Expert Systems with Applications 37 (2010) 8590–8598

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Chaos-based support vector regressions for exchange rate forecasting

Shian-Chang Huang [a,*], Pei-Ju Chuang [a], Cheng-Feng Wu [b], Hiuen-Jiun Lai [a]

[a] Department of Business Administration, National Changhua University of Education, Changhua, Taiwan
[b] Department of Business Administration, National Taiwan University, Taipei, Taiwan

ARTICLE INFO

Keywords:
Chaos theory
Hybrid model
Support vector machine
Exchange rate forecasting
Kernel method

ABSTRACT

This study implements a chaos-based model to predict the foreign exchange rates. In the first stage, the delay coordinate embedding is used to reconstruct the unobserved phase space (or state space) of the exchange rate dynamics. The phase space exhibits the inherent essential characteristic of the exchange rate and is suitable for financial modeling and forecasting. In the second stage, kernel predictors such as support vector machines (SVMs) are constructed for forecasting. Compared with traditional neural networks, pure SVMs or chaos-based neural network models, the proposed model performs best. The root-mean-squared forecasting errors are significantly reduced.

1. Introduction

International transactions are usually settled in the near future. Exchange rate forecasting is very important to evaluate the benefit and risk attached to the international business environment. Owing to the high risk associated with the international transactions, exchange rate forecasting is one of the challenging and important fields in modern time series analysis.

The difficulty of forecasting arises from the inherent non-linearity and non-stationarity in exchange rate or financial time series (Cao, 2003; Huang & Wu, 2010). To solve the problem, this study develops a new forecasting strategy that employs the phase space reconstruction from chaos theory and support vector regression from kernel methods to extract the above financial characteristics for making a good prediction. Namely, this study will develop a chaos-based nonparametric model to predict the exchange rate's future behavior.

For forecasting strategies, Box and Jenkins' Auto-Regressive Integrated Moving Average (ARIMA) technique has been widely used for time series forecasting. However, ARIMA is a general univariate model and it is developed based on the assumption that the time series being forecasted are linear and stationary, usually not satisfied for financial data. In recent years, neural networks (NN) has found useful applications in financial time series forecasting, including Hill, O'Connor, and Remus (1996), Kamruzzaman and Sarker (2003), Wang and Leu (1996), Yao and Tan (2000), Zhang and Hu (1998), and Zimmermann, Neuneier, and Grothmann (2001). Recently, the support vector machine (SVM) method (e.g., Cristianini & Shawe-Taylor, 2000; Schoelkopf, Burges, & Smola, 1999; Vapnik, 1995; Wang, 2005), another form of neural networks, has been gaining popularity and has been regarded as the state-of-the-art technique for regression and classification applications. It is believed that the formulation of SVM embodies the structural risk minimization principle, thus combining excellent generalization properties with a sparse model representation. Recent applications of SVM in financial forecasting include: Cao (2003), Chang and Tsai (2008), Huang (2008), Huang and Wu (2008), Kim (2003), and Ince and Trafalis (2005).

Chaos theory is relatively new in science. It is only very recently, in the late 1980s, that interest in chaos theory as a financial analysis tool has emerged. This is so since the theory offers a new way by which the behavior of financial markets can be predicted (at least, over the short-term) and that this non-linear, financial model goes beyond statistics for it can reveal hidden patterns and trends in financial data which could not be captured by conventional statistical techniques. Scheindman and Lebaron (1989) and Frank and Stengos (1988) have found the chaotic behavior in financial market such as stock market, foreign exchange markets and futures market. Chaotic time series prediction becomes an extremely important research area and obtains widespread application (Liu, Dong, & Chen 2007; Shangfei & Peiling, 1998; Xu & Xiong, 2003). Recent applications of chaos theory in financial markets include: Federici and Gandolfo (2002), Kumar, Tan, and Ghosh (1999), Ma and Xu (2007), Pavlidis, Tasoulis, and Vrahatis (2005), and Torkamani, Mahmoodzadeh, Pourroostaei, and Lucas (2007).

For time series feature extraction, using chaos theory a chaotic attractor may be obtained by measuring the chaotic exchange rate series. The properties of the chaotic attractor can be retained

* Corresponding author. Address: Department of Business Administration, College of Management, National Changhua University of Education, No. 2, Shi-Da Road, Changhua 500, Taiwan. Tel.: +886 4 7232105x7420, +886 953092968 (mobile); fax: +886 4 7211292.
E-mail address: shhuang@cc.ncue.edu.tw (S.-C. Huang).

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.06.001

through a reconstruction procedure. This procedure is known as the delay coordinate embedding (Takens, 1981) resulting in a reconstructed phase space (or state space) which contains a reconstructed chaotic attractor preserving both geometrical and dynamical properties of the original chaotic attractor. In forecasting, one can only observe the exchange rate series, which are unable to exhibit the inherent essential character of the exchange rate dynamic system. The phase space reconstruction provides us a mean to study the unobserved variables of the system, and thus is suitable for financial forecasting.

The major innovation of this paper lies in combing the phase space reconstruction with kernel regressors for exchange rate forecasting. In the first stage, the delay coordinate embedding transforms the input space (raw data) to a feature space (or state space) suitable for financial modeling and forecasting. In the second stage of the new method, kernel regressors that best fit the transformed series are constructed for final forecasting. Compared with neural networks, pure SVMs or chaos-based neural network models, the proposed model performs best. The root-mean-squared forecasting errors are significantly reduced.

