Expert Systems with Applications 42 (2015) 2121–2135

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A high-order multi-variable Fuzzy Time Series forecasting algorithm based on fuzzy clustering

S. Askari, N. Montazerin *

Mechanical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), 424 Hafez Avenue, Tehran 1591634311, Iran

ARTICLE INFO

Article history:
Available online 6 October 2014

Keywords:
Fuzzy sets
Fuzzy Time Series
Fuzzy clustering
Fuzzy C-Means
Multi-variable algorithms
High-order algorithms

ABSTRACT

A High–order algorithm for Multi-Variable Fuzzy Time Series (HMV-FTS) is presented based on fuzzy clustering to eliminate some well-known problems with the existing FTS algorithms. High-order algorithms can handle only one-variable FTS and multi-variable algorithms can handle only one-order FTS. HMV-FTS does both tasks simultaneously. FTS algorithms cannot incorporate existing information about future value of a variable in the forecasting process while HMV-FTS can. Defuzzification of the fuzzy value of a forecast to cluster centers or midpoint of intervals and use of intervals are other controversial problems with the existing FTS algorithms. These are eliminated by constructing fuzzy sets from partition matrices and letting each data point to contribute in defuzzification based on its membership grade in the fuzzy sets. In multi-variable FTS algorithms, one variable is considered as main variable which is forecasted and the other variables are secondary; while HMV-FTS treats all variables equally and more than one variable can be forecasted at the same time. It is shown that HMV-FTS is suitable for system identification, forecasting and interpolation. This algorithm is more accurate than popular FTS algorithms and other forecasting tools and systems such as ANFIS, Type II fuzzy model and ARIMA model.

1. Introduction

Fuzzy Time Series (FTS) is first introduced in Song and Chissom (1993a, 1993b, 1994) as a forecasting tool for enrollments. Since then, it has been subject of numerous researches in the area of forecasting especially when the available data is imprecise, vague and without identifiable trend. Initial structure of FTS is highly complex and computationally expensive. It is changed into a more efficient model in Chen (1996) which is generally accepted by researchers and is the common form of FTS. This model is used in many forecasting problems including, shipping index (Duru, 2010), pollution (Domanska & Wojtylak, 2012), university enrollments (Aladag, Basaran, Egrioglu, Yolcu, & Uslu, 2008), rice production (Singh, 2007a), stock exchange (Huarng & Yu, 2006a; Huarng, Yu, & Hsu, 2007; Park, Lee, Song, & Chun, 2010; Qiu, Liu, & Wang, 2012; Wong, Tu, & Wang, 2010), etc.

FTS algorithms usually employ intervals of universe of discourse to construct fuzzy sets. Appropriate intervals and their lengths are always challenging problems and researchers attempt to improve accuracy of FTS by selecting proper intervals and adjustment of their lengths. For example, Particle Swarm Optimization (PSO) is used to find proper intervals and adjust interval lengths (Huang et al., 2011; Kuo et al., 2009), Tabu Search and fuzzy inference systems are used to find length of intervals (Avazbeigi, Hashemi Doulabi, & Karimi, 2010). Some other techniques for determining best intervals and interval lengths are found in (Egrioglu, Aladag, Yolcu, Uslu, & Basaran, 2010; Wang, Liu, & Pedrycz, 2013). Some FTS algorithms are based on fuzzy clustering in which no interval is used and instead the data is fuzzified to the cluster centers (Bulut, Duru, & Yoshida, 2012; Chen & Tanuwijaya, 2011; Cheng, Cheng, & Wang, 2008; Egrioglu, Aladag, Yolcu, Uslu, & Erilli, 2011; Li, Kuo, Cheng, & Chen, 2010, 2008). The most important advantage of these algorithms over interval based algorithms is that no interval is required. An algorithm is presented to handle high order FTS using consecutive differences of the algorithm parameters instead of fuzzy clustering and intervals (Singh, 2007a, 2007b, 2008, 2009).

