Financial stock Online Forecasting by Observer Kalman filter Identification

Ying-Kai Chen[1],Yu-Lung Wu[2] and Shang-Jen Chuang[3]

[1]Zuoying Branch of Kaohsiung Armed Forces General Hospital
No. 533 Zuoying Rd., Zuoying District, Kaohsiung City,Taiwan
[1]ccyykk0528@yahoo.com.tw

[2] Department of Information Management
I-Shou University
No. 1, Sec. 1, Syuecheng Rd., Dashu District, Kaohsiung City 84001, Taiwan
[2]wuyulung@isu.edu.tw

[3]Department of Electronic Communication Engineering
National Kaohsiung Marine University
No.142, Haijhuan Rd., Nanzih Dist., Kaohsiung City 81157, Taiwan
[3]david@mail.nkmu.edu.tw

Abstract. Generally, stock market indexes are usually non-linear functions, which are unsuitable to predict with linear systems. In this research, The Observer/Kalman filter Identification (OKID) is a linear system, yet it was simulated with nonlinear systems to reach the results of predictions.

This paper uses OKID and Back-Propagation Neural Network(BPNN) as an algorithm to predict the Financial Stock, and analyze their advantages and disadvantages.

Keywords: Prediction, Observer, Kalman filter, Identification.

Introduction

In recent years the economic recovery causes, tradition saving being replaced by financial investment. Stock markets have many advantages including openness of information, higher return on investment and good liquidity. It is a good investment opportunity for people.

However, stock prices are influenced by many reasons. It is a nonlinear system which is difficult to predict. Experts want to predict the market by analytical methods and prediction models, if the accuracy can be improved. It will be easier to be speculating in the stock markets.

OKID method is a Linear Time-Invariant (LTI) Discrete-time state space model **[1]**with Multi-Input and Multi-Output (MIMO)**[2]**, It can identify unknown systems through the Eigen system Realization Algorithm(ERA).

Procedure of OKID algorithm:

1. **Compute the observer Markov parameters.**
2. **Identify both system and observer-gain Markov parameters.**
3. **Realize a state-space model of the system and the corresponding Observer gain by ERA method.**

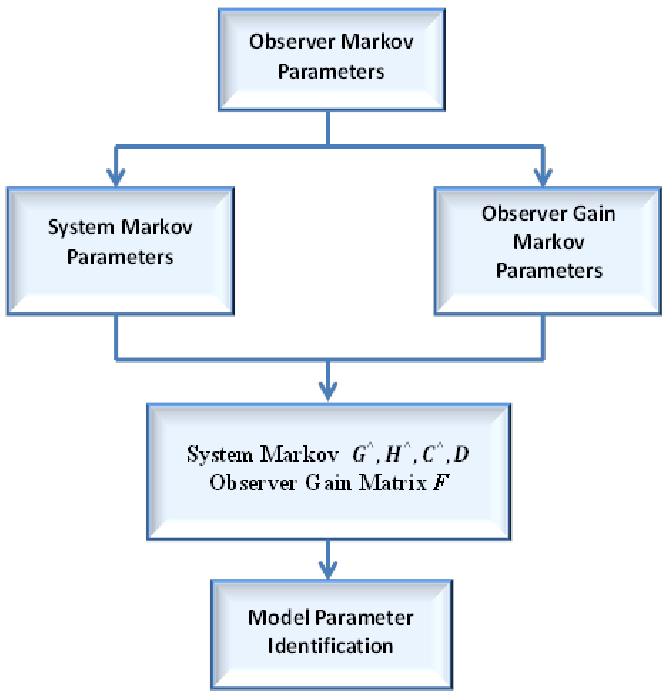


Fig 1. Flowchart for the OKID

We can denote the linear discrete multivariable systems **[3]**by Eq. **(a)** and **(b)**:

$$x(k+1) = \boldsymbol{A}x(k) + \boldsymbol{B}u(k) \quad \textbf{(a)}$$

$$y(k)=\boldsymbol{C}x(k) + \boldsymbol{D}u(k) \quad \textbf{(b)}$$

Setting $x(0) = 0, k = 0,1,2, \ldots, l-1$, the equation can be rewritten in the form

$$x(0) = 0$$
$$y(0) = Du(0)$$
$$x(1) = Bu(0)$$
$$y(1) = CBu(0) + Du(1)$$
$$x(2) = ABu(0) + Bu(1)$$

$$y(2) = CABu(0) + Cbu(1) + Du(2)$$
$$\vdots$$
$$x(l-1) = \sum_{i=1}^{l-1} CA^{i-1} Bu(l-1-i) \quad \textbf{(c)}$$
$$y(l-1) = \sum_{i=1}^{l-1} CA^{i-1} Bu(l-1-i) + Du(l-1) \quad \textbf{(d)}$$

