FUZZY LINEAR REGRESSION BASED ON LEAST ABSOLUTES DEVIATIONS

S. M. TAHERI AND M. KELKINNAMA

Abstract. This study is an investigation of fuzzy linear regression model for crisp/fuzzy input and fuzzy output data. A least absolutes deviations approach to construct such a model is developed by introducing and applying a new metric on the space of fuzzy numbers. The proposed approach, which can deal with both symmetric and non-symmetric fuzzy observations, is compared with several existing models by three goodness of fit criteria. Three well-known data sets including two small data sets as well as a large data set are employed for such comparisons.

1. Introduction and Literature Review

Regression analysis is an important tool in evaluating the functional relationship between a certain variable called the dependent variable, and a set of other variables called explanatory variables. In statistical regression analysis, the estimation of the parameters and the prediction of variables are done based on a set of crisp data. On the other hand, it is usually assumed that the parameters of the underlying model are exact numbers (i.e. the relationship between variables is crisp). But in systems in which human intelligence participates, we usually encounter the following two cases:

1) the observations due to the variables are fuzzy, and/or
2) the relationship between variables is imprecise.

We need, therefore, to investigate some soft methods for dealing with these situations. Fuzzy set theory provides appropriate tools for regression analysis when the relationship between variables is vaguely defined and/or the observations are reported as imprecise quantities.

After introducing fuzzy set theory, several approaches to fuzzy regression have been developed by some authors. Let us briefly review some important studies in fuzzy regression models. For the first time, Tanaka et al. [32] proposed a linear regression model with fuzzy parameters. Their method, in which the observed data are crisp, has been developed in different directions by some authors, (see for example [10, 16, 31, 37]). Tanaka et al.’s approach is essentially based on transforming the problem of fitting a fuzzy model on a data set to a linear programming problem. Pourahmad et al. [25, 26] introduced two fuzzy logistic regression models which are

Accepted: March 2009; Revised: July 2010 and January 2011; Accepted: June 2011

Key words and phrases: Fuzzy regression, Least absolutes deviations, Metric on fuzzy numbers, Similarity measure, Goodness of fit.
used when the explanatory (input) variables are crisp but the value of the binary response (output) variable is reported as a number between zero and one or by a linguistic term.

Another approach to fuzzy regression is introduced by Celmins [3] and Diamond [12], using a generalized least squares method. In fuzzy least squares approach, the optimal model is usually derived based on a metric on the space of fuzzy numbers. For more on this approach and some applications see, for example, [1, 5, 11, 14, 15, 23, 36].

Both of the above approaches to fuzzy regression are sensitive in terms of outlier data points. In such cases, when there exist some outliers in the data set, it is usually preferred to use a robust approach. Traditionally, the regression analysis based on the method of least absolutes deviations (LAD) was used as a robust method with respect to the least squares (LS) method in modeling a data set in which there were some outliers ([13, 28]). The problem of existing outliers may occur when we want to model a set of fuzzy (imprecise) data. So, it is necessary to develop fuzzy regression methods based on the LAD approach. So far, there are just a few researches in this topic, which are reviewed below. Chang and Lee [6] studied the fuzzy LAD regression based on a ranking method for fuzzy numbers. Torabi and Behboodian [33] proposed a LAD-based method to estimate the fuzzy parameters in a linear model with fuzzy input and fuzzy output. Choi and Buckley [9] suggested the LAD estimators to construct the fuzzy regression model, and investigated the performance of their regression models in comparison with least squares method. Taheri and Kelkinnama [30] studied a least absolutes regression model, based on crisp input-fuzzy output data. Hassanpour et al. [17] proposed a goal programming approach to fuzzy regression modeling for non-fuzzy input-fuzzy output data, which is in some sense based on absolute errors. They applied a similar goal programming approach to fuzzy regression with fuzzy input-output data, too [18]. By using the generalized Hausdorff metric, Chachi and Taheri [4] investigated a least absolutes approach to multiple fuzzy regression modeling for crisp input-fuzzy output data.

The main contribution of this work is to investigate the fuzzy regression model with crisp/fuzzy input and fuzzy output, as follows

\[ \hat{Y}_i = a_0 \oplus (a_1 X_{i1}) \oplus (a_2 X_{i2}) \oplus \cdots \oplus (a_p X_{ip}) \oplus E, \]

on the basis of observed data set \{ (X_i, Y_i) : i = 1, \ldots, n \}, \( X_i = (1, X_{i1}, X_{i2}, \ldots, X_{ip}) \), where \( X_{ij}, i = 1, \ldots, n, j = 1, \ldots, p \), and \( Y_i, i = 1, \ldots, n \) are fuzzy numbers. In model (1), \( a_0, a_1, \ldots, a_p \) are crisp coefficients, and \( E \) is the fuzzy error term. Note that, by considering the fuzzy error term \( E \), the model could be used for cases in which all input observations are crisp too.

Here, we briefly review the important studies which have been done on the above fuzzy regression model. Kao and Chyu [19] proposed a two-stage approach for analyzing such a model. In stage one, using the defuzzified observations by centroid method, they estimated the crisp coefficients by employing the ordinary least squares method. Then, in the second stage, they added the fuzzy error term \( E = (0, \alpha, \beta) \) to the estimated model and for estimating the spreads of \( E \), they used
the error criterion which have been proposed by Kim and Bishu [20], (see section 5.1). They designed a nonlinear programming which minimized the sum of errors between the observed and estimated fuzzy responses. Kim et al. [21] adopted the Kao and Chyu’s approach and replaced least squares estimators by least absolutes estimators in stage one.

It should be mentioned that these two approaches have some disadvantages. First, note that the Kim and Bishu’s error criterion (which is the absolute difference of membership functions of two fuzzy numbers) is independent of distance between fuzzy numbers when the intersection of these numbers is empty. Therefore, their criterion doesn’t measure the distance correctly.

The second disadvantage of the Kao and Chyu’s approach is that in their method, for computing error criterion, one must determine a specific form of relative position of $Y_i$ and $\hat{Y}_i$. But, in practice, there are more than one form of such a position and therefore one can not easily determine which shape would result in minimizing the sum of errors.

Choi and Buckley [9] designed a two-stage procedure to obtain least absolutes estimators in a certain fuzzy regression model. In stage one, like Kao and Chyu’s procedure, they defuzzified fuzzy observations using the centroid method. Since for non-symmetric fuzzy numbers, the left and right spreads may affect in such a defuzzification, they found least absolute deviations estimators more efficient than the least squares deviations estimators. In stage two, because of the first problem in Kim and Bishu’s criterion, i.e. occasional independency of distance, they used another distance to estimate the spreads of the fuzzy error term. They minimized the sum of absolute differences between left (and right) endpoints of supports of the observed and estimated fuzzy responses to estimate the left (and right) spread of $E$. In fact, their approach is a three-stage approach. It should be mentioned that, by using the defuzzification in the first stage, one may lose some information which is involved in the fuzzy data.

It is noticeable that the above approaches don’t use a metric on the space of fuzzy numbers. In this regard, Arabpour and Tata [1] applied Diamond’s metric [12] to obtain least square deviations estimators in some fuzzy regression models. In special cases, for fuzzy input-fuzzy output data, they considered the model $Y = a + bX$, where $a$ and $b$ are crisp numbers. But, since this model does not contain any fuzzy error term, it can not be applied when the input data are crisp.

The purpose of this paper is to introduce a new metric, based on absolute deviations, on the space of fuzzy numbers, and then apply such a metric to develop a new approach to analyze the fuzzy regression model (1). As we will show, unlike the Kao and Chuy, and Choi and Buckley’s approaches, we determine the crisp regression coefficients and fuzzy error term simultaneously based on minimizing the sum of distances between the observed and estimated fuzzy responses, i.e. $Y_i$ and $\hat{Y}_i$.

