Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

ASSOSIATIVE SEARCH MODELS IN TRADING

Natalya N. Bakhtadze, Vladimir A. Lototsky, Ravil T. Valiakhmetov

Institute of Control Sciences, 65, Profsoyuznaya, Moscow,117997 GSP, Russia
Tel.: +(7)(095)3349201. E-mail: lotfone@ ipu.rssi.ru

Abstract. The paper presents process identification algorithms for trading based on virtual models design using process data archives and knowledge bases. Fuzzy Associative Search Methods are used for identification algorithms development.

Keywords. Process identification, knowledgebase, virtual models, fuzzy associative search models

1. INTRODUCTION

Short-term prediction of financial indexes or stock prices is a most important and difficult problem in financial mathematics. A lot of methods and rules had been developed within 2 basic prediction philosophies (fundamental and technical analyses). The classical ***fundamental analysis*** is based on well developed mathematical statistics and time series theory tools (Shiryaev, 1995).

For long-term prediction, investigations had shown high autocorrelation coefficient between pricing factors in regression models. This indicates the stable predictable dynamics.

The price dynamics specified by the correlation between factors is difficult to discover and leads to long-term price deviation from the median value. This results in low effectiveness of regression models application in trading. One-step recurrent identification algorithms show better results but they not always comply with time and precision requirements. One reason for this is weak formalization of some pricing factors such as the influence of political situation on financial markets.

Technical analysis helps to identify various market pattern abnormalities using visual graphical analyses of historical data and automatic identification of standard geometrical forms (such as head and shoulders, flags, trends), i.e., by identifying the most common patterns typical for a certain state of the market. This method have not yet obtained official recognition, albeit its good effectiveness for medium and short-term forecasting on foreign exchange market (more than 90% of FOREX traders in London are using technical analyses for short-term prediction 2–4 weeks ahead). But the main postulate of technical analyses – "market repeats", i.e., the present assets price is determined by its dynamics in the past – is obviously noteworthy. A method using prediction models based on the imitation of analyst's or trader's associative thinking can be considered as an alternative to technical analysis.

Identification algorithms employed in modern control systems often use expert knowledge both from human expert or from a knowledgebase. In the second case, a trader can choose between recommended control action (a decision to buy or to sell) or a forecast based on market state monitoring.

Two knowledge types are distinguished: declarative and procedural (Larichev *et al*., 2001). The first type includes the description of various facts, events, and

978-1-1234-7890-2/08/$20.00 © 2008 IFAC 4280 10.3182/20080706-5-KR-1001.0857

observations, while skills and experience refer to the second type. Experts (in our case, analysts and traders) differ from novices by their structure and way of thinking and, in particular, the searching strategy (Patel and Ramoni, 1997). If a person is not experienced, (s)he would uses the so-called 'backward reasoning'. (S)he reviews different possible answers and makes a decision in favor of a specific answer based on the information received from the process at the current time step. On the contrary, an expert does not need to analyze current information in the process of decision-making, rather (s)he uses the so-called 'forward reasoning' method which implies that the decision-making strategy is created subconsciously and this strategy is nonverbal. Therefore, in terms of the method of *computational view of thought* (Hunt, 1989), the effectiveness of system will to a great extent be determined by expert's qualification and by the available a priori information. Within the framework of this method, the cognitive psychology determines *knowledge* as a certain set of actually existing elements-symbols stored in human memory, processed during thinking and determining the behavior. The symbols, in turn, could be determined by their structure and the nature of neuron links. (Simon, 1997).

Knowledge processing in an intelligent system consists in the recovery (*associative search*) of knowledge by its fragment (Gavrilov, 2002). The knowledge can be defined as an associative link between *images.* The associative search process can take place either as a process of image recovery using partially specified symptom (or knowledge fragment recovery by incomplete information; this process is usually emulated in various associative memory models) or as searching others images (linked associatively with the input image) related with other time steps. Those images make sense of a cause or an effect of an input image.

In (Gavrilov, 2002) a model is offered which describes the associative thinking process as a sequential process of remembering based on *associations* – pairs of images defined by a set of symptoms. Such model can be considered as an intermediate level between neuron network models and logical models used in classical artificial intelligence systems. In this paper, we discuss an approach to developing on-line support of trader's decision-making based on the dynamic simulation of associative search and the identification technique based of virtual models.