There are two important issues, namely, number of variables in FTS and order of FTS algorithm that are always controversial and no definite solution is known for them. Multi-variable algorithms are presented for the former (Bulut et al., 2012; Chen & Tanuwijaya, 2011; Cheng et al., 2008; Egrioglu, Aladag, Yolcu, Basaran, & Uslu, 2009; Huarng et al., 2007; Yu & Huarng, 2008); and high-order algorithms are proposed for the latter (Aladag et al., 2008; Chen, 2002; Chen, Cheng, & Teoh, 2008; Egrioglu

* Corresponding author.
E-mail addresses: mntzrn@aut.ac.ir, mntzrn@hotmail.com (N. Montazerin).

http://dx.doi.org/10.1016/j.eswa.2014.09.036
0957-4174/© 2014 Elsevier Ltd. All rights reserved.

et al., 2009, 2010; Park et al., 2010; Singh, 2009; Teoh, Chen, Cheng, & Chu, 2009).

High-order algorithms handle only one-variable FTS and multi-variable algorithms handle only one-order FTS. The objective of this paper is to present a High-order Multi-Variable algorithm for Fuzzy Time Series (HMV-FTS). Problem of intervals and their lengths are still studied and their roles are not fully understood. We use fuzzy clustering algorithms to construct fuzzy sets of HMV-FTS and avoid intervals and their lengths. In interval based FTS algorithms, mid-point of the interval is considered as defuzzified value of fuzzy forecast and other members of the interval play no role in defuzzification. In clustering based FTS algorithms, fuzzy forecast is defuzzified to the cluster centers without any contribution of members of universe of discourse and their membership grades in the cluster. We use defuzzification of the fuzzy set instead of cluster centers to get crisp value of forecast. In this approach, each member of the universal set of each variable contributes to defuzzification depending on its membership grade in the cluster.

Rest of the paper is organized as follows: Basic definitions and algorithms of FTS are discussed in Section 2. In Section 3, HMV-FTS algorithm is presented. Examples of Identification of systems and forecasting of processes are presented in Sections 4 and concluding remarks are given in Section 5.

2. FTS definitions and algorithms

2.1. FTS definitions

Definition 1. Let $Y_j(t) \in \Re,\ 1 \leqslant t \leqslant N,\ 1 \leqslant j \leqslant m$ (N is number of data vectors and m is number of dependent and independent variables together) be the universe of discourse of the jth variable on which fuzzy sets $A_{j,i}(t)$, $1 \leqslant i \leqslant c_j$ (c_j is number of fuzzy sets of the jth variable) are defined and $F(t)$ be a collection of $A_{j,i}(t)$s, then $F(t)$ is defined as a Fuzzy Time Series on $Y_j(t)$s. In general, $F(t)$ is a linguistic variable with linguistic values, $A_{j,i}(t)$.

Definition 2. If $F(t)$ is caused by $F(t-1)$, the Fuzzy Logical Relationship (FLR) between them is represented by $F(t-1) \to F(t)$ which is a first order FLR. This relation can also be written as $F(t) = F(t-1) \circ R(t-1,t)$, where $\circ$ is composition operator and $R(t-1,t)$ is a fuzzy relationship. In this FLR, $F(t-1)$ and $F(t)$ are called current state and next state which are denoted by A_{j,i_1} and A_{j,i_2}, respectively. Similarly, an n-order FLR, $F(t-n)$, $F(t-n+1), \ldots, F(t-2), F(t-1) \to F(t)$ is shown by

$$A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n} \to A_{j,i_{n+1}} \tag{1}$$

where,

$$A_{j,i_1} = F(t-n), A_{j,i_2} = F(t-n+1), \ldots, A_{j,i_{n-1}} = F(t-2),$$
$$A_{j,i_n} = F(t-1) \cdot A_{j,i_{n+1}} = F(t) \tag{2}$$

Usually, first and last FLRs of an FTS are defined as $\Phi, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n} \to A_{j,i_{n+1}}$ and $A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n} \to \Phi$, where Φ is a null value.

Definition 3. FLRs with identical current states are grouped into a Fuzzy Logical Relationship Group (FLRG). For instance:

$$\begin{cases} A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n} \to A_{j,q_1} \\ A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n} \to A_{j,q_2} \\ \vdots \\ A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n} \to A_{j,q_r} \end{cases} \Rightarrow A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n} \to A_{j,q_1}, A_{j,q_2}, \ldots, A_{j,q_r} \tag{3}$$

where, $A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n} \to A_{j,q_1}, A_{j,q_2}, \ldots, A_{j,q_r}$ is a n-order FLRG.

2.2. FTS basic algorithms

FTS algorithms are generally composed of the following steps (Chen, 1996). Each variable is handled individually in both interval and clustering based algorithms. The algorithm takes training data set including input/output data as input and yields forecasted values of the dependent variable as output which is usually last row of the data set.