In addition, we investigate some appropriate criteria to evaluate the goodness of fit of the fuzzy regression models.

Concerning the above discussions, as we will see in details in section 6, our approach can remove the above disadvantages.
This paper is organized as follows: in the next section, we provide some preliminaries about fuzzy numbers and algebraic operations on fuzzy numbers. In section 3, we define a new metric on the space of fuzzy numbers and investigate its properties. In section 4, we propose a new LAD-based approach to fuzzy regression modeling. In section 5, we investigate some criteria for evaluating goodness of fit of fuzzy regression models. In section 6, using three well-known data sets, we describe the performance of the proposed method with respect to some usual methods. Finally, section 7 provides a brief conclusion.

2. Preliminaries

In this section, we recall some necessary definitions which are used throughout this paper. For more details, see e.g. [38].

Let \( X \) be a universal set. A fuzzy set \( A \) of \( X \) is defined by its membership function \( A : X \rightarrow [0, 1] \). Let \( A \) and \( B \) be two fuzzy sets of \( X \). Then, \( A \) is called a subset of \( B \), \( A \subseteq B \), iff \( A(x) \leq B(x), \forall x \), and \( A \) is called equal to \( B \), \( A = B \), if \( A(x) = B(x), \forall x \). The fuzzy set \( A \) is called normal if there exists an element \( x \) s.t. \( A(x) = 1 \). In the following, we assume that the universal set is the real numbers, \( \mathbb{R} \). The \( \alpha \)-level set of \( A \) is defined by \( A_\alpha = \{ x : A(x) \geq \alpha \} \), for \( 0 < \alpha \leq 1 \), and for \( \alpha = 0 \), \( A_0 \), is defined as the closure of the set \( \{ x : A(x) > 0 \} \). The fuzzy set \( A \) of \( \mathbb{R} \) is called convex if all \( A_\alpha \), \( 0 < \alpha \leq 1 \), are convex sets.

Definition 2.1. A fuzzy number \( M \) is a normal convex fuzzy set of \( \mathbb{R} \) with a piecewise continuous membership function.

Definition 2.2. Let \( L \) (and \( R \)) be decreasing, shape functions from \( \mathbb{R}^+ \cup \{0\} \) to \( [0, 1] \) with \( L(0) = 1; L(x) < 1 \) for all \( x > 0 \), \( L(x) > 0 \) for all \( x < 1 \); \( L(1) = 0 \) or \( L(x) > 0 \) for all \( x \) and \( L(+\infty) = 0 \). Then a fuzzy number \( M \) is called of \( LR \)-type if for \( m, \alpha > 0, \beta > 0 \) in \( \mathbb{R} \)

\[
M(x) = \begin{cases} 
L \left( \frac{m-x}{\alpha} \right), & x \leq m \\
R \left( \frac{x-m}{\beta} \right), & x > m 
\end{cases}
\]

where \( m \) is called the mode of \( M \), and \( \alpha \) and \( \beta \) are called the left and right spreads, respectively.

Symbolically, \( M \) is denoted by \( (m, \alpha, \beta)_{LR} \). We denote the set of all fuzzy numbers by \( F(\mathbb{R}) \) and the set of all \( LR \) fuzzy numbers with specified left and right reference functions \( L \) and \( R \), by \( F_{LR}(\mathbb{R}) \).

In special case, where \( L(x) = R(x) = \max(0, 1 - |x|) \), \( M \) is called triangular fuzzy number and is denoted by \( (m, \alpha, \beta)_T \). If, in addition, \( \alpha = \beta \), \( M \) is called symmetric triangular fuzzy number and is denoted by \( (m, \alpha)_T \).

Algebraic operations on fuzzy numbers are defined based on the extension principle. For our purposes, we recall two well-known results.

Definition 2.3. Let \( M, N \in F(\mathbb{R}) \), \( f : \mathbb{R} \rightarrow \mathbb{R} \) be a unary operation and \( * : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \) be a binary operation. Then
Fuzzy Linear Regression Based on Least Absolutes Deviations

a) $f(M)$ is defined by a fuzzy set of $R$ with the following membership function

$$f(M)(y) = \sup_{x:y=f(x)} M(x)$$

b) $M \odot N$ ($\odot$: extended operation of *) is defined as a fuzzy set of $R$ with the following membership function

$$(M \odot N)(z) = \sup_{x,y:x \ast y = z} \min\{M(x), N(y)\}$$

**Proposition 2.4.** Let $M = (m, \alpha, \beta)_{LR}$ and $N = (n, \gamma, \delta)_{LR}$ be two LR-type fuzzy numbers and $\lambda \in R$. Then

- a) $\lambda M = \begin{cases} (\lambda m, -\lambda \beta, -\lambda \alpha)_{RL}, & \lambda < 0 \\ I_{\{0\}}(\cdot), & \lambda = 0 \\ (\lambda m, \lambda \alpha, \lambda \beta)_{LR}, & \lambda > 0 \end{cases}$

- b) $M \oplus N = (m, \alpha, \beta)_{LR} \odot (n, \gamma, \delta)_{LR} = (m + n, \alpha + \gamma, \beta + \delta)_{LR}$.

- c) If $N = (n, \gamma, \delta)_{RL}$, then

$$M \ominus N = (m, \alpha, \beta)_{LR} \odot (-n, \delta, \gamma)_{LR} = (m - n, \alpha + \delta, \beta + \gamma)_{LR}.$$ 

3. A New Metric on $F_{LR}(R)$

In this section, we define a new metric on the set of all LR fuzzy numbers, with specified shape functions $L$ and $R$, $F_{LR}(R)$, based on $L_1$-norm. We will employ this metric to construct the fuzzy regression model in the next section.

**Definition 3.1.** The metric $D_{LR}$ on $F_{LR}(R)$ is defined by

$$D_{LR}(X,Y) = \frac{1}{3}\left\{ |m_x - m_y| + |(m_x - l\alpha_x) - (m_y - l\alpha_y)| \\ + |(m_x + r\beta_x) - (m_y + r\beta_y)| \right\}$$

(2)

where, $l = \int_0^1 L^{-1}(w)dw$, $r = \int_0^1 R^{-1}(w)dw$, $X = (x, \alpha_x, \beta_x)_{LR}$, and $Y = (y, \alpha_y, \beta_y)_{LR}$.

**Theorem 3.2.** ($F_{LR}(R), D_{LR}$) is a metric space.

**Proof.** It is sufficient to show that

- (1) $\forall X, Y \in F_{LR}(R)$, $D_{LR}(X,Y) > 0$,
- (2) $D_{LR}(X,Y) = 0 \iff X = Y$,
- (3) $\forall X, Y \in F_{LR}(R)$, $D_{LR}(X,Y) = D_{LR}(Y,X)$,
- (4) $\forall X, Y, Z \in F_{LR}(R)$, $D_{LR}(X,Z) \leq D_{LR}(X,Y) + D_{LR}(Y,Z)$.

The first three items are obviously held. We prove the triangular inequality, i.e. item (4). We have

$$\frac{1}{3}|m_x - m_z| \leq \frac{1}{3}|m_x - m_y| + \frac{1}{3}|m_y - m_z|,$$
\[ \frac{1}{2}(m_x - \alpha_x) - (m_y - 2 \alpha_y) \leq \frac{1}{2}(m_x - \alpha_x) - (m_y - \alpha_y) + \frac{1}{2}(m_y - \alpha_y) - (m_z - \alpha_z), \]
\[ \frac{1}{2}(m_x + \alpha_x) - (m_y + \beta_y) \leq \frac{1}{2}(m_x + \alpha_x) - (m_y + \beta_y) + \frac{1}{2}(m_y + \beta_y) - (m_z + \beta_z). \]

Now, we add the left sides and also the right sides of the above three inequalities to obtain the triangular inequality. \( \square \)

**Theorem 3.3.** \( (F_{LR}(\mathbb{R}), D_{LR}) \) is a complete metric space.

**Proof.** We must show that any Cauchy sequence in \( F_{LR}(\mathbb{R}) \) converges to a member of \( F_{LR}(\mathbb{R}) \). Let \( \{X_n; n \geq 1\} \) be a Cauchy sequence in \( F_{LR}(\mathbb{R}) \) where \( X_n = (m_{x_n}, \alpha_{x_n}, \beta_{x_n})_{LR} \). Then \( \frac{1}{2}|m_{x_n} - m_{x_n'}| \leq D_{LR}(X_n, X_n') \to 0 \) as \( n, n' \to \infty \).