Представленный в настоящем докладе метод является алгоритмическим аналогом графического подхода, реализуемого методом технического анализа. Приводимые в разделе

2. NONLINEAR PREDICTION ALGORITHM BASED ON VIRTUAL MODELS DEVELOPMENT

An identification algorithm for complex nonlinear dynamic objects such as continuous and batch processes was presented in (Chadeev, 2004). The identification algorithm with continuous real-time self-tuning is based on *virtual models* design.

At every time step, a new virtual model is created. To build a model for a specific time step, a temporary "ad hoc" database of historic and current process data is generated. After calculating the output forecast based on object's current state, the database is deleted without saving.

The linear dynamical prediction model looks as follows:

$$y_t = a_0 + \sum_{i=1}^{r} a_i y_{t-i} + \sum_{j=1}^{s} \sum_{p=1}^{P} b_{jk} x_{t-j,p}\,, \qquad (1)$$

where y_t is the object's output forecast at the t-th step, x_t is the input vector, r is the output memory depth, s is the input memory depth, P is the input vector length.

The original dynamic algorithm consists in the design of an approximating hypersurface of input vector space and the related one-dimensional outputs at every time step (see Figure 1). To build a virtual model for a specific time step, the points close in a manner to the current input vector are selected. The output value at the next step is further calculated using least mean squares (LMS).

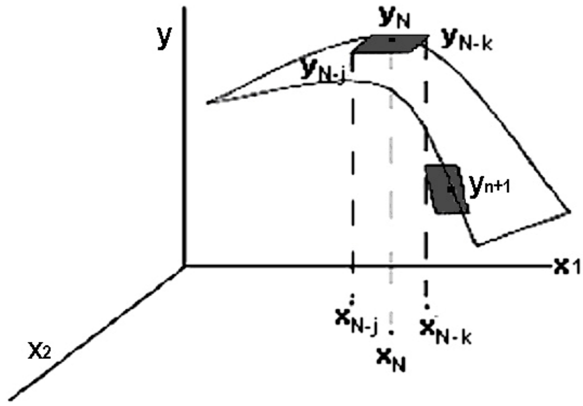

Fig. 1. Approximating hypersurface design

3. ASSOCIATIVE SEARCH TECHNIQUE FOR VIRTUAL MODELS DESIGN

For trading problems, we use a method based on the associative thinking model.

High-speed approximating hypersurface design algorithms enabling the usage of fuzzy models for various process applications were offered in (Bakhtadze *et al*., 2007).

The following quantity

$$d_{t,t-j} = \sum_{p=1}^{P} \left| x_{tp} - x_{t-j,p} \right|, \; j = 1,\ldots,s\,, \qquad (2)$$

was introduced as distance (metric in $\Re^P$) between points of P-dimensional input space, where, generally, $s < t$, and x_{tp} are the components of the input vector at the current time step t.

Assume that for the current input vector $\boldsymbol{x}_t$:

$$\sum_{p=1}^{P} \left| x_{tp} \right| = d_t\,. \qquad (3)$$

To build an approximating hypersurface for $\boldsymbol{x}_t$, we select such vectors $\boldsymbol{x}_{t-j}$, $j = 1,\ldots,s$ from the input data archive that for a given D_t the following condition will hold:

$$d_{t,t-j} \le d_t + \sum_{p=1}^{P} \left| x_{t-j,p} \right| \le d_t + D_t,\ j = 1,\dots,s\,. \qquad (4)$$

The 2-D case is illustrated below (Fig. 2).

