Expert Systems with Applications 37 (2010) 5430–5435

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A fuzzy support vector regression model for business cycle predictions

Kuo-Ping Lin [a,*], Ping-Feng Pai [b,c]

[a] Department of Information Management, Lunghwa University of Science and Technology, Taoyuan 333, Taiwan
[b] Department of Information Management, National Chi Nan University, University Rd., Puli, Nantou 545, Taiwan
[c] Department of International Business Studies, National Chi Nan University, University Rd., Puli, Nantou 545, Taiwan

ARTICLE INFO

Keywords:
Business cycle
Fuzzy set theory
Support vector regression

ABSTRACT

Business cycle predictions face various sources of uncertainty and imprecision. The uncertainty is usually linguistically determined by the beliefs of decision makers. Thus, the fuzzy set theory is ideally suited to depict vague and uncertain features of business cycle predictions. Consequently, the estimation of fuzzy upper and lower bounds become an essential issue in predicting business cycles in an uncertain environment. The support vector regression (SVR) model is a novel forecasting approach that has been successfully used to solve time series problems. However, the SVR approach has not been widely applied in fuzzy forecasting problems. This study employs support vector regressions to calculate fuzzy upper and lower bounds; and presents a fuzzy support vector regression (FSVR) model for forecasting indices of business cycles. A numerical example of a business cycle prediction in Taiwan was used to demonstrate the forecasting performance of the FSVR model. The empirical results are satisfactory. Therefore, the FSVR model is an effective alternative in forecasting business cycles under uncertain circumstances.

1. Introduction

Accuracy in forecasting business cycles is an important issue in economic study, and statistical methods have usually been employed to analyze them. Many investigations have been done in the analysis of business cycles (Banerji & Hiris, 2001; Layton, 1996, 1998; Seip & McNown, 2007; Wu & Tseng, 2002; Yang & Kim, 2005). However, business cycles are often determined by a panel of macroeconomic experts, and thus, it is difficult to predict the index of business cycles. The difficulty arises from assumptions made from the probability distributions and business cycle data, which are usually vague. The index of business cycles in Taiwan is composed of nine exogenous variables, and five lights are used to represent different economic activities. The five lights include some uncertain factors in predicting business cycles. Hence, the fuzzy set theory (Zadeh, 1965) is a proper approach to analyze Taiwan business cycles.

Unlike most of traditional technologies SVR (Vapnik, Golowich, & Smola, 1996) implementing neural network models, SVR adopts a structural risk minimization principle, which seeks to minimize the upper bounds of the generalization error rather than minimize the training error. In recent years, SVR schemes have been extended to cope with forecasting problems, and have provided many promising results in customer demand (Levis & Papageorgiou, 2005), finance (Huang, Nakamori, & Wang, 2005; Kim, 2003; Tay & Cao, 2002), intermittent demand (Hua & Zhang, 2006), tourism demand (Pai & Hong, 2005), air quality (Lu & Wang, 2005), wind speed (Mohandes, Halawani, Rehman, & Hussain, 2004), plant control systems (Xi, Poo, & Chou, 2007), rainfall (Hong & Pai, 2007), prices for the electricity market (Gaoa, Bompard, Napoli, & Cheng, 2007), and flood control (Yu, Chen, & Chang, 2006). Hong and Hwang (2003) proposed a support vector fuzzy regression machine model for modifying convex optimization problems of multivariate fuzzy linear regression models. Empirical results indicate that the developed model derives satisfying solutions efficiently. Jeng, Chuang, and Su (2003) developed a support vector interval regression network to efficiently handle interval output data. Yao and Yu (2006) developed a fuzzy regression based on asymmetric support vector machines, which overcome limitations of traditional nonlinear fuzzy regression, and can be effectively used for parameter estimation. Chuang (2008) presented an interval support vector regression network model, which can handle interval input and output data. Hao and Chiang (2008) developed a fuzzy regression analysis model based on support vector learning techniques, and suggested that the developed model can perform automatic and accurate control in fuzzy regression analysis tasks.

