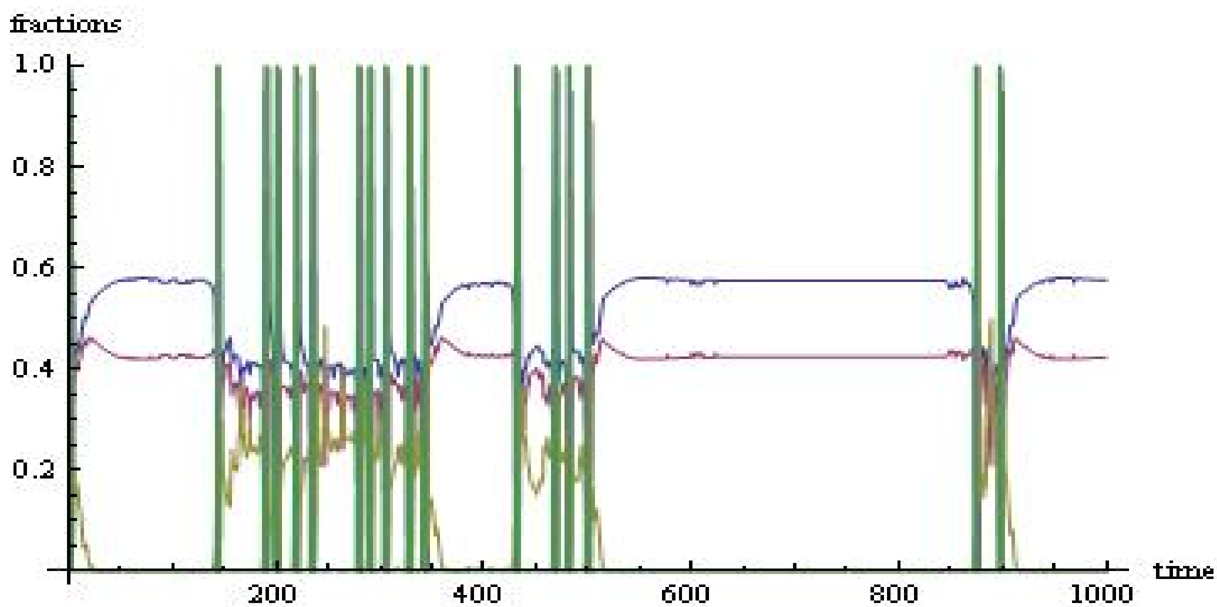

Figure 4: 2nd simulation – plot of fractions of trader types; $\beta = 30$, $a = 3$, $\sigma^2 = 3$, $r = 1.01$, $w = 0$; Rule of thumb – blue, NN1i – violet, Linear regression – beige, NN3i – green (blue line and violet line almost merge).

Model	Average profit
Rule of thumb	1.29802
NN1i	1.27812
Linear regression	0.772721
NN3i	0.551962

Table 2: Average profits of models.

5 Conclusion

There is no steady state in the first simulation as it is usual for small intensities of choice and simple strategies. We can see big oscillation of fractions of third and fourth strategy at times around 170, 350 and 550 at Figure 2 and also we can see oscillation of a price of the risky asset at the same times at Figure 1, it is probably caused by the simulated data.

There are usually repeating cycles when we choose higher intensity of choice. However, there is no repeating cycle in the simulation with $\beta = 30$ as can be seen from Figures 3 and 4. It is again probably caused by the simulated data.

There is no dominance of any model on the simulated market, but if we choose rule of thumb and never change this strategy we will earn more than with any other strategy. We can also say that rule of thumb and linear regression outperform their equivalent from neural networks. But we must in the same breath add that only very simple neural networks were used in this paper.

References

[1] Cars Hommes. Heterogeneous agent models in economics and finance. Discussion Paper TI 05-056/1, Tinbergen Institute, 2005.

[2] Jiri Sima and Roman Neruda. *Teoreticke otazky neuronovych siti*. Matfyzpress, Praha, 1996.