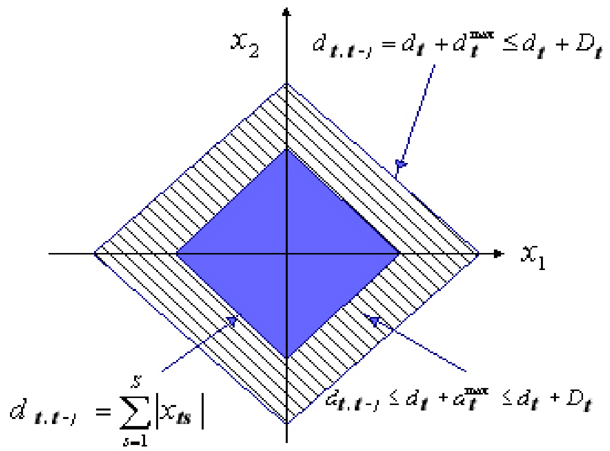

Fig.2. An approximating hypersurface building

The preliminary value of D_j is determined on the basis of process knowledge. If the selected domain does not contain enough inputs for applying LMS, i.e., the corresponding SLAE has no solution, then the chosen points selection criterion can be slackened by increasing the threshold D_t.

To increase the speed of the virtual models-based algorithm, an approach is applied based on employing a model of analyst's or trader's associative thinking for predicting.

For modeling the associative search procedure imitating the intuitive prediction of process status by a trader, we assume that the sets of process variable values, which are the components of an input vector, as well as the system outputs at previous time steps altogether create *a set of symptoms, making an image of the object output at the next step.*

The associative search process consists in the recovery of all symptoms describing the specific object based on its images. Denote the image initiating the associative search by R_0 and the corresponding resulting image of the associative search by R. A pair of images (R_0,R) will be further called *association A* or $A(R_0,R)$. The set of all associations on the set of images forms the *memory* of the *knowledgebase* of the intelligent system.

At the system learning phase, an archive of images is created. In our case, a set of n input vectors selected form the process history according to the algorithm described in Section 1 will be considered as an image.

At the prediction stage, the input vector x_t will be considered as an initial image R_0^a of the associative search, while approximating hypersurface formed by the input vectors from the process history built with the help of the algorithm from Section 1 will be the final image R^a of the associative search. This means that to build a virtual model, one should select the existing hypersurfaces stored in the archive at the learning phase rather than individual vectors close to x_t. The selected hypersurface is an image of the current input vector which is used for output prediction. The algorithm implements the process of image R^a recovery based on R_0^a, i.e., the associative search process, and can be described by a predicate $\Xi=\{\Xi_i(R_{0i}^a,R_i^a,T^a)\}$, where $R_{0i}^a \subseteq R_0$, $R_i^a \subseteq R$, and T^a is the duration of the associative search.

For the algorithm described in Section 2, this predicate is a function asserting the truth or the falsity of input vector's membership of the specific domain in the inputs space. Therefore, the associative search process is reduced to the selection of a certain set of input vectors satisfying the condition (4) from the process archive. If the process history contains no image satisfying (4), then either the threshold D_t should be increased, or for a certain image of our input vector, some symptom should be replaced with a more relevant one. Formally, this means that the "worst" (i.e., located farther away from the current input then the rest ones w.r.t. the chosen criterion) vector from the process history will be deleted and replaced with a more relevant one, and so on.

Therefore, the analyst's decision-making process about to buy or to sell at any time step t could be constructed as associative search (process of remembering) of images (similar situations). The coordinates of approximating hypersurfaces used at previous steps are kept in archive.

4. ASSOCIATIVE SEARCH PROCEDURES IN SHORT TERM FORECASTING

In short-term price prediction, not only the current situation but also price dynamics is very important. The conventional regression models are not precise enough to handle this problem.

We apply the associative search procedure with more complicated requirements to approximating hypersurface selection. We select from the archive such hypersurface (corresponding to some $\boldsymbol{x}_{t-j}$, $j = 1,\dots,s$), that (i) it contains input vector at the current time step t, and (ii) the hypersurface corresponding to $\boldsymbol{x}_{N-j-1}$, $j = 1,\dots,s$ contains the input vector at the previous time step t-1. Formally, this means that the predicate becomes more complex:

$$\Xi\,(R_0^a,R^a,T^a) = \left\{ \sum_{p=1}^{P} \left| x_{t-j,p} \right| \le D_t - \sum_{p=1}^{P} \left| x_{tp} \right|;\ \sum_{p=1}^{P} \left| x_{t-j-1,p} \right| \le D_{t-1} - \sum_{p=1}^{P} \left| x_{tp} \right| \right\} \quad (5)$$

There is principal opportunity to find more precise rules in the process of price changing by increasing the memory, say, to l steps ($l<t$).

