AMO - Advanced Modeling and Optimization, Volume 17, Number 1, 2015

LINEAR FUZZY REGRESSION MODEL

U. M. PIRZADA AND JAITA SHARMA

Abstract. This paper deals with the linear fuzzy regression model with fuzzy data and crisp non-negative parameters. To find the best values of the parameters of the fuzzy model, we minimize the sum of squared errors of lower and upper level functions of fuzzy error function.

1. Introduction

Regression analysis is one of the most applicable tool used in methods of estimation. It deals with the investigation of the dependence of a variable upon one or more variables. The aim of regression analysis is to estimate the parameters those occur in its particular mathematical form. Usually, method of least square is applied to determine the values of parameters.

The variables involve in the regression model are not always crisp in nature. To make more realistic regression model, we fuzzify the variables. Instead of dealing with exact data points, we can consider the approximate data points which represents in terms of fuzzy numbers. In this case, we get the fuzzy regression model which can be further classified into linear or non-linear form.

Fuzzy regression analysis is introduced by Tanaka in [14]. His approach was based on linear programming. After that, many authors [2, 6, 7, 8, 11, 12, 15] have studied this theory using different approaches. Both variables and parameters are fuzzy in the model discussed by [2, 12].

In this paper, we consider the following linear fuzzy regression model with fuzzy data points and crisp parameters.

$$\tilde{y}_i = \beta_0 \oplus \beta_1 \odot \tilde{x}_i,\ i = 1, 2, ..., n \tag{1.1}$$

where $(\tilde{x}_i, \tilde{y}_i)$, are fuzzy numbers, n is number of data points, β_0, β_1 are crisp non-negative parameters.

Diamond [2] has applied the concept of least square under a suitable metric to develop a methodology and determine the regression parameters. In this paper, to determine the best values of parameters, β_0, β_1, we consider the fuzzy error for each data point. Then we find the sum of squared errors of lower and upper level functions of fuzzy error function. Our approach of computation is simpler than the previous approaches in the sense that we minimize the error using crisp function corresponding to fuzzy function.

2000 *Mathematics Subject Classification.* 03E72, 90C70.

Key words and phrases. Fuzzy numbers and Regression analysis and Least square approach.

This research is supported by National Board for Higher Mathematics (NBHM) , Department of Atomic Energy (DAE), India.

AMO - Advanced Modeling and Optimization. ISSN: 1841-4311.

Outline of the paper is as follows:

Section 2 is preliminary, contains basic concepts regarding fuzzy numbers. In Section 3, we study the linear fuzzy regression model with fuzzy data points and determine the values of crisp non-negative parameters. Example is given in the same section. Conclusion is given in Section 4.

2. Preliminaries

In this section, we cite some basic definitions regarding fuzzy numbers.

Definition 1. [3] Let $\mathbb{R}$ be the set of real numbers and $\tilde{a} : \mathbb{R} \to [0,1]$ be a fuzzy set. We say that $\tilde{a}$ is a fuzzy number if it satisfies the following properties:

(i) $\tilde{a}$ is normal, that is, there exists $x_0 \in \mathbb{R}$ such that $\tilde{a}(x_0) = 1$;
(ii) $\tilde{a}$ is fuzzy convex, that is, $\tilde{a}(tx + (1-t)y) \geq min\{\tilde{a}(x), \tilde{a}(y)\}$, whenever x, y $\in \mathbb{R}$ *and* $t \in [0,1]$;
(iii) $\tilde{a}(x)$ is upper semi-continuous on $\mathbb{R}$, that is, $\{x/\tilde{a}(x) \geq \alpha\}$ is a closed subset of $\mathbb{R}$ for each $\alpha \in (0,1]$;
(iv) $cl\{x \in \mathbb{R}/\tilde{a}(x) > 0\}$ forms a compact set.