Thus, \( \{m_{x_n}\}_{n=1}^{\infty} \) is a Cauchy sequence in \( \mathbb{R} \) and \( m_{x_n} \to m_x \), as \( n \to \infty \). On the other hand, \( \frac{1}{2}|m_{x_n} - \alpha_{x_n}| \leq D_{LR}(X_n, X_n') \to 0 \) as \( n, n' \to \infty \), and so \( \{m_{x_n} - \alpha_{x_n}\}_{n=1}^{\infty} \) is a Cauchy sequence in \( \mathbb{R} \). Since \( \{m_{x_n}\}_{n=1}^{\infty} \) is a Cauchy sequence and \( l \) is constant, it follows that \( \{\alpha_{x_n}\}_{n=1}^{\infty} \) is also a Cauchy sequence in \( \mathbb{R} \), and \( \alpha_{x_n} \to \alpha_x \), as \( n \to \infty \). Similarly, \( \beta_{x_n} \to \beta_x \), as \( n \to \infty \).

Therefore, \( D_{LR}(X_n, X) = \frac{1}{2}|m_x - m_x| + |(m_{x_n} - \alpha_{x_n}) - (m_x - \alpha_x)| + |(m_{x_n} + \beta_{x_n}) - (m_x + \beta_x)| \xrightarrow{n \to \infty} 0. \)

Thus, \( X_n \xrightarrow{n \to \infty} X \), where \( X = (m_x, \alpha_x, \beta_x)_{LR} \in F_{LR}(\mathbb{R}) \). \( \square \)

**Remark 3.4.** In special cases, when two fuzzy numbers reduce to crisp numbers, i.e. \( \alpha_x = \beta_x = \alpha_y = \beta_y = 0 \), the metric \( D_{LR} \) reduces to the ordinary metric in \( \mathbb{R} \); that is, \( d(m_x, m_y) = |m_x - m_y| \).

**Remark 3.5.** Yang and Ko [26-28] have introduced a metric on \( F_{LR}(\mathbb{R}) \) similar to the metric \( D_{LR} \), but based on \( L_2 \)-norm, (i.e. squared distance). They have applied the metric for studying some fuzzy regression methods and fuzzy clustering, (see Coppi et al. [11] for a generalization of their method to the case with \( LR \) fuzzy response.) It should be mentioned that our metric and the Yang and Ko one, have an advantage with respect to the Diamond’s metric [12] in which the shape of fuzzy numbers has not been considered in determination of the distance between two fuzzy numbers.

4. Fuzzy Least Absolutes Deviations Regression

In this section, based on the metric introduced in section 3, we propose and investigate a new least absolutes deviations approach to fuzzy regression modeling, for fuzzy input and fuzzy output data in which the parameters of the model are assumed to be crisp numbers. We will assume that all fuzzy data are symmetric LR fuzzy numbers.

Consider the set of observed data \( \{(X_i, Y_i) : i = 1, \ldots, n\} \), where \( X_i = (1, X_{i1}, X_{i2}, \ldots, X_{ip}) \) and \( X_{ij} = (x_{ij}, s_{x_{ij}})_{LL}, i = 1, \ldots, n, j = 1, \ldots, p \). Also, \( Y_i = (y_{ij}, s_{y_{ij}})_{LL}. \)

Our aim is to fit a fuzzy linear regression model with crisp coefficients to the aforementioned data set, as

\[ \hat{Y}_i = a_0 \oplus (a_1 X_{i1}) \oplus (a_2 X_{i2}) \oplus \cdots \oplus (a_p X_{ip}) \oplus E, \]

where \( E = (0, \alpha, \beta)_{LL} \) is the error term. Using Proposition 2.4, we obtain the following form for \( \hat{Y}_i \).
\[
\hat{Y}_i = (\sum_{j=0}^{p} a_j x_{ij}) + \sum_{j=1}^{p} |a_j| x_{ij} + \alpha \sum_{j=1}^{p} |a_j| x_{ij} + \beta)_{LL}, \quad i = 1, \ldots, n.
\]

where \( x_{i0} = 1 \). To find the best fuzzy linear model of the form (1), using the metric \( D_{LR} \), we minimize the sum of distances between \( Y_i \) and \( \hat{Y}_i \), i.e., \( \sum_{i=1}^{n} D_{LR}(Y_i, \hat{Y}_i) \), which is equivalent to minimize the following expression

\[
\sum_{i=1}^{n} \frac{1}{3} (|y_i - \sum_{j=0}^{p} a_j x_{ij}| + |y_i - \hat{y}_i| - (\sum_{j=1}^{p} |a_j| x_{ij} + \alpha))
\]

\[+ (|y_i + \hat{y}_i| - \sum_{j=1}^{p} a_j x_{ij} + (\sum_{j=1}^{p} |a_j| x_{ij} + \beta))).
\]

By solving this minimization problem, the crisp coefficients and the parameters \( \alpha \) and \( \beta \) are called the positive and negative deviation variables, respectively. The reason is that if \( y_i - \sum_{j=0}^{p} a_j x_{ij} \geq 0 \), then \( d_{iM}^+ = y_i - \sum_{j=0}^{p} a_j x_{ij} \) and \( d_{iM}^- = 0 \), and if \( y_i - \sum_{j=0}^{p} a_j x_{ij} < 0 \), then \( d_{iM}^+ = 0 \) and \( d_{iM}^- = -(y_i - \sum_{j=0}^{p} a_j x_{ij}) \). In other words, at least one of these variables will be zero. Similarly, by defining nonnegative deviation variables \( d_{iL}^+ \), \( d_{iL}^- \) for the second absolute expression in (4), and \( d_{iR}^+ \), \( d_{iR}^- \) for the third one, we can reformulate these two expressions as follows

\[
(|y_i - \sum_{j=0}^{p} a_j x_{ij} - (\sum_{j=0}^{p} a_j x_{ij} - \sum_{j=1}^{p} |a_j| x_{ij} + \alpha)|) = d_{iL}^+ + d_{iL}^-
\]

\[
(|y_i - \sum_{j=0}^{p} a_j x_{ij} - (\sum_{j=0}^{p} a_j x_{ij} - \sum_{j=1}^{p} |a_j| x_{ij} + \alpha)|) = d_{iL}^+ + d_{iL}^-
\]

\[
(|y_i + \sum_{j=0}^{p} a_j x_{ij} + (\sum_{j=0}^{p} a_j x_{ij} + \sum_{j=1}^{p} |a_j| x_{ij} + \beta)|) = d_{iR}^+ + d_{iR}^-
\]

Finally, the problem of minimizing the expression (4) is reformed to the following mathematical programming problem

\[
\begin{align*}
\min \quad & \sum_{i=1}^{n} \frac{1}{3} (d_{iM}^+ + d_{iM}^- + d_{iL}^+ + d_{iL}^- + d_{iR}^+ + d_{iR}^-)