$$\Xi\,(R_0^a,R^a,T^a)= \left\{ \sum_{p=1}^{P} \left| x_{t-j,p} \right| \le D_t - \sum_{p=1}^{P} \left| x_{tp} \right|; \right.$$
$$\left. \sum_{p=1}^{P} \left| x_{t-j-1,p} \right| \le D_{t-1} - \sum_{p=1}^{P} \left| x_{tp} \right|;\dots, \sum_{p=1}^{P} \left| x_{t-l,p} \right| \le D_{t-l} - \sum_{p=1}^{P} \left| x_{tp} \right| \right\} \qquad (6)$$

5. FUZZY VIRTUAL MODELS

Assume that one or more variables in (1) are fuzzy. In real life this may mean the fuzzification of weekly recommendation provided by major investment banks that in this case are considered as experts.

Generally, (1) can be represented as a fuzzy Takagi–Sugeno (TS) model (Takagi and Sugeno, 1985). The fuzzy TS model consists of a set of production rules with linear finite difference equations in the right-hand member (for simplicity, a single input case, i.e., P=1, is considered):

If $y(t-1)$ is Y_1^θ,…, $y(t-r)$ is Y_r^θ,

$x(t)$ is X_0^θ,…, $x(t-s)$ is X_r^θ,

then $y^\theta(t) = a_0^\theta + \sum_{k=1}^{r} a_k^\theta y(t-k) + \sum_{l=0}^{s} b_l^\theta x(t-l)$,

$$\theta = 1,\dots,n, \tag{7}$$

where $\boldsymbol{a}^\theta = (a_0^\theta, a_1^\theta,\dots,a_r^\theta)$, $\boldsymbol{b}^\theta = (b_0^\theta, b_1^\theta,\dots,b_s^\theta)$ are adjustable parameter vectors; $\boldsymbol{y}(t-r) = (1, y(t-1),\dots, y(t-r))$ is state vector; $\boldsymbol{x}(t-s) = (x(t), x(t-1),\dots, x(t-s))$ is an input sequence; $Y_1^\theta,\dots,Y_r^\theta$, $X_0^\theta,\dots,X_r^\theta$ are fuzzy sets.

By redenoting input variables: $(u_0(t), u_1(t), \dots, u_m(t)) = (1, y(t-1), \dots, y(t-r), x(t),\dots, x(t-s))$, finite difference equation's coefficients: $(c_0^\theta, c_1^\theta,\dots,c_m^\theta) = (a_0^\theta, a_1^\theta,\dots,a_r^\theta, b_1^\theta,\dots,b_s^\theta)$, and membership functions:

$(U_1^\theta(u_1(t)),\dots,U_m^\theta(u_m(t)))=$
$=(Y_1^\theta(y(t-1)),\dots,Y_r^\theta(y(t-r)), X_0^\theta(x(t)),\dots,X_s^\theta(x(t-s)))$, where m=r+s+1,

one obtains the analytic form of the fuzzy model (4), intended for calculating the output $\hat{y}(t)$:

$$\hat{y}(t) = \boldsymbol{c}^T \tilde{\boldsymbol{u}}(t), \tag{8}$$

where $\boldsymbol{c} = (c_0^1,\dots,c_0^n,\dots,c_m^1,\dots,c_m^n)^T$ is the vector of the adjustable parameters;

$\tilde{\boldsymbol{u}}^T(t) = (u_0(t)\beta^1(t),\dots u_0(t)\beta^\rho(t),\dots u_m(t)\beta^1(t),\dots u_m(t)\beta^n(t))$ is the extended input vector;

$$\beta^\theta(t) = \frac{U_1^\theta(u_1(t)) \otimes \dots \otimes U_m^\theta(u_m(t))}{\sum_{\theta=1}^{N}\left(U_1^\theta(u_1(t)) \otimes \dots \otimes U_m^\theta(u_m(t))\right)} \tag{9}$$

is a fuzzy function, where ⊗ denotes the minimization operation or fuzzy product.