where cl denotes closure of a set. The set of all fuzzy numbers on $\mathbb{R}$ is denoted by $F(\mathbb{R})$. For all $\alpha \in (0,1]$, α-level set $\tilde{a}_\alpha$ of any $\tilde{a} \in F(\mathbb{R})$ is defined as $\tilde{a}_\alpha = \{x \in \mathbb{R}/\tilde{a}(x) \geq \alpha\}$. The 0-level set $\tilde{a}_0$ is defined as the closure of the set $\{x \in \mathbb{R}/\tilde{a}(x) > 0\}$. By definition of fuzzy numbers, we can prove that, for any $\tilde{a} \in F(\mathbb{R})$ and for each $\alpha \in (0,1]$, $\tilde{a}_\alpha$ is compact convex subset of $\mathbb{R}$, and we write $\tilde{a}_\alpha = [\tilde{a}^L_\alpha, \tilde{a}^U_\alpha]$. $\tilde{a} \in F(\mathbb{R})$ can be recovered from its α-level sets by a well-known decomposition theorem (ref. [4]), which states that $\tilde{a} = \cup_{\alpha \in [0,1]} \alpha \cdot \tilde{a}_\alpha$ where union on the right-hand side is the standard fuzzy union.

Definition 2. [13] The membership function of a triangular fuzzy number $\tilde{a}$ is defined as

$$\tilde{a}(r) = \begin{cases} \frac{(r-a_1)}{(a-a_1)} \; if \; a_1 \leq r \leq a \\ \frac{(a_2-r)}{(a_2-a)} \; if \; a < r \leq a_2 \\ 0 \; otherwise \end{cases}$$

which is denoted by $\tilde{a} = (a_1, a, a_2)$. The α-level set of $\tilde{a}$ is then

$$\tilde{a}_\alpha = [(1-\alpha)a_1 + \alpha a, (1-\alpha)a_2 + \alpha a].$$

Definition 3. According to Zadeh's extension principle, we define addition, subtraction and product of two fuzzy numbers $\tilde{a}, \tilde{b} \in F(\mathbb{R})$ by their α-level sets as follows:

$$\begin{aligned} (\tilde{a} \oplus \tilde{b})_\alpha &= [\tilde{a}^L_\alpha + \tilde{b}^L_\alpha, \tilde{a}^U_\alpha + \tilde{b}^U_\alpha] \\ (\tilde{a} \ominus \tilde{b})_\alpha &= [\tilde{a}^L_\alpha - \tilde{b}^U_\alpha, \tilde{a}^U_\alpha - \tilde{b}^L_\alpha] \\ (\tilde{a} \otimes \tilde{b})_\alpha &= [\min\{\tilde{a}^L_\alpha\tilde{b}^L_\alpha, \tilde{a}^L_\alpha\tilde{b}^U_\alpha, \tilde{a}^U_\alpha\tilde{b}^L_\alpha, \tilde{a}^U_\alpha\tilde{b}^U_\alpha\}, \max\{\tilde{a}^L_\alpha\tilde{b}^L_\alpha, \tilde{a}^L_\alpha\tilde{b}^U_\alpha, \tilde{a}^U_\alpha\tilde{b}^L_\alpha, \tilde{a}^U_\alpha\tilde{b}^U_\alpha\}] \end{aligned}$$

where $\tilde{a}_\alpha = [\tilde{a}^L_\alpha, \tilde{a}^U_\alpha], \tilde{b}_\alpha = [\tilde{b}^L_\alpha, \tilde{b}^U_\alpha]$, for $\alpha \in [0,1]$.

Moreover, we define scalar multiplication of fuzzy number $\tilde{a}$ with $\lambda \in \mathbb{R}$ as follows:

$$\begin{aligned} (\lambda \odot \tilde{a})_\alpha &= [\lambda \cdot \tilde{a}^L_\alpha, \lambda \cdot \tilde{a}^U_\alpha], \; if \; \lambda \geq 0 \\ &= [\lambda \cdot \tilde{a}^U_\alpha, \lambda \cdot \tilde{a}^L_\alpha], \; if \; \lambda < 0 \end{aligned}$$

Fuzzy-valued function is defined as follows:

Definition 4. Let V be a real vector space and $F(\mathbb{R})$ be a set of fuzzy numbers. Then a function $\tilde{f} : V \to F(\mathbb{R})$ is a fuzzy-valued function defined on V.
Corresponding to such a function $\tilde{f}$ and $\alpha \in [0, 1]$, we define two real-valued functions $\tilde{f}^L_\alpha$ and $\tilde{f}^U_\alpha$ on V as $\tilde{f}^L_\alpha(x) = (\tilde{f}(x))^L_\alpha$ and $\tilde{f}^U_\alpha(x) = (\tilde{f}(x))^U_\alpha$ for all $x \in V$ are called lower and upper α-level functions respectively.

