Computers & Industrial Engineering 110 (2017) 207–215

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Investigation on objective function and assessment rule in fuzzy regressions based on equality possibility, fuzzy union and intersection concepts

Hamed Shakouri, Reza Nadimi *, Seyed Farid Ghaderi

Department of Industrial Engineering, Tehran University, Iran

ARTICLE INFO

Article history:
Received 24 September 2016
Received in revised form 31 January 2017
Accepted 27 May 2017
Available online 31 May 2017

Keywords:
Fuzzy linear regression
Fuzzy sets
Possibility of equality
Fuzzy number

ABSTRACT

Various methods have been proposed to estimate parameters of fuzzy linear regression models by possibility theory. A novel definition of the possibility of equality of two fuzzy numbers is introduced here. This novel equality of possibility is applied to objective function in the fuzzy linear regression model to obtain coefficients of the model. To evaluate the results of the model with other models, a more precise method of measuring error is developed to calculate difference between fuzzy numbers, based on the fuzzy union and intersection concepts. Several numerical examples are presented to illustrate the capability of the proposed approach in comparison to the other methods by obtaining more accurate results. Finally an application of the method is given for SAFA Rolling & Pipe Mills Company.

1. Introduction

Fuzzy regression models have received much attention after introduction by Tanaka, Uegima, and Asia (1982). These models have been successfully applied in various fields of forecasting (Al-Kandaria, Solimanb, & El-Hawary, 2004; Beenstock, Goldin, & Nabot, 1999; Carcedo & Otero, 2005; Manusov & Mogilenko, 2002; Nazarko & Zalewski, 1999; Olsina, Garces, & Haubrich, 2006; Sedelnikov & Manusov, 2004; Shakouri & Nadimi, 2007, 2013; Shakouri, Nadimi, & Ghaderi, 2009). Fuzzy regression model is an extension of the statistical regression model in which one or both of input and output data are considered as fuzzy numbers.

Possibilistic and least-squares are the two general well-known forms of fuzzy regression models; the former attempts to minimize the fuzziness of the model (Tanaka, 1987; Tanaka, Hayashi, & Watada, 1989; Tanaka & Watada, 1988; Tanaka et al., 1982), whereas the latter aims minimizing the gap between observed and estimated outputs (Chang & Lee, 1994; Coppia, Ursob, Giordania, & Santorob, 2006; D'Urso and Gastaldi, 2000; Diamond, 1988; Ming, Friedman, & Kandel, 1997; Yang & Lin, 2002).

Verification of the estimated outputs are spun around both the model's parameters and measuring criterion. Various researches have been focused on the least-squares models by considering the concept of possibility theory (Zadeh, 1978) to predict the model parameters. Several criteria, also, have been taken to measure the accuracy of results and finally to verify the model (Icen & Demirhan, 2016). However, the shortcoming in either sides, prediction of the parameters or measuring criterion, leads toward incorrect decision in the model verification and comparison among different fuzzy regression model's results.

Possibility of equality of two fuzzy numbers which was introduced by Dubois and Prade (1980), briefly called DP criterion in this paper, has been variously applied to estimate the parameters of fuzzy regression models. Kim and Bishu (1998) (KB) measuring criterion has also been used to measure the difference between observed data and estimated outputs.

Chen, Hsueh, and Chang (2013) applied two-stage method to generate a fuzzy regression model based on the distance concept. At the first stage, fuzzy numbers were defuzzified into the crisp number and then ordinary regression was implemented on the defuzzified numbers. In the second stage, a mathematical programming model was applied to minimize the total error. They applied the KB criterion to compare the results of their model.

Savic and Pedrycz (1991) employed a linear regression model based on a two-stage structure to minimize the vagueness criterion. They supposed distance between the center and the spread of observed and estimated fuzzy numbers as a criterion to compare results of the approach with other models. They neglected, however, to consider impact of membership functions in their criterion.

* Corresponding author.
E-mail addresses: hshakouri@ut.ac.ir (H. Shakouri), r.nadimi@alumni.ut.ac.ir (R. Nadimi), ghaderi@ut.ac.ir (S.F. Ghaderi).

http://dx.doi.org/10.1016/j.cie.2017.05.032
0360-8352/

Kim and Bishu (1998) revised fuzzy linear regression analysis in terms of the fuzzy membership values between estimated and observed data. Finally, they defined a proportional error term to measure and evaluate the results of their own approach with other methods.

Modarres, Nasrabadi, and Nasrabadi (2005) applied the possibility of equality (introduced by Dubois and Prade (1980)) between two fuzzy numbers, as a degree of fitness of their model in a quadratic programming model. They used KB measure to prove the performance of their method.

Shakouri and Nadimi (2009) suggested a non-equality possibility index to estimate fuzzy parameters. They applied that index as an objective function to find a minimum degree of acceptable uncertainty. They also utilized KB criterion to illustrate the accuracy of their results.

Jung, Yoon, and Choi (2015) expressed the limitation of the KB criterion in measuring of two fuzzy numbers. They just addressed the lack of sensitivity of the KB criterion in overlapping area without considering the effect of uncertainty variation in non-overlapping area which will be consider in this paper.

However, DP criterion concentrates on the common area of the two fuzzy numbers to obtain the possibility of equality without considering the uncommon area of the two fuzzy numbers. Moreover, KB criterion focuses only the intersection concept to assess results of fuzzy regression models. Close scrutiny of these two criteria reveals that without taking into account the union concept and just relying on intersection notion leads to less reliable results to validate the model. These shortcomings, which will be discussed in the rest of the present paper, have been the main stimuli to research on proposing an alternative definition and a different model.