The remainder of the paper is organized as follows. Section 2 introduces the new prediction model, including the chaos theory, delay coordinate embedding, and the support vector regression. Section 3 describes the data used in the study, and discusses the experimental findings. Conclusions are given in Section 4.

2. Chaos-based support vector forecasting models

We first introduce the chaos theory and the delay coordinate embedding for phase space reconstruction, and then the methodology of support vector regression for financial forecasting.

2.1. Chaos theory

A chaotic system reveals a relatively complex behavior through the dynamic of non-linear system. The orbits of the system attract to a complex higher-dimensional subset called a strange attractor. The chaotic system is sensitivity to initial conditions; points that are arbitrarily close initially become exponentially further apart with increasing time, leading to the amplification of very small perturbations into global uncertainties. The importance of studying chaotic behavior lies in the fact that chaotic behavior is much more widespread, and may even be the norm in the real world, especially in financial markets.

2.2. Delay coordinate embedding

Takens embedding theorem (Takens, 1981) provides theoretical foundation for the analysis of time series generated by non-linear dynamical systems. Later Sauer, Yorke, and Casdagli (1991) show a phase space can be reconstructed from a univariate chaotic time series. Let an univariate time series $\{x_i\}_{i=1}^{N}$, where N is the length of the time series, is generated from a d-dimension chaotic attractor, a phase space R^d of the attractor can be reconstructed by using delay coordinate defined as

$$\mathbf{Z}_i = \left(x_i, x_{i-\tau}, \ldots, x_{i-(m-1)\tau}\right), \tag{1}$$

where m is known as the embedding dimension of reconstructed phase space, τ is the delay constant. The delay coordinate method is currently the most widely used choice for chaotic time series analysis. It has one good property that the signal to noise ratio on each component is equal.

If we find the embedding dimension m, according to the method of phase space reconstruction, we can reconstruct an m-dimensional phase space from the time series $\{x_t\}$. Let the vector of the phase space marked as $\mathbf{Z}_i = (x_i, x_{i-\tau}, \ldots, x_{i-(m-1)\tau})$. And there exists a smooth function $F{:}R^m \to R^m$, satisfying $F(\mathbf{Z}_t) = \mathbf{Z}_{t+1}$. We can infer from F a function $f{:}R^m \to R$, satisfying $x_{t+1} = f(\mathbf{Z}_t)$. As a result, choosing the embedding dimension m is very important, with which we can predict x_{t+1} from the history data. Takens considered that the sufficient condition for embedding dimension is $m \geqslant 2d + 1$. However, too large embedding dimension needs more observations and complex computation. Moreover, if m is too large, chaotic data add redundancy and degrade the performance of many algorithms. As a result if the original attractor has dimension d, then an embedding dimension of $m = 2d + 1$ will be adequate for reconstructing the attractor.

For the estimation of delay constant τ, if τ is too small, each coordinate is almost the same and the trajectories of the reconstructed space are squeezed along an identity line; this phenomenon is known as redundancy. If τ is too large, in the presence of chaos and noise, the dynamics at any one time become effectively and causally disconnected from the dynamics at a later time, so that even simple geometric objects look extremely complicated; this phenomenon is known as irrelevance.

2.3. Estimate τ and m

To reconstruct the phase space, good choices for time lag τ and embedding dimension m are needed. This study uses the first minimum of the mutual information function (Abarbanel, 1996) $I(\tau)$ to determine τ:

$$I(\tau) = \sum_{n=1}^{N-\tau} P(x_n, x_{n+\tau}) \log_2\left(\frac{P(x_n, x_{n+\tau})}{P(x_n)P(x_{n+\tau})}\right), \tag{2}$$

where $P(x_n)$ is the probability density of x_n, while the $P(x_n, x_{n+\tau})$ is the probability density of x_n and $x_{n+\tau}$.

The false nearest neighbor (FNN, Kennel, Brown, & Abarbanel, 1992) method is a approach to find the optimal embedding dimension. Their idea is quite intuitive. Suppose the minimal embedding dimension for a given time series is m_0. This means that in a m_0-dimensional delay space the reconstructed attractor is a one-to-one image of the attractor in the original phase space. Especially, the topological properties are preserved. Thus the neighbors of a given point are mapped onto neighbors in the delay space. Due to the assumed smoothness of the dynamics, neighborhoods of the points are mapped onto neighborhoods again. Of course the shape and the diameter of the neighborhoods is changed according to the Lyapunov exponents. But suppose now you embed in an m-dimensional space with $m < m_0$. Due to this projection the topological structure is no longer preserved. Points are projected into neighborhoods of other points to which they would not belong in higher dimensions. These points are called *false neighbors*. If now the dynamics is applied, these false neighbors are not typically mapped into the image of the neighborhood, but somewhere else, so that the average "diameter" becomes quite large.

The idea of the algorithm false nearest is the following. For each point $\mathbf{Z}_i$ in the time series look for its nearest neighbor $\mathbf{Z}_j$ in an m-dimensional space. Calculate the distance $\|\mathbf{Z}_i - \mathbf{Z}_j\|$. Iterate both points and compute

$$R_i = \frac{\left\|\mathbf{Z}_{i+1} - \mathbf{Z}_{j+1}\right\|}{\left\|\mathbf{Z}_i - \mathbf{Z}_j\right\|}. \tag{3}$$

If R_i exceeds a given heuristic threshold R_t, this point is marked as having a false nearest neighbor. The criterion that the embedding dimension is high enough is that the fraction of points for which $R_i > R_t$ is zero, or at least sufficiently small.