\text{s.t.} \quad & y_i - \sum_{j=0}^{p} a_j x_{ij} = d_{iM}^+ - d_{iM}^-, \\
& (y_i - \sum_{j=0}^{p} a_j x_{ij} - (\sum_{j=0}^{p} a_j x_{ij} - \sum_{j=1}^{p} |a_j| x_{ij} + \alpha)) = d_{iL}^+ - d_{iL}^-, \\
& (y_i + \sum_{j=0}^{p} a_j x_{ij} + (\sum_{j=0}^{p} a_j x_{ij} + \sum_{j=1}^{p} |a_j| x_{ij} + \beta)) = d_{iR}^+ - d_{iR}^-,
\end{align*}
\]
4.1. Extension to Non-Symmetric Fuzzy Observations. In this part, we extend the proposed method to the cases in which the observations are non-symmetric fuzzy numbers. Consider the aforementioned data set in which 

\[
X_{ij} = (x_{ij}, \alpha_{x_{ij}}, \beta_{x_{ij}})_{LL}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, p \text{,}
\]

\[
Y_i = (y_i, \alpha_{y_i}, \beta_{y_i})_{LL}, \quad i = 1, \ldots, n, \text{ and } E = (0, \alpha, \beta)_{LL}.
\]

According to Propositions 2.4, the spreads of multiplication of any real number \(a_j\) and a (triangular) fuzzy number depend on the sign of \(a_j\). Suppose that we know the signs of coefficients, then we can determine the sets \(P = \{j: a_j > 0, j = 1, \ldots, p\}\) and \(N = \{j: a_j < 0, j = 1, \ldots, p\}\). Thus, we can write

\[
Y_i = \left( \sum_{j=0}^{p} a_j x_{ij}, \sum_{j=0}^{p} a_j \alpha_{x_{ij}} - \sum_{j=0}^{p} a_j \beta_{x_{ij}} + \alpha, \sum_{j=0}^{p} a_j \beta_{x_{ij}} - \sum_{j=0}^{p} a_j \alpha_{x_{ij}} + \beta \right)_{LL},
\]

The linear programming problem, therefore, is obtained similar to the problem (5)

\[
\begin{align*}
\min \quad & \sum_{i=1}^{n} \left( d^M_i + d^M_i - d^L_i + d^L_i - d^{R+}_i + d^{R-}_i \right) \\
\text{s.t.} \quad & y_i - \sum_{j=0}^{p} a_j x_{ij} = d^M_i - d^M_i, \\
& (y_i - l \alpha_{y_i}) - (\sum_{j=0}^{p} a_j x_{ij} - l (\sum_{j=0}^{p} a_j \alpha_{x_{ij}} - \sum_{j=0}^{p} a_j \beta_{x_{ij}} + \alpha)) = d^L_i - d^L_i, \\
& (y_i + l \beta_{y_i}) - (\sum_{j=0}^{p} a_j x_{ij} + l (\sum_{j=0}^{p} a_j \beta_{x_{ij}} - \sum_{j=0}^{p} a_j \alpha_{x_{ij}} + \beta)) = d^{R+}_i - d^{R-}_i, \\
& d^M_i, d^L_i, d^{R+}_i, d^{R-}_i \geq 0, \quad i = 1, \ldots, n, \\
& a_j \in \mathbb{R}, \quad j = 0, \ldots, p, \quad \alpha, \beta \geq 0.
\end{align*}
\]

Remark 4.1. Our proposal to determine the signs of \(a_j\)'s is as follows:

We could employ the classic least absolutes deviations regression, using the modes of fuzzy explanatory (inputs) and fuzzy response (output) variables. The modes of fuzzy numbers have a degree of membership equal to one and have the most importance, so that they could determine the major trend of regression line (plane). By estimating the coefficients, we then know the trend between the fuzzy response variable and \(j\)th explanatory variable, i.e., the sign of \(a_j\) (similar procedure has been employed by some authors, see for example [8]).

Sometimes, the signs of some estimated \(a_j\)’s are different from the predetermined one’s. This often occurs when the corresponding estimated coefficients from classical regression are almost zero. In this situation, with respect to the sign of the new estimated \(a_j\)’s, one can swap \(j\) in \(P\) or \(N\), then solve the new optimization problem. The procedure is continued until the sign of all estimated \(a_j\)’s are the same as their predetermined signs in \(P\) and \(N\).
5. Goodness of Fit

In this section, first, we investigate an index of goodness of fit introduced by Kim and Bishu [20]. Then, based on a similarity of measure between two fuzzy sets, we introduce a new criterion as goodness of fit index for evaluating the fuzzy regression models. In addition, we employ another goodness of fit measure, introduced and applied by Chen and Hsueh [7].

5.1. Error Index. A popular criterion in the literature, for evaluating fuzzy regression models is the Kim and Bishu’s error criterion [20], which is based on the absolute differences between the membership functions of the estimated and observed fuzzy responses.

Definition 5.1. For the fuzzy linear regression model (1), let $Y_i$ and $\hat{Y}_i$ be the observed and estimated fuzzy response for the $i$th observation, respectively. Then, the related error measure is defined by

$$E_i = \frac{\int_{S_{Y_i} \cup S_{\hat{Y}_i}} |\hat{Y}_i(x) - Y_i(x)| \, dx}{\int_{S_{Y_i}} Y_i(x) \, dx}$$

where, $Y_i(x)$ and $\hat{Y}_i(x)$ are the membership functions of $Y_i$ and $\hat{Y}_i$, respectively, and $S_{Y_i}$ and $S_{\hat{Y}_i}$ are their supports.

Definition 5.2. For the fuzzy regression model (1), mean of the errors between estimated and observed values, as a measure for goodness of fit of the model, is defined by

$$ME = \frac{1}{n} \sum_{i=1}^{n} E_i.$$ 

5.2. Similarity Measure. Similarity measures (in some literature: capability indices) are used to measure the similarity between fuzzy sets [24, 27]. In this work, we use the following index to evaluate the goodness of fit of fuzzy regression models.

Definition 5.3. [24] Suppose that $A$ and $B$ are two fuzzy numbers. The similarity measure between $A$ and $B$ is defined by

$$S_{UI}(A,B) = \frac{\text{Card}(A \cap B)}{\text{Card}(A \cup B)}$$

where

$$\text{Card}(A) = \int_{\mathbb{R}} A(x) \, dx$$

in which the ”min” operator is used for intersection of two fuzzy sets and the ”max” operator is used for the union of them.

Theorem 5.4. Let $A$, $B$, and $C$ be three fuzzy numbers. Then

1. $0 \leq S_{UI}(A, B) \leq 1$,
2. $S_{UI}(A, B) = S_{UI}(B, A)$,
3. $\forall x, (A \cap B)(x) = 0 \iff S_{UI}(A, B) = 0$, 

(4) \( A = B \iff S_{U1}(A, B) = 1 \).
(5) \( A \subseteq B \subseteq C \implies S_{U1}(A, C) \leq \min\{S_{U1}(A, B), S_{U1}(B, C)\} \).

Proof. (1) and (2) are obvious. We prove the other items.

3) \((\iff)\) is trivial.

\((\subseteq)\) Let \( S_{U1}(A, B) = 0 \), then \( \text{Card}(A \cap B) = 0 \) and so in the continuous case (the discrete case is similar)

\[
\int_{\mathbb{R}} (A \cap B)(x) dx = 0.
\]

Since \( (A \cap B)(x) \geq 0 \), from the above equality we have for all \( x \), \( (A \cap B)(x) = 0 \).

4) \((\iff)\) is trivial.

\((\subseteq)\) Let \( S_{U1}(A, B) = 1 \), then \( \text{Card}(A \cap B) = \text{Card}(A \cup B) \), and so in the continuous case (the discrete case is similar)

\[
\int_{\mathbb{R}} [(A \cup B)(x) - (A \cap B)(x)] dx = 0.
\]

But, \( (A \cap B) \subseteq (A \cup B) \) and so \( (A \cap B) = (A \cup B) \) and hence \( A = B \).