If for t = 0, the vector $\boldsymbol{c}(0)$ = 0, the correcting nm×nm matrix $Q(0)$ (m is the number input vectors, n is the number of production rules), and the values of $u(t)$, t = 1 ,…, N are specified, the parameter vector $\boldsymbol{c}(t)$ is calculated using the known multi-step LSM (Buckley, 1993):

$$\boldsymbol{c}(t) = \boldsymbol{c}(t-1) + Q(t)\tilde{\boldsymbol{u}}(t)[y(t) - \boldsymbol{c}^T(t-1)\tilde{\boldsymbol{u}}(t)] \tag{10}$$

$$Q(t) = Q(t-1) - \frac{Q(t-1)\tilde{\boldsymbol{u}}(t)\tilde{\boldsymbol{u}}^T(t)Q(t-1)}{1+\tilde{\boldsymbol{u}}^T(t)Q(t-1)\tilde{\boldsymbol{u}}(t)},$$

$$Q(0) = \gamma I,\ \gamma >> 1, \tag{11}$$

where I is the unit matrix.

The above equations show that even in case of one-dimensional input and few production rules, a lot of observations are needed to apply LSM that makes the fuzzy model too unwieldy. Therefore, only a part of the whole set of rules ($r < n$) should be chosen according to a certain criterion.

The application of the associative search techniques where one or more model parameters are fuzzy, is reduced to such determination of the predicate Ξ={$\Xi_i(R_0^a,R_i^a,T^a)$}, that the number of production rules in the TS model is significantly reduced according to some criterion.

For example, the following matrix:

$$\begin{matrix} \beta_1^{\Theta_t} & \dots & \beta_P^{\Theta_t} \\ \dots & \dots & \dots \\ \beta_1^{\Theta_{t-s}} & \dots & \beta_P^{\Theta_{t-s}} \end{matrix} \tag{12}$$

can be defined for P-dimensional input vectors at time steps t–j, j = 1,…,s. If the rows of this matrix are ranged, say, w.r.t. $\sum_{p=1}^{P}\left|\beta_p^{\Theta_i}\right|$ decrease and a certain number of rows, are selected, then such selection combined with the condition (4) will determine the predicate Ξ and, respectively, the criterion for selecting the images (sets of input vector) from the history.

Let us range the rows of this matrix, for example, subject to the criterion of descending the values $\sum_{p=1}^{P}\left|\beta_p^{\Theta_i}\right|$, and let us select a certain number of rows. Such selection together with the condition (4) defines the predicate Ξ={$\Xi_i(R_0^a,R_i^a,T^a)$), and, respectively, the images selection criterion (sets of input vectors) from the archive.

6. MODELING RESULTS

The following case study illustrates the effectiveness of the proposed method for estimating company's stock price 1 day ahead as against the traditional recurrent algorithms. The comparison of the technique proposed with other known approaches (ANN, kernel estimates, etc.) will be the object of the authors' future research.

To examine stock price movement, several USA companies listed at NYSE (New York Stock Exchange) were considered. The data source was www.yahoo.com.

Assume the model looks as follows:

$$S_N = a_N{}^1 * S_{N-1} + \mathrm{I} * b_N{}^1 + CO_{N-1} * b_N{}^2 + SP_{N-1} * b_N{}^3 + Div_N * b_N{}^4 \tag{13}$$

where S_N is the stock price at the time instant (day) N, S_{N-1} is the stock price at the day N–1; I is a numerical constant equal to 1; CO_{N-1} is crude oil cost at the instant N–1; SP_{N-1} is a normalized value of S&P 500 index on the day N–1; Div_N is the company's dividend on the day N. This input can have only 2 values within the observation period. The dividend announcement date is the date of this input's change. $a_N^1, b_N^1, b_N^2, b_N^3, b_N^4$ are the model's numerical coefficients to be determined. For the identification, the following known algorithm:

$$K_N = K_{N-1} + \frac{Y_N - (K_{N-1}, X_N)}{1 + (X_N, X_N)} \cdot X_N, \qquad (14)$$

where K_N is the parameter vector, Y_N is the output, X_N is the input vector, was applied as well as the virtual models-based one. The graphs in Figures 3 and 4 show the application results of both algorithms for Mittal Steel Company, a major international steel producer with market capitalization of 58 bill. USD.