In this study, a fuzzy support vector regression model is presented to forecast an index of business cycles. Support vector regression was used to calculate fuzzy upper and lower bounds, and then make predictions by fuzzy *H*-level set (*H*-cut). In addition, genetic algorithms (GA) were employed to select three parameters of SVR models. The remainder of this paper is organized as follows.

* Corresponding author.
E-mail addresses: kplin@mail.lhu.edu.tw (K.-P. Lin), paipf@ncnu.edu.tw (P.-F. Pai).

0957-4174/$ - see front matter Crown Copyright © 2010 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.02.071

A brief introduction of the theory of SVR is given in Section 2. The fuzzy support vector regression model is derived in Section 3. A numerical example of business cycle predictions and empirical results are presented in Section 4. Some concluding remarks are offered in Section 5.

2. Support vector regression models

Presented by Vapnik (1995), the support vector machine was originally applied to pattern recognition problems. Then, the support vector machine model has been successfully extended for dealing with nonlinear regression problems (Vapnik et al., 1996). The support vector regression model is based on the idea of mapping the original data x nonlinearly into a higher dimensional feature space. The SVR approach is to approximate an unknown function by a training data set $\{(x_i, Y_i), i = 1, \ldots, N\}$. The regression function can be formulated as follows:

$$F = w\phi(x_i) + b, \tag{1}$$

where $\phi(x_i)$ denotes the feature of the inputs, and w and b indicate coefficients. The coefficients (w_i and b) are estimated by minimizing the following regularized risk function.

$$R(F) = C\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(Y_i, F_i) + \frac{1}{2}\|w\|^2, \tag{2}$$

where

$$L_\varepsilon(Y_i, F_i) = \begin{cases} 0 & \text{if } |Y_i - F_i| \leqslant \varepsilon \\ |Y_i - F_i| - \varepsilon & \text{otherwise} \end{cases}, \tag{3}$$

where C and ε are user-defined parameters. The parameter ε is the difference between actual values and values calculated from the regression function. This difference can be viewed as a tube around the regression function. The points outside the tube are regarded as training errors. In Eq. (2), $L_\varepsilon(Y_i, F_i)$ is called an ε-insensitive loss function, and can be illustrated as Fig. 1.

The loss equals zero if the approximate value is within the ε-tube. Additionally, the second item of Eq. (2), $\frac{1}{2}\|w\|^2$, is adopted to estimate the flatness of a function which can avoid overfitting. Therefore, C indicates a parameter determining the trade-off between the empirical risk and the model flatness. Two positive slack variables (ξ_i and ξ_i^*), representing the distance from actual values to the corresponding boundary values of the ε-tube, are then introduced. These two slack variables equal zero when the data points fall within the ε-tube. Eq. (2) is then reformulated into the following constrained form:

$$\begin{aligned} &Min \quad f(w, \xi, \xi^*) = \frac{1}{2}\|w\|^2 + C\left(\sum_{i=1}^{N}(\xi_i + \xi_i^*)\right) \\ &Subjective\ to\ w\phi(x_i) + b - Y_i \leqslant \varepsilon + \xi_i^*, \quad i = 1, 2, \ldots, N, \\ &\qquad Y_i - w\phi(x_i) - b \leqslant \varepsilon + \xi_i, \quad i = 1, 2, \ldots, N, \\ &\qquad \xi_i, \xi_i^* \geqslant 0, \quad i = 1, 2, \ldots, N. \end{aligned} \tag{4}$$

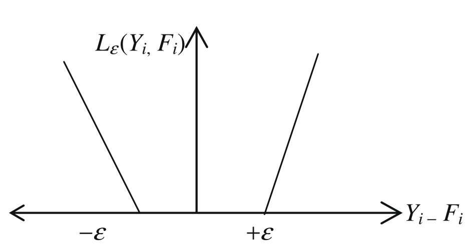

Fig. 1. The ε-insensitive loss function.