3. Linear fuzzy regression model

We consider the following linear fuzzy regression model (1.1):

$$\tilde{y}_i = \beta_0 \oplus \beta_1 \odot \tilde{x}_i,\ i = 1, 2, ..., n$$

where $(\tilde{x}_i, \tilde{y}_i)$ are fuzzy data - fuzzy numbers, n is number of data points, $\beta_0, \beta_1 \geq 0$ are crisp non-negative parameters.

The fuzzy estimated error associated with each data is defined as follows:

$$\tilde{\epsilon}_i = [\tilde{y}_i \ominus (\beta_0 \oplus \beta_1 \odot \tilde{x}_i)],$$

where $i = 1, ..., n$. In terms of α-level sets, fuzzy estimated error is

$$(\tilde{\epsilon}_i)_\alpha = [(\tilde{y}_i)^L_\alpha - (\beta_0 + \beta_1(\tilde{x}_i)^U_\alpha), (\tilde{y}_i)^U_\alpha - (\beta_0 + \beta_1(\tilde{x}_i)^L_\alpha)],$$

where $(\tilde{y}_i)^L_\alpha - (\beta_0 + \beta_1(\tilde{x}_i)^U_\alpha)$ are lower values of fuzzy estimated error for i^{th} data point and $(\tilde{y}_i)^U_\alpha - (\beta_0 + \beta_1(\tilde{x}_i)^L_\alpha)$ are upper values of fuzzy estimated error for i^{th} data point. To determine the values of parameters β_0, β_1 so that the fuzzy estimated error is approximately zero, we define sum of squared error in terms of lower and upper α-level functions of the fuzzy estimated error. Therefore, we have

$$S_\alpha(\beta_0, \beta_1) = \sum_{i=1}^{n} ((\tilde{y}_i)^L_\alpha - (\beta_0 + \beta_1(\tilde{x}_i)^U_\alpha))^2 + ((\tilde{y}_i)^U_\alpha - (\beta_0 + \beta_1(\tilde{x}_i)^L_\alpha))^2$$

The first order necessary conditions in the integral forms are :

$$\int_0^1 \frac{\partial}{\partial \beta_0} S_\alpha(\beta_0, \beta_1) d\alpha = 0$$

and

$$\int_0^1 \frac{\partial}{\partial \beta_1} S_\alpha(\beta_0, \beta_1) d\alpha = 0$$

Simplifying, we get the following equations

$$Y = 2n\beta_0 + \beta_1 X, YX = X\beta_0 + \beta_1 X2, \tag{3.1}$$

where

$$Y = \int_0^1 \sum_{i=1}^{n} (\tilde{y}^L_{i\alpha} + \tilde{y}^U_{i\alpha}) d\alpha,$$

$$X = \int_0^1 \sum_{i=1}^{n} (\tilde{x}^L_{i\alpha} + \tilde{x}^U_{i\alpha}) d\alpha,$$

$$XY = \int_0^1 \sum_{i=1}^{n} (\tilde{y}^L_{i\alpha} \tilde{x}^U_{i\alpha} + \tilde{y}^U_{i\alpha} \tilde{x}^L_{i\alpha}) d\alpha \ and$$

TABLE 1. Fuzzy error $\tilde{\epsilon}_i$, $i = 1, ..., 8$

$\tilde{\epsilon}^L$	$\tilde{\epsilon}$	$\tilde{\epsilon}^U$
-0.4355	0.4165	1.2686
0.0083	0.8604	1.7124
-0.2519	1.4522	3.1563
-1.4560	-0.6039	0.2481
-0.5121	0.3399	1.1920
-3.2724	-1.5683	0.1358
-0.2724	0.5797	1.4317
-2.3285	-2.3285	-0.6244

$$X2 = \int_0^1 \sum_{i=1}^{n} ((\tilde{x}_{i\alpha}^L)^2 + (\tilde{x}_{i\alpha}^U)^2) d\alpha.$$

Solving these equations, we get the parameters β_0 and β_1.