Rewrite the formula to matrix：

$$\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{U} \quad \textbf{(e)}$$

thus

$$\boldsymbol{y} = [y(0) \quad y(1) \quad y(2) \quad \dots \quad y(l-1)]$$
$$\boldsymbol{Y} = [D \quad CB \quad CAB \quad \dots \quad CA^{l-2}B]$$

$$\boldsymbol{U} = \begin{bmatrix} u(0) & u(1) & u(2) & \dots & u(l-1) \\ & u(0) & u(1) & \dots & u(l-2) \\ & & u(0) & \cdots & u(l-3) \\ & & & \ddots & \vdots \\ & & & & u(0) \end{bmatrix}$$

Matrix $\boldsymbol{y}$ is the $rl \times l$ output matrix between input and output, where m is the number of output variables, l is the number of input variables; Matrix $\boldsymbol{Y}$ is the matrix of Markov parameters, r is the number of input variables. Matrix $\boldsymbol{U}$ is upper triangular matrix of $rl \times l$

Using state feedback we can express Eq. (a). as:

$$x(k+1) = Ax(k) + Bu(k) + Gy(k) - Gy(k)$$
$$= (A + GC)x(k) + (B + GD)u(k) - Gy(k) \quad \textbf{(f)}$$
$$x(k+1) = \overline{A}x(k) + \overline{B}v(k) \quad \textbf{(g)}$$

where

$$\overline{A} = A + GC$$
$$\overline{B} = [B + GD - G], \quad v(k) = \begin{bmatrix} u(k) \\ y(k) \end{bmatrix}$$

The matrix of input / output relations can be described by Eq. **(h):**

$$\boldsymbol{y} = \overline{\mathbf{Y}}\mathbf{V} \quad \textbf{(h)}$$

Where

$$\boldsymbol{y} = [y(0) \quad y(1) \quad y(2) \quad \dots \quad y(l-1)]$$

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 02 Issue: 06 | Sep-2015 www.irjet.net p-ISSN: 2395-0072

$$\overline{\mathbf{Y}} = \begin{bmatrix} D & C\overline{B} & C\overline{AB} & \dots & C\overline{A}^{p-1}\overline{B} & \dots & C\overline{A}^{l-2}\overline{B} \end{bmatrix}$$

$$\mathbf{V} = \begin{bmatrix} u(0) & u(1) & u(2) & \dots & u(p) & \dots & u(l-1) \\ & v(0) & v(1) & \dots & v(p-1) & \dots & v(l-2) \\ & & v(0) & \dots & v(p-2) & \dots & v(l-3) \\ & & & \ddots & \vdots & \dots & \vdots \\ & & & & v(0) & \dots & v(l-p-1) \\ & & & & \ddots & & \vdots \\ & & & & & & v(0) \end{bmatrix}$$

When$\geq p$, $\overline{A}^k\overline{B} \approx 0$, using Eq. **(e)**wecan obtain the Observer Markov Parameters $\overline{\boldsymbol{Y}}$

Where

$$\boldsymbol{y} = \overline{\boldsymbol{Y}}\ \boldsymbol{V} \quad \textbf{(i)}$$

Then, we can conclude

$$\boldsymbol{y} = [y(0) \quad y(1) \quad y(2) \quad \dots \quad y(p) \dots \quad y(l-1)]$$

$$\overline{\boldsymbol{Y}} = \begin{bmatrix} D & C\overline{B} & C\overline{AB} & \dots & C\overline{A}^{p-1}\overline{B} \end{bmatrix}$$

$$\boldsymbol{V} = \begin{bmatrix} u(0) & u(1) & u(2) & \dots & u(p) & \dots & u(l-1) \\ & v(0) & v(1) & \dots & v(p-1) & \dots & v(l-2) \\ & & v(0) & \dots & v(p-2) & \dots & v(l-3) \\ & & & \ddots & \vdots & \dots & \vdots \\ & & & & v(0) & \dots & v(l-p-1) \end{bmatrix}$$

Kalman filter.is a time domain filter**[5]**using the optimal recursive data processing algorithm.It is good for solving problems with good efficiency and accuracy.