5) From \( A \subseteq B \subseteq C \), it follows that \( A \cup C = B \cup C \) and \( A \cap C \subseteq B \cap C \). Thus

\[
S_{U1}(A, C) = \frac{\text{Card}(A \cap C)}{\text{Card}(A \cup C)} \leq \frac{\text{Card}(B \cap C)}{\text{Card}(A \cup C)} = \frac{\text{Card}(B \cap C)}{\text{Card}(B \cup C)} = S_{U1}(B, C).
\]

On the other hand, from \( A \subseteq B \subseteq C \), we have \( A \cup B \subseteq A \cup C \) and \( A \cap C = A \cap B \). Thus

\[
S_{U1}(A, C) = \frac{\text{Card}(A \cap C)}{\text{Card}(A \cup C)} = \frac{\text{Card}(A \cap B)}{\text{Card}(A \cup C)} \leq \frac{\text{Card}(A \cap B)}{\text{Card}(A \cup B)} = S_{U1}(A, B).
\]

So, \( S_{U1}(A, C) \leq \min\{S_{U1}(A, B), S_{U1}(B, C)\} \). \( \square \)

Definition 5.5. For the fuzzy linear regression model (1), the mean of similarity measures (MSM) is defined by

\[
\text{MSM} = \frac{1}{n} \sum_{i=1}^{n} S_{U1}(Y_i, \hat{Y}_i).
\]

Note that, \( 0 \leq \text{MSM} \leq 1 \). We will use \( \text{MSM} \) index to evaluate the goodness of fit of fuzzy regression models.

5.3. Distance Criterion. Concerning the weakness of Choi and Buckley’s error index in reflecting the distance between fuzzy numbers in the case of no intersection, Chen and Hsueh [7] proposed a distance criterion based on \( \alpha \)-levels.
Definition 5.6. For the fuzzy regression model (1), let $D_i$ be the average error for the $i$th case by $m$ $\alpha$-levels, as

$$D_i = \frac{1}{2m} \sum_{k=1}^{m} (|\hat{Y}^L_{\alpha_k} - Y^L_{\alpha_k}| + |\hat{Y}^U_{\alpha_k} - Y^U_{\alpha_k}|)$$

where, for a certain $\alpha$-level, $(\hat{Y}^L_{\alpha_k})$ and $(\hat{Y}^U_{\alpha_k})$ are the lower and upper bounds of the $i$th estimated fuzzy response, and $(Y^L_{\alpha_k})$ and $(Y^U_{\alpha_k})$ are the lower and upper bounds for $i$th observed fuzzy response. Then, the $MD$ index, as a measure of goodness of fit for the model (1), is defined by

$$MD = \frac{1}{n} \sum_{i=1}^{n} D_i.$$

Chen and Hsueh [7] applied only two $\alpha$-levels, $\alpha = 0, 1$, for obtaining their models and comparing some fuzzy models. In the next section, we will use these two $\alpha$-levels for comparing different fuzzy regression models.

6. Numerical Examples and Comparison Studies

In this section, we illustrate our proposed approach and its performances using some well-known data sets. In addition, we compare our approach with several existing fuzzy regression approaches based on the criteria explained in section 5.

First, we consider an example in which the observations of input variables are crisp, but the observations of the dependent variable are fuzzy. In the second example, we will consider the case in which both sets of observations are fuzzy. In the third example, we use a large non-symmetric data set.

Example 6.1. The real data set in Table 1 is used by Kim and Bishu [20]. The observations of independent variables are crisp but the observations of the dependent variable are presented as symmetric triangular fuzzy numbers.

Choi and Buckley [9] claimed that for modeling this data set the least absolutes fuzzy regression works better than the Kim and Bishu’s approach (which is the

<table>
<thead>
<tr>
<th>No.</th>
<th>$Y$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(5.83, 3.56)</td>
<td>2.00</td>
<td>0.00</td>
<td>15.25</td>
</tr>
<tr>
<td>2</td>
<td>(0.85, 0.52)</td>
<td>0.00</td>
<td>5.00</td>
<td>14.13</td>
</tr>
<tr>
<td>3</td>
<td>(13.93, 8.50)</td>
<td>1.13</td>
<td>1.50</td>
<td>14.13</td>
</tr>
<tr>
<td>4</td>
<td>(4.00, 2.44)</td>
<td>2.00</td>
<td>1.25</td>
<td>13.63</td>
</tr>
<tr>
<td>5</td>
<td>(1.65, 1.01)</td>
<td>2.19</td>
<td>3.75</td>
<td>14.75</td>
</tr>
<tr>
<td>6</td>
<td>(1.58, 0.96)</td>
<td>0.25</td>
<td>3.50</td>
<td>13.75</td>
</tr>
<tr>
<td>7</td>
<td>(8.18, 4.99)</td>
<td>0.75</td>
<td>5.25</td>
<td>15.25</td>
</tr>
<tr>
<td>8</td>
<td>(1.85, 1.13)</td>
<td>4.25</td>
<td>2.00</td>
<td>13.50</td>
</tr>
</tbody>
</table>

Table 1. Data Set in Example 6.1
least squares fuzzy regression method) since there is one outlier data point (No. 3). Using the Choi and Buckley’ method, the optimal model (CB) is

\[ \hat{Y}_{CB} = -2.8273 \odot 0.3878 \ x_1 \oplus 1.0125 \ x_2 \oplus 0.6185 \ x_3 \oplus (0, 1.0696, 2.0042)^T. \]

The above relation is derived based on the model 2.2 in [7], and it is noticeable this model yields more efficient results than those of based on the model 2.1 in their work.

Chen and Hsueh [8] proposed a least squares approach to fuzzy regression models with crisp coefficients. They minimized total estimation errors in terms of the sum of \( E_i \)'s, where \( E_i \) is denoted by the average squared distances (errors) between the observed and estimated fuzzy responses for the \( i \)th observation based on a few \( \alpha \)-levels. Their optimal model (CH) is

\[ \hat{Y}_{CH} = -16.7956 \odot 1.0989 \ x_1 \odot 1.1798 \ x_2 \odot 1.8559 \ x_3 \oplus (0, 2.8888)^T. \]

Hassanpour et al. [17] proposed a least absolutes regression method that minimizes the differences between centers (modes) of the observed and estimated fuzzy responses and also between the spreads of them, using a goal programming approach. They took into account fuzzy coefficients for crisp inputs in their model. Employing their method for the above data set yields the following model (HMY)

\[ \hat{Y}_{HMY} = (-2.8273, 0.3877, 0.3877)^T \ x_1 \odot (1.0125, 0.0000)^T \ x_2 \oplus (0.6185, 0.1790)^T \ x_3. \]

On the other hand, by employing the proposed method in section 4, the optimal model (TK) is obtained as

\[ \hat{Y}_{TK} = -15.5578 \odot 0.2444 \ x_1 \odot 0.9976 \ x_2 \oplus 1.5142 \ x_3 \oplus (0, 1.1300)^T. \]

Now, for evaluating the performance of our proposed model (TK), we compare it with the other models, based on the three goodness of fit criteria introduced in section 5.

First, we compared the models using the error index, \( ME \). The results are given in Table 2. It is obvious that, based on this criterion, the proposed model (TK) has a smaller \( ME \), and, therefore, it is fitted to the data better than the other models.

Second, the similarity of measures between the observed values and estimated values for different models were calculated. The results are shown in Table 3. As one can see, the \( MSM \) for the proposed model is 0.3710 which is better than the \( MSM \) for the other models.

For more comparison, we calculated the \( MD \) criterion for different models. By comparing the results given in Table 4, it is obvious that our model is better than the other models based on distance criterion.