Example 1. *We consider a linear fuzzy regression model (1.1) for following eight sets of fuzzy data points (triangular fuzzy numbers):*

$\tilde{x} = \{(1.5, 2.0, 2.5), (3.0, 3.5, 4.0), (4.5, 5.5, 6.5), (6.5, 7.0, 7.5), (8.0, 8.5, 9.0), (9.5, 10.5, 11.5), (10.5, 11.0, 11.5), (12.0, 12.5, 13.0)\}$

$\tilde{y} = \{(3.5, 4.0, 4.5), (5.0, 5.5, 6.0), (6.5, 7.5, 8.5), (6.0, 6.5, 7.0), (8.0, 8.5, 9.0), (7.0, 8.0, 9.0), (10.0, 10.5, 11.0), (9.0, 9.5, 10.0)\}$

Using equations (3.1), we get $\beta_0 = 2.1753$ *and* $\beta_1 = 0.7041$.

The calculated values of $\tilde{Y}$ *using fuzzy regression equation are :*

$\tilde{y} = \{(3.2314, 3.5835, 3.9355), (4.2876, 4.6396, 4.9917), (5.3437, 6.0478, 6.7519), (6.7519, 7.1039, 7.4560), (7.8080, 8.1601, 8.5121), (8.8642, 9.5683, 10.2724), (9.5683, 9.9203, 10.2724), (10.6244, 10.9765, 11.3285)\}$.

The fuzzy errors between actual values $\tilde{y}$ *and calculated values of* $\tilde{Y}$ *are shown in Table 1.*

Here we observed that fuzzy errors are not sufficiently small for some data points. For instance, for the third, sixth and eighth data points, absolute fuzzy errors are greater compared to other data points (see Table 1). To find the more accurate results, we can increase the fuzzy data points or we can make a quadratic fuzzy regression model.

4. CONCLUSIONS

We have studied the linear fuzzy regression model to find the crisp non-negative parameters. We have minimized sum of squared error of level functions of fuzzy error function. Our future interest is to study and analysis fuzzy linear and non-linear regression models using different methodologies to find more accurate results.

References

[1] Diamond P., Kloeden, 1994, Metric spaces of fuzzy sets: Theory and Applications. World Scientific.

[2] Diamond P., 1988, Fuzzy least squares. Information Sciences, 46 141-157.

[3] George A. A., 2003, Fuzzy Ostrowski Type Inequalities. Computational and Applied mathematics, 22 279-292.

[4] George A. A., 2005, Fuzzy Taylor Formulae. CUBO, A Mathematical Journal, 7 1-13.

[5] Hsien-Chung Wu, 2004, An (α, β)-Optimal Solution Concept in Fuzzy Optimization Problems. Optimization, 53 203-221.

[6] Kim B., Bishu R.R., 1998, Evaluation of fuzzy linear regression models by comparing membership functions. Fuzzy Sets and Systems 100, 343352.

[7] Kim K.J., Moskowitz H., Koksalan M., 1996, Fuzzy versus statistical linear regression. European Journal of Operational Research 92 , 417434.

[8] Peters G., 1994, Fuzzy linear regression with fuzzy intervals. Fuzzy Sets and Systems, 63 45-55.

[9] Puri M. L., Ralescu D. A., 1983, Differentials of fuzzy functions. J. of Math. Analysis and App., 91 552-558.

[10] RamíK J. and Rimanek J., 1985, Inequality relation between fuzzy numbers and its use in fuzzy optimization. Fuzzy Sets and Systems, 16 123-138.

[11] Savic D., Pedrycz W., 1991, Evaluation of fuzzy linear regression models. Fuzzy Sets and Systems, 39 51-63.

[12] Sakawa M. , Yano H. , 1992, Multiobjective fuzzy linear regression analysis for fuzzy inputoutput data. Fuzzy Sets and Systems, 47 173-181.

[13] Saito, S. and Ishii H., 2001, L-Fuzzy Optimization Problems by Parametric Representation. IEEE. 1173-1177

[14] Tanaka H., Uegima S., Asai K., 1982, Linear regression analysis with fuzzy model. IEEE Trans. Systems, Man and Cybernetics, 12 903-907.

[15] Yen K.K., Ghoshray S., Roig G., 1999, A linear regression model using triangular fuzzy number coefficients, Fuzzy Sets and Systems, 106 167-177.

Department of Applied Mathematics, Faculty of Tech. & Engg., M. S. University of Baroda, Vadodara-390 001, Gujarat, India

E-mail address: salmapirzada@yahoo.com

Department of Applied Mathematics, Faculty of Tech. & Engg., M. S. University of Baroda, Vadodara-390 001, Gujarat, India

E-mail address: jaita.sharma@gmail.com