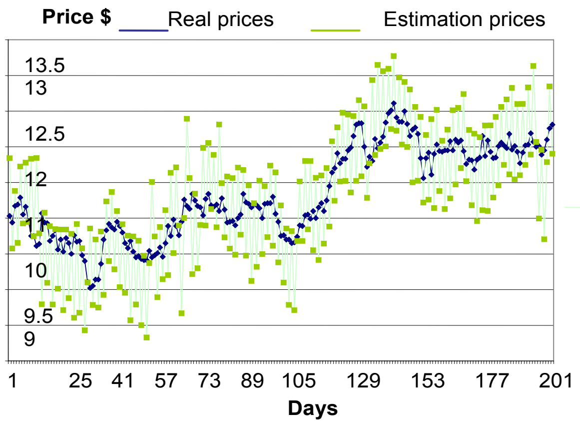

Fig.3. Application result of the algorithm (14)

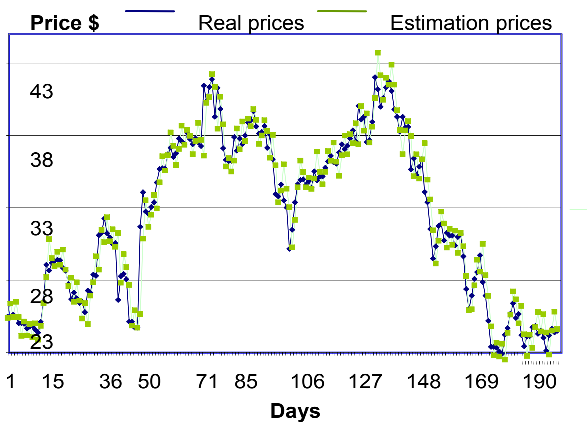

Fig.4. Application result of the virtual models-based algorithm

The error was calculated as:

$$M = \frac{\sum_{i=1}^{N} |S_i - Y_i|}{N},$$ where S_i is the stock price, Y_i is the predicted output, N is the number of days.

The expression for the error has a clear economic meaning: this is the average number of money units lost by the bidder per 1 day at long or short position where the position amount equals 1 stock.

CONCLUSION

For short term prediction in trading, identification algorithms based on associative search procedure can be used. In case of fuzzy definition of one or more variables, successful associative search is possible only with knowledge database built and augmented by analyst or trader during a trading session.

REFERENCES

Shiryaev A.N. (1995). *Probability. 2nd ed.* Springer-Verlag, New York.

Larichev, O. I., Asanov, A., Naryzhny, Y., Strahov, S. (2001). *ESTHER - Expert System for the Diagnostics of Acute Drug Poisonings.* In: Preprints of the 21 SGES International Conference on Knowledge Based Systems and Applied Artificial Intelligence. Pp. 159—168.

Patel V.L., Ramoni M.F. (1997). Cognitive Models of Directional Inference in Expert Medical Reasoning. In: *Expertise in Context: Human and Machine.* (Feltovich P., Ford K., Hofman R. (Ed.)). AAAI Press, Menlo Parc, CA.

Hunt E. (1989). Cognitive Science: Definition, Status and Questions. *Annual Review of Psychology.* Volume no 40, pp. 603-629.

Simon H.A. (1997). The Sciences of the Artificial (3rd ed.). The MIT Press, Cambridge, MA.

Gavrilov A.V. (2002). The Model of Associative Memory of Intelligent System. In: Preprints of the 6-th Russian-Korean International Symposium on Science and Technology. Volume no 1, pp. 174-177.

Chadeev V.M. (2004). Digital Identification of Nonlinear Dynamic Plants. *Automation and Remote Control,* no. 12, pp. 85 –93.

Bakhtadze N., Lototsky V., Maximov E., Pavlov B. (2007). Associative Search Models in Industrial systems. In: Preprints of IFAC Workshop of Intelligent Manufacturing Systems, pp. 156-161.

Takagi T., Sugeno M. (1985). Fuzzy identification of systems and its applications to modeling and control.. *IEEE Trans. Systems Man Cybernetics.*Volume no 15, no.1, pp. 116–132.

Buckley J.J. (1993). Sugeno-type controller are universal controllers. *Fuzzy Sets and Systems*, no. 53, pp. 299-303.