As mentioned above, in the data set in Table 1, there exists one outlier datum. Similar to classical regression, it seems that in fuzzy regression modeling, the least absolutes deviations estimators are more suitable than the least squares ones, when there are some outliers in the data set.
In the above example, the models \((CB)\), \((HMY)\), and \((TK)\) (which are LAD-based) work better than the model \((CH)\) (which is LS-based) in terms of different goodness of fit criteria.

<table>
<thead>
<tr>
<th>No.</th>
<th>CB</th>
<th>CH</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5683</td>
<td>1.4273</td>
<td>0.2332</td>
<td>0.7648</td>
</tr>
<tr>
<td>2</td>
<td>2.6957</td>
<td>6.2552</td>
<td>4.6581</td>
<td>1.7125</td>
</tr>
<tr>
<td>3</td>
<td>1.1777</td>
<td>1.1847</td>
<td>1.2857</td>
<td>1.1329</td>
</tr>
<tr>
<td>4</td>
<td>0.4099</td>
<td>0.6258</td>
<td>0.3431</td>
<td>0.6219</td>
</tr>
<tr>
<td>5</td>
<td>1.0171</td>
<td>3.0368</td>
<td>2.2312</td>
<td>1.3494</td>
</tr>
<tr>
<td>6</td>
<td>1.3307</td>
<td>3.6758</td>
<td>1.6449</td>
<td>0.2718</td>
</tr>
<tr>
<td>7</td>
<td>1.3080</td>
<td>1.1329</td>
<td>1.5395</td>
<td>1.2264</td>
</tr>
<tr>
<td>8</td>
<td>0.4123</td>
<td>1.9204</td>
<td>1.6706</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**Table 2.** Comparison Between the Models in Example 6.1, Using the Error Index \(ME\)

<table>
<thead>
<tr>
<th>No.</th>
<th>CB</th>
<th>CH</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4317</td>
<td>0.1186</td>
<td>0.7668</td>
<td>0.2654</td>
</tr>
<tr>
<td>2</td>
<td>0.1894</td>
<td>0.0234</td>
<td>0.1146</td>
<td>0.2989</td>
</tr>
<tr>
<td>3</td>
<td>0.0013</td>
<td>0.0615</td>
<td>0.0046</td>
<td>0.0000</td>
</tr>
<tr>
<td>4</td>
<td>0.5981</td>
<td>0.5545</td>
<td>0.7072</td>
<td>0.4035</td>
</tr>
<tr>
<td>5</td>
<td>0.4252</td>
<td>0.1194</td>
<td>0.2366</td>
<td>0.2220</td>
</tr>
<tr>
<td>6</td>
<td>0.3231</td>
<td>0.0434</td>
<td>0.3684</td>
<td>0.7780</td>
</tr>
<tr>
<td>7</td>
<td>0.0000</td>
<td>0.1645</td>
<td>0.0024</td>
<td>0.0000</td>
</tr>
<tr>
<td>8</td>
<td>0.7020</td>
<td>0.2987</td>
<td>0.3052</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

**Table 3.** Comparison Between the Models in Example 6.1, Using the Index \(MSM\)

<table>
<thead>
<tr>
<th>No.</th>
<th>CB</th>
<th>CH</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0119</td>
<td>3.4791</td>
<td>0.4154</td>
<td>1.8225</td>
</tr>
<tr>
<td>2</td>
<td>0.5086</td>
<td>2.6793</td>
<td>1.0048</td>
<td>0.3051</td>
</tr>
<tr>
<td>3</td>
<td>0.7412</td>
<td>7.5132</td>
<td>9.9747</td>
<td>9.8647</td>
</tr>
<tr>
<td>4</td>
<td>0.6707</td>
<td>0.8278</td>
<td>0.4382</td>
<td>0.9825</td>
</tr>
<tr>
<td>5</td>
<td>0.2637</td>
<td>2.0981</td>
<td>0.8153</td>
<td>0.8504</td>
</tr>
<tr>
<td>6</td>
<td>0.6900</td>
<td>2.7390</td>
<td>0.9788</td>
<td>0.1499</td>
</tr>
<tr>
<td>7</td>
<td>0.9480</td>
<td>3.6913</td>
<td>7.1816</td>
<td>6.0669</td>
</tr>
<tr>
<td>8</td>
<td>0.2336</td>
<td>1.1898</td>
<td>0.6434</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**Table 4.** Comparison Between the Models in Example 6.1, Using the Index \(MD\)
Example 6.2. In this example, using a well-known data set (including fuzzy input-fuzzy output data) we perform a comparison between our proposed model with some existing models.

The data set in Table 5 used by several authors are initially presented by Sakawa and Yano [29]. In this data set, fuzzy input and fuzzy output observations are symmetric triangular fuzzy numbers.

The results of fitting several models to this data set are as follows:

Using Choi and Buckley’ method ($CB$), the following model is obtained

$$\hat{Y}_{CB} = 3.9444 \oplus 0.4444 X \oplus (0, 0.2778)_{T}.$$  

The Arabpour and Tata’s model ($AT$) [1], which is a fuzzy least squares regression model, is

$$\hat{Y}_{AT} = 3.4873 \oplus 0.5306 X.$$  

Based on the Chen and Hsueh’s method [8], the optimal model ($CH$) is obtained as

$$\hat{Y}_{CH} = 3.5750 \oplus 0.5196 X \oplus (0, 0.3006)_{T}.$$  

Finally, the Hassanpour et al.’s model ($HMY$) [18] is obtained as

$$\hat{Y}_{HMY} = 3.9444 \oplus 0.4444 X.$$  

On the other hand, based on the proposed method in section 4, the optimal model ($TK$) is obtained as

$$\hat{Y}_{TK} = 3.9444 \oplus 0.4444 X \oplus (0, 0.2778)_{T}. \quad (12)$$  

Note that the ($TK$) and ($CB$) models are the same. Also the coefficients of ($HMY$) model are similar to those of ($TK'$s) and ($CB'$s).

Now, we compare the above models based on three criteria: $ME$, $MSM$, and $MD$. The amounts of errors between observed and estimated fuzzy responses, for several models, are given in Table 6. As one can see, the mean of errors $ME$ of the model ($HMY$) is smaller than that of the other models. In what follows, we compare the above four models based on the other criteria.

The values of similarity measures for the models are given in Table 7. As one can see, the mean of similarity measures for the proposed model is much larger than
the others so that based on this criterion the proposed model ($TK$) is supposed to be the best model.

Finally, we use the values of distances between the observed and estimated fuzzy responses to calculate the $MD$ index for the models. The results are given in Table 8 indicating that, based on the distance criterion, the model ($TK$) is the best model.

Now, for more comparison as well as to study the effect of outliers, we change the fifth data as $Y = (18.5, 0.5)^T$, and reconstruct the above models. The summary of results are given in Table 9.

To compare the effect of the outlier, we calculate the difference between each criterion for all five models with and without the outlier. The results of three criteria are given in Tables 10-12.

As it can be seen, the models ($TK$) and ($CB$) are somewhat similar. Note that the models ($TK$), ($CB$), and ($HMY$), which are least absolutes regression models (LAD-based), have affected by the outlier less than the two models ($AT$) and ($CH$), which are least squares regression models (LS-based). This shows the robustness of least absolutes fuzzy regression models with respect to outliers in this data set.