Current study presents new equality possibility concept to generate a new objective function in fuzzy regression model. This paper also suggests a more precise method by combination of both union and intersection concepts to compare the results of the proposed approach with other methods.

The paper is organized as follows. The concept of fuzzy number and equality possibility index are briefly introduced in Section 2. Then shortcoming of the DP equality possibility index is explained in more details. Section 3 presents fuzzy linear regression (FLR) model in brief. Section 4 is devoted to describe our proposed solution to FLR problems, along with formulating a new method to evaluate the performance of fuzzy regression models. Numerical examples as well as an industrial case study are given in Section 5 to compare the results obtained by the proposed approach with that of other methods. The concluding remarks are clarified at the end.

2. Fuzzy numbers

Based on Dubois and Prade's definition (Dubois & Prade, 1980), $\tilde{A}$ is assumed a fuzzy number if it satisfies the following criteria:

First: normality, which means: $\exists\, x \in \boldsymbol{R}$ such that $\mu_{\tilde{A}}(x) = 1$.
Second: convexity, it means $\forall x_1, x_2 \in \boldsymbol{R}, \forall h \in [0;1]$.
$\mu_{\tilde{A}}(h\, x_1 + (1-h)\, x_2) \geq \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2))$.

It is common to define a *LR*-type fuzzy number (LR stands for the left and right of the shape function by which a membership function is defined. Please see Eq. (1)) as $\tilde{A} = (c_L, a, c_R)_{LR}$ where a, c_L and c_R are the center, left spread and right spread of fuzzy number, respectively (c_L and $c_R > 0$).

A fuzzy number, $\tilde{A}$, is called symmetric triangular fuzzy number provided that $c_L = c_R = c$. Therefore $\tilde{A} = (a, c)_L$ is a symmetric triangular fuzzy number if:

$$\mu_{\tilde{A}}(x) = L((a-x)/c) = 1 - \frac{|a-x|}{c}, a - c \leqslant x \leqslant a + c. \quad (1)$$

In this paper, symmetric triangular fuzzy numbers are only considered for simplicity and linearity which is normally expected for FLR problems.

2.1. Equality possibility criterion

Dubois and Prade (1980) proposed the index below in order to define the possibility of equality between two fuzzy numbers.

$$Poss_{DP}(\tilde{A}_1 = \tilde{A}_2) = \sup_{x \in R^1} \min\{\mu_{\tilde{A}_1}(x), \mu_{\tilde{A}_2}(x)\}, \quad (2)$$

where *Poss* is short for Possibility and the subscript *DP* denotes Dubois-Prade.

Some researchers applied above index as a degree of fitness in the objective function or the constraints of mathematical programming to calculate the parameters of fuzzy regression models. Although, it often works in a good manner, but in some cases it may lead towards inaccurate results. For the purpose of this paper, $\tilde{A}_1$ is presumed a fuzzy parameter that has to be estimated by fuzzy regression analysis based on the equality possibility index.

Assuming that $\tilde{A}_2$ and $\tilde{A}_3$ are two candidates to estimate $\tilde{A}_1$ based on *DP*'s index (See Fig. 1). According to the *DP*'s index, there is no preference between $\tilde{A}_2$ and $\tilde{A}_3$ as a candidate for estimation of $\tilde{A}_1$ because both candidates have same equality of possibility with $\tilde{A}_1$. Whereas the common area between $\tilde{A}_2$ and $\tilde{A}_1$ is larger than the common area between $\tilde{A}_3$ and $\tilde{A}_1$.

Focusing attention just on the supremum of the minimum of the membership functions is the major reason leading to such an inaccuracy. Therefore, considering both the common and uncommon areas is proposed to solve this problem and get more accurate results. Indeed, this problem is a strong motivation to think of a solution which will be presented and discussed in Section 4.

3. Fuzzy linear regression model

Although there may be complex relationships in the real world, the principle of parsimony in modelling (Grasa, 1989), a linear regression is always preferred to any other sophisticated model, as far as it can describe a phenomenon well (Shakouri & Menhaj, 2008). Thus, we focus herein on FLR models. Tanaka et al. (1982) introduced a formulation for FLR model as:

$$\tilde{Y}_i^* = f(\boldsymbol{X_i}, \tilde{A}) = \tilde{A}_0 X_{i0} + \tilde{A}_1 X_{i1} + \cdots + \tilde{A}_n X_{in}, \quad (3)$$

where $\tilde{Y}_i^*, i = 1, \ldots, m$ and $\tilde{A}_j = (a_j, c_j)_L, j = 0,1, \ldots, n$ are the estimated fuzzy output (dependent variable) and the set of symmetric fuzzy coefficients respectively. Independent variables are denoted by the vector $\boldsymbol{X}_i = [X_{i0}, X_{i1}, \ldots, X_{in}]^T$. By applying the extension principle

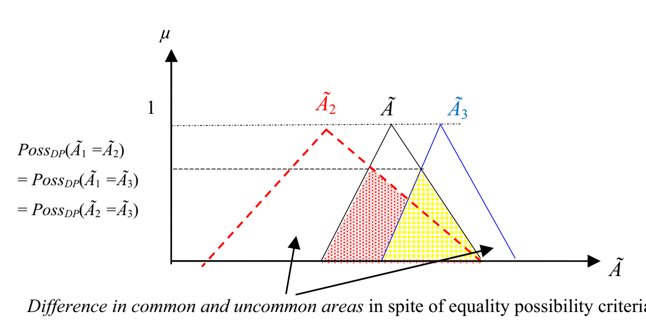

Fig. 1. DP's Index along with both common areas and uncommon areas.