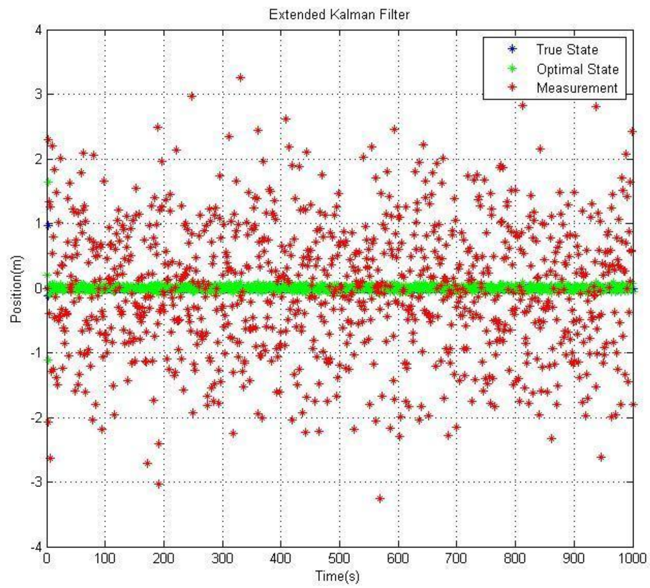


Fig 2.Extended Kalman Filter

Test method. This paper uses two algorithm to predict the highest, lowest and closing prices for **2885.TW and 2890.TW**.The data ranges are collected from **Oct 27th, 2009 to April 20th, 2015**,1308 in total. After removing the noise data, the data were divided into two parts for training and testing; there are 1202 for training data and 106 for testing data. The data from October 27th, 2009 to August 10th, 2014 were used for training, while those from August 11th, 2014 to April 20th, 2015 were used for testing. There are11 inputs, which are the opening, highest, lowest, closing prices, volume of one day before the data day and two days before the data day, and opening price of the data day.

Basal structures. The OKID's initial variable**q is 32**,which is determined by trial and error method, while the Markov parameters matrix size were**96 x 96**.This paper uses Mean Absolute Percentage Error(MAPE) to calculate the error between actual value and predict value.

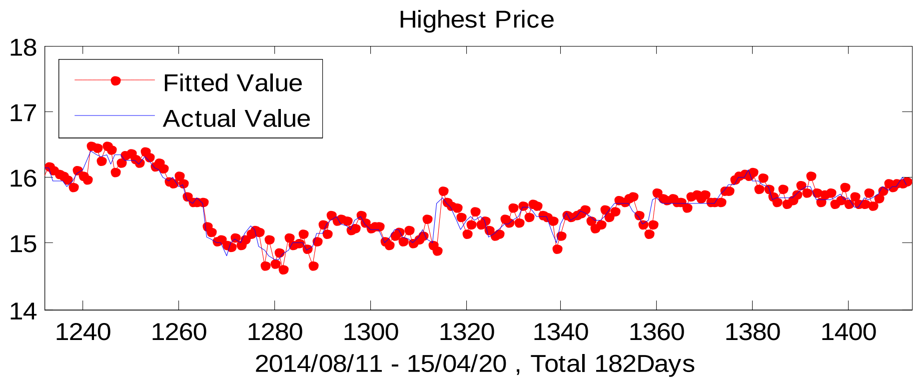
$$\text{MAPE} = \frac{1}{n}\sum_{t=0}^{n}\left|\frac{A_t-F_t}{A_t}\right| \times 100\%$$

A_t: **The actual value.**

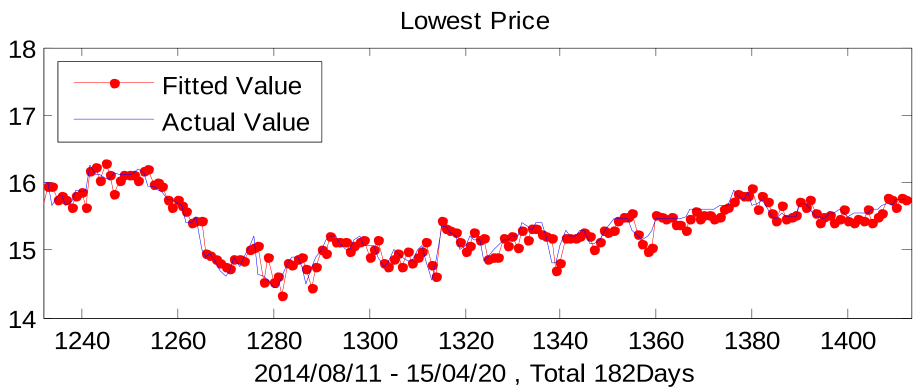
F_t: **The fitted value.**

n : The number of observations.