<table>
<thead>
<tr>
<th>No.</th>
<th>AT</th>
<th>CH</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.4078</td>
<td>1.7431</td>
<td>1.4444</td>
<td>1.9444</td>
</tr>
<tr>
<td>2</td>
<td>0.6124</td>
<td>0.4118</td>
<td>0.5556</td>
<td>0.0008</td>
</tr>
<tr>
<td>3</td>
<td>1.4063</td>
<td>1.5112</td>
<td>1.3676</td>
<td>1.5056</td>
</tr>
<tr>
<td>4</td>
<td>1.5200</td>
<td>1.8863</td>
<td>1.3672</td>
<td>1.6043</td>
</tr>
<tr>
<td>5</td>
<td>1.3502</td>
<td>1.5565</td>
<td>1.4444</td>
<td>1.9016</td>
</tr>
<tr>
<td>6</td>
<td>1.3850</td>
<td>1.4722</td>
<td>0.9631</td>
<td>1.0047</td>
</tr>
<tr>
<td>7</td>
<td>1.5300</td>
<td>2.1202</td>
<td>1.4444</td>
<td>2.0000</td>
</tr>
<tr>
<td>8</td>
<td>1.4753</td>
<td>1.6531</td>
<td>0.5556</td>
<td>0.0024</td>
</tr>
<tr>
<td>$ME$</td>
<td>1.3360</td>
<td>1.5443</td>
<td>1.1428</td>
<td>1.2455</td>
</tr>
</tbody>
</table>

Table 6. Comparison Between Different Models in Example 6.2, Using the Error Index $ME$

<table>
<thead>
<tr>
<th>No.</th>
<th>AT</th>
<th>CH</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0431</td>
<td>0.0976</td>
<td>0.0000</td>
<td>0.0141</td>
</tr>
<tr>
<td>2</td>
<td>0.4239</td>
<td>0.6747</td>
<td>0.4444</td>
<td>1.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.0417</td>
<td>0.0926</td>
<td>0.0273</td>
<td>0.0672</td>
</tr>
<tr>
<td>4</td>
<td>0.0039</td>
<td>0.0584</td>
<td>0.0275</td>
<td>0.1099</td>
</tr>
<tr>
<td>5</td>
<td>0.0608</td>
<td>0.1533</td>
<td>0.0000</td>
<td>0.0253</td>
</tr>
<tr>
<td>6</td>
<td>0.0508</td>
<td>0.1055</td>
<td>0.1999</td>
<td>0.2630</td>
</tr>
<tr>
<td>7</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>8</td>
<td>0.0193</td>
<td>0.1238</td>
<td>0.4444</td>
<td>0.9976</td>
</tr>
<tr>
<td>$MSM$</td>
<td>0.0804</td>
<td>0.1632</td>
<td>0.1429</td>
<td>0.3095</td>
</tr>
</tbody>
</table>

Table 7. Comparison Between Different Models in Example 6.2, Using the Index $MSM$
<table>
<thead>
<tr>
<th>No.</th>
<th>AT</th>
<th>CH</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5485</td>
<td>0.6130</td>
<td>0.8332</td>
<td>0.8332</td>
</tr>
<tr>
<td>2</td>
<td>0.1951</td>
<td>0.1085</td>
<td>0.1390</td>
<td>0.0002</td>
</tr>
<tr>
<td>3</td>
<td>1.0944</td>
<td>1.0705</td>
<td>1.1114</td>
<td>1.1114</td>
</tr>
<tr>
<td>4</td>
<td>0.7015</td>
<td>0.7080</td>
<td>0.5552</td>
<td>0.5552</td>
</tr>
<tr>
<td>5</td>
<td>0.5026</td>
<td>0.5135</td>
<td>0.7782</td>
<td>0.7782</td>
</tr>
<tr>
<td>6</td>
<td>1.0586</td>
<td>1.0245</td>
<td>0.6106</td>
<td>0.6106</td>
</tr>
<tr>
<td>7</td>
<td>1.1761</td>
<td>1.2160</td>
<td>1.6672</td>
<td>1.6672</td>
</tr>
<tr>
<td>8</td>
<td>0.6198</td>
<td>0.5625</td>
<td>0.1392</td>
<td>0.0006</td>
</tr>
<tr>
<td>MD</td>
<td>0.7371</td>
<td>0.7271</td>
<td>0.7292</td>
<td>0.6946</td>
</tr>
</tbody>
</table>

Table 8. Comparison Between Different Models in Example 6.2, Using the Index $MD$

\[
\hat{Y}_{CB} = 3.9444 \oplus 0.4444 \times (0, 0.2772, 0.2780)_{T}
\]

\[
\hat{Y}_{AT} = 4.0290 \oplus 0.6242 \times X
\]

\[
\hat{Y}_{CH} = 4.0915 \oplus 0.6160 \times (0, 0.2400)_{T}
\]

\[
\hat{Y}_{HMY} = 3.9444 \oplus 0.4444 \times X
\]

\[
\hat{Y}_{TK} = 3.9444 \oplus 0.4444 \times (0, 0.2778)_{T}
\]

Table 9. The Obtained Models for the Data Set with Outlier in Example 6.2

<table>
<thead>
<tr>
<th>Model</th>
<th>CB</th>
<th>CH</th>
<th>AT</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>without outlier</td>
<td>1.2455</td>
<td>1.5443</td>
<td>1.3360</td>
<td>1.1428</td>
<td>1.2455</td>
</tr>
<tr>
<td>with outlier</td>
<td>1.2576</td>
<td>1.6903</td>
<td>1.4120</td>
<td>1.1428</td>
<td>1.2455</td>
</tr>
<tr>
<td>absolute difference</td>
<td>0.0121</td>
<td>0.1460</td>
<td>0.0760</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 10. The Values of $ME$ for the Models in Example 6.2

<table>
<thead>
<tr>
<th>Model</th>
<th>CB</th>
<th>CH</th>
<th>AT</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>without outlier</td>
<td>0.3095</td>
<td>0.1632</td>
<td>0.0804</td>
<td>0.1429</td>
<td>0.3095</td>
</tr>
<tr>
<td>with outlier</td>
<td>0.3064</td>
<td>0.1454</td>
<td>0.0976</td>
<td>0.1429</td>
<td>0.3095</td>
</tr>
<tr>
<td>absolute difference</td>
<td>0.0031</td>
<td>0.0178</td>
<td>0.0172</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 11. The Values of $MSM$ for the Models in Example 6.2

<table>
<thead>
<tr>
<th>Model</th>
<th>CB</th>
<th>CH</th>
<th>AT</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>without outlier</td>
<td>0.6946</td>
<td>0.7271</td>
<td>0.7371</td>
<td>0.7292</td>
<td>0.6946</td>
</tr>
<tr>
<td>with outlier</td>
<td>1.9452</td>
<td>2.3060</td>
<td>2.3214</td>
<td>1.9792</td>
<td>1.9446</td>
</tr>
<tr>
<td>absolute difference</td>
<td>1.2506</td>
<td>1.5789</td>
<td>1.5843</td>
<td>1.2500</td>
<td>1.2500</td>
</tr>
</tbody>
</table>

Table 12. The Values of $MD$ for the Models in Example 6.2
Table 13. Data Set in Example 6.3