I. Definition of the universe of discourse for jth variable, $\Lambda_j = [L_{\min,j} - L_{1,j}, L_{\max,j} + L_{2,j}]$, where $L_{\min,j}$ and $L_{\max,j}$ are minimum and maximum of the jth variable and $L_{1,j}$ and $L_{2,j}$ are the lowest positive values chosen such that Λ_j can be divided into the number of desired intervals.

II. Partitioning Λ_j into c_j intervals of equal lengths, $\lambda_{j,i}$, $1 \leqslant i \leqslant c_j$, such that $\cup_{i=1}^{c_j} \lambda_{j,i} = \Lambda_j$.

III. Definition of fuzzy sets $A_{j,i}$, $1 \leqslant i \leqslant c_j$ for the jth variable as in Definition 1 based on the above intervals. Fuzzy sets are defined as:

$$\begin{aligned} A_{j,1} &= \frac{a_{11}}{\lambda_{j,1}} + \frac{a_{12}}{\lambda_{j,2}} + \ldots + \frac{a_{1c_j}}{\lambda_{j,c_j}} \\ A_{j,2} &= \frac{a_{21}}{\lambda_{j,1}} + \frac{a_{22}}{\lambda_{j,2}} + \ldots + \frac{a_{2c_j}}{\lambda_{j,c_j}} \\ &\vdots \\ A_{j,c_j} &= \frac{a_{c_j1}}{\lambda_{j,1}} + \frac{a_{c_j2}}{\lambda_{j,2}} + \ldots + \frac{a_{c_jc_j}}{\lambda_{j,c_j}} \end{aligned} \tag{4}$$

c_j is usually taken as 7. In the literature, these fuzzy sets are defined as:

$$\begin{aligned} A_{j,1} &= \frac{1}{\lambda_{j,1}} + \frac{0.5}{\lambda_{j,2}} + \frac{0}{\lambda_{j,3}} + \ldots + \frac{0}{\lambda_{j,c_j}} \\ A_{j,2} &= \frac{0.5}{\lambda_{j,1}} + \frac{1}{\lambda_{j,2}} + \frac{0.5}{\lambda_{j,3}} + \frac{0}{\lambda_{j,4}} + \ldots + \frac{0}{\lambda_{j,c_j}} \\ &\vdots \\ A_{j,i} &= \frac{0}{\lambda_{j,1}} + \frac{0}{\lambda_{j,2}} + \ldots + \frac{0}{\lambda_{j,i-2}} + \frac{0.5}{\lambda_{j,i-1}} + \frac{1}{\lambda_{j,i}} + \frac{0.5}{\lambda_{j,i+1}} + \frac{0}{\lambda_{j,i+2}} + \ldots + \frac{0}{\lambda_{j,c_j}} \\ &\vdots \\ A_{j,c_j} &= \frac{0}{\lambda_{j,1}} + \frac{0}{\lambda_{j,2}} + \ldots + \frac{0}{\lambda_{j,c_j-2}} + \frac{0.5}{\lambda_{j,c_j-1}} + \frac{1}{\lambda_{j,c_j}} \end{aligned} \tag{5}$$

IV. Fuzzification of the data: If a given value in the jth row of the data, $Y_{j,t}$, belongs to the interval $\lambda_{j,i}$, $Y_{j,t} \in \lambda_{j,i}$, it is fuzzified to the fuzzy set $A_{j,i}$.

V. Construction of the FLRs as described in Definition 2 and FLRGs as in Definition 3.

VI. Finding current state of FLR of the forecast time, $A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n}$.

VII. Finding fuzzy value(s) of the forecast: FLRG with current state identical to that of forecast FLR (all with the same $A_{j,i_1}, A_{j,i_2}, \ldots, A_{j,i_{n-1}}, A_{j,i_n}$ of step VI) is found. Then, next state of this FLRG, $\Sigma = \{A_{j,q_1}, A_{j,q_2}, \ldots, A_{j,q_r}\}$ is considered as the fuzzy value(s) of the forecast. Note that sometimes Σ may be an empty set. Finding crisp value of forecast with an empty Σ will be explained later in the FTS algorithms.

VIII. Defuzzification of the fuzzy value(s) of the forecast: Fuzzy value(s) of the forecast, Σ, are defuzzified into the crisp