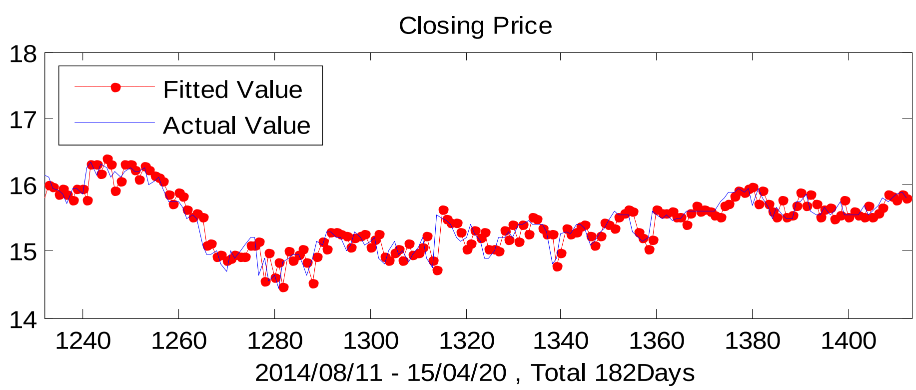
Results and Discussion



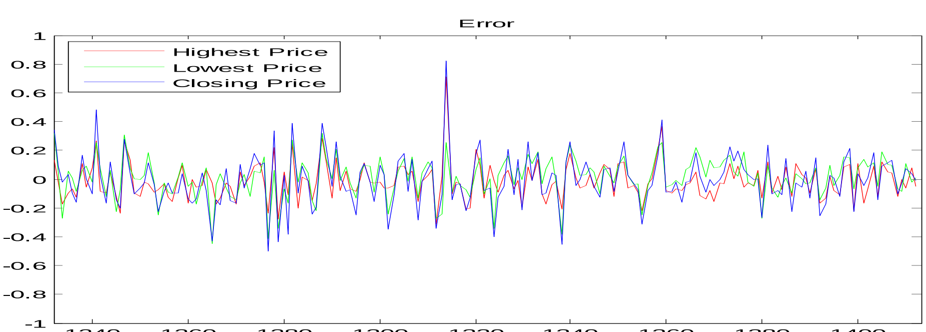
(3-1)



(3-2)



(3-3)



(3-4)

Fig 3. (1)- (3)Comparison Chart of OKID. (4)Error

Date Algorithm		09/10/27-14/08/10 MAPE for Training	14/08/11-15/04/20 MAPE for Testing
OKID	High	**0.9038**	**0.6145**
	Low	**0.9680**	**0.6751**
	Close	**1.2661**	**0.8650**
BPNN	High	1.609	1.938
	Low	1.689	1.997
	Close	1.835	2.336

Fig 4.The MAPE of OKID and BPNN

In the result, the MAPE of OKID for training and testing are less than **1%**.It is an accurate prediction.

Conclusion.

This paper uses OKID and BPNN to predict the FinancialStock Market, and shows that OKID is more accurate than BPNN, but BPNN architecture is easy to design. As a result, we can choose the suitable algorithm based on actual need to solve the problem.

References

[1]J.-N. Juang , M. Phan , L. G. Horta and R. W. Longman "Identification of observer Kalman filter Markov parameters: Theory and experiments", J.Guid., Contr. Dyn., vol. 16, no. 2, pp.320 -329 1993

[2]N. Robertson, K. C. Park and K. F. Alvin "MIMO system structural identification: A revisit via wavelet transforms",Proc. 14th Int. Modal Anal.Conf.,1996

[3]W. T. Hsu "A New Digital Redesign Tracker by Using Observer/Kalman Filter Identification for Unknown Nonlinear MIMO Systems", Master thesis, National Cheng Kung University, 2006

[4]S.Y. Lee"The system modeling and Identification of structure dynamics behavior ",Master thesis, National Central University, 2000

[5]D. G. Luenberger "An introduction to observers", IEEE Trans Autom.Control,vol. AC-16, no. 6, 1971