This constrained optimization problem can be solved using the following primal Lagrangian form:

$$\begin{aligned} Min\ &\frac{1}{2}\|w\|^2 + C\left(\sum_{i=1}^{N}(\xi_i + \xi_i^*)\right) - \sum_{i=1}^{N}\beta_i[w\phi(x_i) + b - Y_i + \varepsilon + \xi_i] \\ &- \sum_{i=1}^{N}\beta_i^*[Y_i - w\phi(x_i) - b + \varepsilon + \xi_i^*] - \sum_{i=1}^{N}(\alpha_i\xi_i + \alpha_i^*\xi_i^*). \end{aligned} \tag{5}$$

Eq. (5) is minimized with respect to primal variables w, b, ξ, and ξ^*, and is maximized with regard to non-negative Lagrangian multipliers α_i, α_i^*, β_i, and β_i^*. Finally, Karush–Kuhn–Tucker conditions are applied to Eq. (4), and the dual Lagrangian form given by Eq. (6).

$$\begin{aligned} &Max \quad \sum_{i=1}^{N} Y_i(\beta_i - \beta_i^*) - \varepsilon\sum_{i=1}^{N}(\beta_i + \beta_i^*) - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\beta_i - \beta_i^*)(\beta_j - \beta_j^*)K(x_i, x_j) \\ &Subjective\ to\ \sum_{i=1}^{N}(\beta_i - \beta_i^*) = 0, \\ &\qquad f0 \leqslant \beta_i \leqslant C, \quad i = 1, 2, \ldots, N, \\ &\qquad 0 \leqslant \beta_i^* \leqslant C, \quad i = 1, 2, \ldots, N. \end{aligned} \tag{6}$$

The Lagrange multipliers in Eq. (6) satisfy the equality $\beta_i * \beta_i^* = 0$. The Lagrange multipliers, β_i and β_i^*, are determined, and an optimal weight vector of the regression hyperplane is written by Eq. (7).

$$w^* = \sum_{i=1}^{N}(\beta_i - \beta_i^*)K(x, x_i). \tag{7}$$

Thus, the regression function is given by:

$$F(x, \beta, \beta^*) = \sum_{i=1}^{N}(\beta_i - \beta_i^*)K(x, x_i) + b. \tag{8}$$

Herein, $K(x_i, x_j)$ denotes a Kernel function whose value equals the inner product of two vectors, x_i and x_j, in the feature space $\phi(x_i)$ and $\phi(x_j)$, meaning that $K(x_i, x_j) = \phi(x_i) * \phi(x_j)$. Any function that satisfies Mercer's condition (Mercer, 1909) can act as the Kernel function. This work uses the Gaussian function.

3. A fuzzy support vector regression model

Introduced by Zadeh (1965), fuzzy set theory has been applied to deal with uncertain problems in many fields. A brief of fuzzy set theory is depicted as follows. A membership function is defined for all elements a in the referential set U. A fuzzy set (A) can be defined by the membership functions $\mu_A(a)$, which can have values of [0, 1]. Thus, a fuzzy set (A) is said to be convex if all ordinary subsets of A are convex. A fuzzy set is normalized if $\exists a \in U$, $\mu_A(a) = 1$. For an interval, level of confidence, or called α-cut or α-level set at level α (0, 1], an ordinary subset of A can be defined and denoted as $[A]_\alpha$,

$$[A]_\alpha = \{a \in U | \mu_A(a) \geqslant \alpha\}, \quad \alpha \in (0, 1]. \tag{9}$$

Furthermore, a fuzzy number (FN) can be defined as a convex, a normalized fuzzy set on a real line, with an upper semi-continuous membership function and bounded support. Dubois and Prade (1980) have defined a general representation form for FNs, which can be called the L–R type FNs. Here, FNs are represented as follows:

Definition 1 (*The bound form representation*). A symmetrical FN can be written as $A = (\psi - c, \psi, \psi + c)_{LR}$, where $\psi - c$ and $\psi + c$, respectively, denote the lower and upper bounds, ψ the mode or center value, $(\psi - c, \psi + c)$ forms the support, and L, R, respectively, denote the left and right reference (or shape) functions of A. In the case $L \equiv R$, and A has the membership function: