Information Sciences 257 (2014) 127–137

Contents lists available at ScienceDirect

Information Sciences

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

A unified approach to asymptotic behaviors for the autoregressive model with fuzzy data

Hye Young Jung *, Woo Joo Lee, Jin Hee Yoon

Department of Mathematics, Yonsei University, 50 Yonsei-Ro, Seoul 120-749, Republic of Korea

ARTICLE INFO

Article history:
Received 29 August 2013
Accepted 7 September 2013
Available online 17 September 2013

Keywords:
Autoregressive model
Least squares estimation
Fuzzy data
Asymptotic normality
Consistency

ABSTRACT

We propose a unified estimator for the autoregressive model with fuzzy input–output variables based on the least squares method. The least squares estimation is investigated in presence of a unified ρ-distance defined on the space of fuzzy numbers. We investigate asymptotic properties of the unified estimator under some simple conditions as well as a generalization, which reduces to the asymptotic properties under those distances when the distances are the special cases of ρ-distance. Some simulation studies are included to compare the asymptotic properties of estimators formed under several distances being the special cases of ρ-distance.

1. Introduction

Time series analysis has been widely used with many successful applications. In time series analysis, the available experimental data are assumed to be precise. However, in many practical situations we encounter data which are not only random but vague due to ambiguous information or linguistically quantified structure, for example 'about 10', 'greater than 10', 'more or less in-between 5 and '10', or 'fair', 'good', 'excellent', etc. Moreover, sometimes in the context of economic systems, such as stock market price, foreign exchange rate or market sales price of particular commodities, the historical data record officially, close value, rough average value, maximum value or minimum value, instead of the precise series with variation during the time period. Under these circumstances, two different types of uncertainty, namely randomness and fuzziness, start manifesting in time series data. Nevertheless, only random errors are considered. In the presence of fuzzy data probabilistic estimation procedures considered for time series may not be completely suitable. In such scenarios, we would be interested to resort ourselves to the use of models of fuzzy regression.

In recent years the issue of time series analysis using fuzzy sets have been studied by several authors [1,5,9,19]. Interestingly enough, the concept of the parametric estimation has not been fully investigated. In [21,22], the fuzzy ARIMA method was proposed, which uses the fuzzy regression method to fuzzify the parameters of the ARIMA model. In the literature, most of the studies are related to the fuzzy regression that is based on possibility theory and fuzzy set theory. Only a few papers discussed the fuzzy estimation problem for models with both of randomness and fuzziness. Möller and Reuter [15] proposed autoregressive-moving average (p,q) models with general fuzzy input–output data based on the $l_\alpha r_\alpha$-discretization. It is a state-of-the-art model amongst fuzzy time series research with randomness and fuzziness involved. However, [15] did not investigate asymptotic properties of the estimator but invoked estimation of the parameters by means of numerical methods such as Monte Carlo or the modified evolution strategy. It is a valu-

* Corresponding author. Tel.: +82 2 2123 7650; fax: +82 2 392 6634.
E-mail address: hyjung@yonsei.ac.kr (H.Y. Jung).

0020-0255/$ - see front matter
http://dx.doi.org/10.1016/j.ins.2013.09.024

able procedure that we investigate asymptotic properties of the estimator so that we can trust the result of data analysis using parametric model. Many authors have suggested statistical theories related to regression model with fuzzy data, however, the basic theory of time series model with fuzzy data has not been suggested yet. This is the motive that we have studied optimal properties of fuzzy time series. Zhao and Peng [24] extended the standard autoregressive model to the case where the input and output are triangular fuzzy numbers. Moreover, [24] was devoted to the parameter estimation of the model using distance which was defined by Kim et al. [11] and set out the asymptotic properties of the estimation. However, since solely a special case is considered we need to have study about a general case. In this paper, we propose a unified estimator for a autoregressive model with *LR*-fuzzy input/output using the ρ-distance and establish the asymptotic normality and consistency of the proposed estimator. The ρ-distance unifying different general distances applied in statistical inference with fuzzy data was introduced by Näther [17] and the theory of statistical inference with fuzzy data under this ρ-distance has been developed by various authors [8,10,12,14,16,17,20,23]. In addition, we investigate a generalization, which reduces to the least squares estimation under those distances when the distances are the special cases of ρ-distance.

The rest of this paper is structured as follows: In Section 2, some preliminary concepts which are required to develop the main results are presented. In Section 3, the solution for a normal equation is derived by applying least squares method for the autoregressive time series with fuzzy data. In Section 4, the asymptotic normality and consistency of the unified estimator obtained in the previous section are discussed. In addition, some simulation studies are provided in Section 5 to illustrate the results proved in the previous section.

2. Preliminaries

In this section, we recall some notions concerning fuzzy number and distance between two fuzzy numbers.

Definition 1. A mapping $\tilde{u} : \mathbf{R} \to [0,1]$ is called a fuzzy number if $\tilde{u}$ satisfies (i) $\tilde{u}$ is normal, i.e. $\tilde{u}(x_0) = 1$ for some $x_0 \in \mathbf{R}$, (ii) $\tilde{u}$ is convex, (iii) $\tilde{u}$ is upper semicontinuous, (iv) supp $\tilde{u} := \overline{\{x \in \mathbf{R} : \tilde{u}(x) > 0\}}$ is compact.

The set of all fuzzy number is denoted by $\mathcal{F}_c(\mathbf{R})$. A real number $a \in \mathbf{R}$, a can be considered as a special fuzzy number in $\mathcal{F}_c(\mathbf{R})$ whose membership function $a(x)$ equals 1 at a and *0* elsewhere. As a special case, we often use the following parametric class of fuzzy numbers, known as *LR*-fuzzy numbers:

$$\tilde{u}(x) = \begin{cases} L((u^m - x)/(u^m - u^l)) & \text{if } x \leqslant u^m, \\ R((x - u^m)/(u^r - u^m)) & \text{if } x > u^m \end{cases} \quad \text{for } x \in \mathbf{R},$$

where $L, R{:}\mathbf{R}^+ \to [0,1]$ are fixed left-continuous and non-increasing functions with $R(0) = L(0) = 1$ and $R(1) = L(1) = 0$. L and R are called left and right shape functions of $\tilde{u}$, u^m is the mode of $\tilde{u}$ and u^l, u^r are left, right endpoints of $\tilde{u}$. In order to describe an *LR*-fuzzy number $\tilde{u}$ the abridged notation $\tilde{u} = (u^l, u^m, u^r)_{LR}$ is used. The spreads $u^m - u^l$ and $u^r - u^m$ quantify the fuzziness of the number and the resulting fuzzy number could be symmetric or non-symmetric. If the spreads are zero, there is no fuzziness of the number, and it is a real number. We denote the set of all *LR*-fuzzy numbers as $\mathcal{F}_{LR}(\mathbf{R})$.

Definition 2. The α-cut of a fuzzy number $\tilde{u}$ is a set defined as $\tilde{u}_\alpha := \{x \in \mathbf{R} | \tilde{u}(x) \geqslant \alpha\}, 0 < \alpha \leqslant 1$. For $\alpha = 0$, $\tilde{u}_0 := \overline{\{x \in \mathbf{R} | \tilde{u}(x) > 0\}}$.

Note that the α-cut $\tilde{u}_\alpha$ of $\tilde{u} \in \mathcal{F}_{LR}(\mathbf{R})$ is a closed and bounded interval for each $\alpha \in [0,1]$

$$\tilde{u}_\alpha = \left[u^L_\alpha, u^U_\alpha\right] = [(1 - L^{-1}(\alpha))u^m + L^{-1}(\alpha)u^l, (1 - R^{-1}(\alpha))u^m + R^{-1}(\alpha)u^r].$$

For $\tilde{u}, \tilde{v} \in \mathcal{F}_{LR}(\mathbf{R}), \lambda \in \mathbf{R}$, we define

$$(\tilde{u} + \tilde{v})_\alpha = \tilde{u}_\alpha + \tilde{v}_\alpha = \left[u^L_\alpha + v^L_\alpha, u^U_\alpha + v^U_\alpha\right],$$

$$(\lambda\tilde{u})_\alpha = \lambda\tilde{u}_\alpha = \begin{cases} \left[\lambda u^L_\alpha, \lambda u^U_\alpha\right] & \text{if } \lambda \geqslant 0, \\ \left[\lambda u^U_\alpha, \lambda u^L_\alpha\right] & \text{if } \lambda < 0, \end{cases}$$

where $\alpha \in [0,1]$.

In addition, the advantage of *LR*-fuzzy numbers is that the opertions + and $\cdot$ can be expressed by simple operations:

$$(u^l, u^m, u^r)_{LR} + (v^l, v^m, v^r)_{LR} = (u^l + v^l, u^m + v^m, u^r + v^r)_{LR},$$

$$\lambda\tilde{u} = \lambda(u^l, u^m, u^r)_{LR} = \begin{cases} (\lambda u^l, \lambda u^m, \lambda u^r)_{LR} & \text{if } \lambda > 0, \\ (\lambda u^r, \lambda u^m, \lambda u^l)_{LR} & \text{if } \lambda < 0, \\ (0,0,0)_{LR} & \text{if } \lambda = 0. \end{cases}$$

These two operations above are equal to the ordinary addition and scalar multiplication of vectors, respectively.