<table>
<thead>
<tr>
<th>No.</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(50, 8, 8)</td>
<td>(98, 6, 2)</td>
<td>(71, 9, 11)</td>
<td>(70, 11, 13)</td>
<td>(30, 11, 9)</td>
</tr>
<tr>
<td>2</td>
<td>(29, 8, 8)</td>
<td>(76, 6, 2)</td>
<td>(61, 9, 11)</td>
<td>(46, 11, 13)</td>
<td>(20, 13, 10)</td>
</tr>
<tr>
<td>3</td>
<td>(41, 8, 8)</td>
<td>(88, 6, 2)</td>
<td>(73, 9, 11)</td>
<td>(58, 11, 13)</td>
<td>(25, 11, 12)</td>
</tr>
<tr>
<td>4</td>
<td>(60, 9, 7)</td>
<td>(62, 9, 10)</td>
<td>(79, 9, 6)</td>
<td>(66, 8, 9)</td>
<td>(43, 11, 9)</td>
</tr>
<tr>
<td>5</td>
<td>(49, 9, 7)</td>
<td>(50, 9, 10)</td>
<td>(75, 9, 6)</td>
<td>(54, 8, 9)</td>
<td>(38, 12, 8)</td>
</tr>
<tr>
<td>6</td>
<td>(59, 9, 7)</td>
<td>(60, 9, 10)</td>
<td>(85, 9, 6)</td>
<td>(64, 8, 9)</td>
<td>(43, 11, 9)</td>
</tr>
<tr>
<td>7</td>
<td>(61, 9, 11)</td>
<td>(77, 8, 6)</td>
<td>(85, 5, 8)</td>
<td>(18, 7, 13)</td>
<td>(40, 17, 11)</td>
</tr>
<tr>
<td>8</td>
<td>(58, 9, 11)</td>
<td>(75, 8, 6)</td>
<td>(82, 5, 8)</td>
<td>(16, 7, 13)</td>
<td>(38, 11, 12)</td>
</tr>
<tr>
<td>9</td>
<td>(55, 9, 11)</td>
<td>(72, 8, 6)</td>
<td>(79, 5, 8)</td>
<td>(13, 7, 13)</td>
<td>(37, 12, 12)</td>
</tr>
<tr>
<td>10</td>
<td>(66, 8, 7)</td>
<td>(59, 17, 11)</td>
<td>(39, 8, 9)</td>
<td>(87, 14, 11)</td>
<td>(59, 10, 9)</td>
</tr>
<tr>
<td>11</td>
<td>(69, 8, 7)</td>
<td>(63, 17, 11)</td>
<td>(49, 8, 9)</td>
<td>(87, 14, 11)</td>
<td>(59, 10, 9)</td>
</tr>
<tr>
<td>12</td>
<td>(59, 8, 7)</td>
<td>(53, 17, 11)</td>
<td>(39, 8, 9)</td>
<td>(77, 14, 11)</td>
<td>(54, 11, 8)</td>
</tr>
<tr>
<td>13</td>
<td>(74, 4, 6)</td>
<td>(89, 11, 5)</td>
<td>(70, 12, 13)</td>
<td>(82, 14, 10)</td>
<td>(61, 14, 3)</td>
</tr>
<tr>
<td>14</td>
<td>(41, 4, 6)</td>
<td>(57, 11, 5)</td>
<td>(58, 12, 13)</td>
<td>(50, 14, 10)</td>
<td>(34, 10, 8)</td>
</tr>
<tr>
<td>15</td>
<td>(49, 4, 6)</td>
<td>(65, 11, 5)</td>
<td>(66, 12, 13)</td>
<td>(58, 14, 10)</td>
<td>(38, 9, 9)</td>
</tr>
<tr>
<td>16</td>
<td>(76, 8, 7)</td>
<td>(75, 10, 8)</td>
<td>(37, 8, 11)</td>
<td>(75, 5, 10)</td>
<td>(64, 16, 9)</td>
</tr>
<tr>
<td>17</td>
<td>(57, 8, 7)</td>
<td>(56, 10, 8)</td>
<td>(18, 8, 11)</td>
<td>(56, 5, 10)</td>
<td>(56, 13, 7)</td>
</tr>
<tr>
<td>18</td>
<td>(72, 8, 7)</td>
<td>(71, 10, 8)</td>
<td>(33, 8, 11)</td>
<td>(71, 5, 10)</td>
<td>(63, 11, 9)</td>
</tr>
<tr>
<td>19</td>
<td>(78, 7, 8)</td>
<td>(65, 6, 6)</td>
<td>(82, 11, 11)</td>
<td>(64, 8, 12)</td>
<td>(66, 16, 5)</td>
</tr>
<tr>
<td>20</td>
<td>(58, 7, 8)</td>
<td>(45, 6, 6)</td>
<td>(62, 11, 11)</td>
<td>(44, 8, 12)</td>
<td>(49, 12, 9)</td>
</tr>
<tr>
<td>21</td>
<td>(72, 7, 8)</td>
<td>(59, 6, 6)</td>
<td>(76, 11, 11)</td>
<td>(58, 8, 12)</td>
<td>(55, 10, 12)</td>
</tr>
<tr>
<td>22</td>
<td>(90, 8, 5)</td>
<td>(95, 13, 3)</td>
<td>(80, 11, 8)</td>
<td>(72, 7, 13)</td>
<td>(67, 11, 14)</td>
</tr>
<tr>
<td>23</td>
<td>(68, 8, 5)</td>
<td>(73, 13, 3)</td>
<td>(58, 11, 8)</td>
<td>(50, 7, 13)</td>
<td>(53, 10, 9)</td>
</tr>
<tr>
<td>24</td>
<td>(71, 8, 5)</td>
<td>(76, 13, 3)</td>
<td>(61, 11, 8)</td>
<td>(53, 7, 13)</td>
<td>(54, 9, 10)</td>
</tr>
<tr>
<td>25</td>
<td>(92, 8, 6)</td>
<td>(76, 6, 9)</td>
<td>(78, 10, 6)</td>
<td>(27, 9, 15)</td>
<td>(70, 13, 7)</td>
</tr>
<tr>
<td>26</td>
<td>(94, 8, 6)</td>
<td>(78, 6, 9)</td>
<td>(80, 10, 6)</td>
<td>(29, 9, 15)</td>
<td>(68, 9, 10)</td>
</tr>
<tr>
<td>27</td>
<td>(87, 8, 6)</td>
<td>(71, 6, 9)</td>
<td>(73, 10, 6)</td>
<td>(22, 9, 15)</td>
<td>(65, 10, 9)</td>
</tr>
<tr>
<td>28</td>
<td>(94, 6, 5)</td>
<td>(51, 9, 8)</td>
<td>(30, 9, 11)</td>
<td>(29, 9, 16)</td>
<td>(75, 5, 14)</td>
</tr>
<tr>
<td>29</td>
<td>(95, 6, 5)</td>
<td>(52, 9, 8)</td>
<td>(31, 9, 11)</td>
<td>(30, 9, 16)</td>
<td>(84, 10, 7)</td>
</tr>
<tr>
<td>30</td>
<td>(86, 6, 5)</td>
<td>(43, 9, 8)</td>
<td>(22, 9, 11)</td>
<td>(21, 9, 16)</td>
<td>(80, 12, 6)</td>
</tr>
</tbody>
</table>

Example 6.3. In this example, we provide a comparison study using a large data set (see Table 13). This data set, that is presented by Chen and Hsueh [8], contains four input variables and one output variable for which all observations are non-symmetric triangular fuzzy numbers. Chen and Hsueh [8] performed a fuzzy least squares regression model. Their model ($CH$) for this data set is obtained as

$$
\hat{Y}_{CH} = 12.093 \oplus 0.859X_1 \oplus (-0.207)X_2 \oplus (-0.134)X_3 \oplus 0.108X_4 \\
\oplus (0, 1.299, 0.039)_T.
$$

Applying the Choi and Buckley’s method results in the following model ($CB$)

$$
\hat{Y}_{CB} = 14.9668 \oplus 0.8482X_1 \oplus (-0.1515)X_2 \oplus (-0.1937)X_3 \oplus 0.0526X_4 \\
\oplus (0, 0.0059, 0.0053)_T.
$$
Table 14. Goodness of Fit Criteria in Example 6.3

<table>
<thead>
<tr>
<th>Model</th>
<th>CB</th>
<th>CH</th>
<th>HMY</th>
<th>TK</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME</td>
<td>0.2333</td>
<td>0.2736</td>
<td>0.2406</td>
<td>0.2140</td>
</tr>
<tr>
<td>MSM</td>
<td>0.8227</td>
<td>0.7746</td>
<td>0.7936</td>
<td>0.8210</td>
</tr>
<tr>
<td>MD</td>
<td>1.3588</td>
<td>1.6687</td>
<td>1.4464</td>
<td>1.3570</td>
</tr>
</tbody>
</table>

In addition, the Hassanpour et al.’s method yields the following model (HMY)

\[ \hat{Y}_{HMY} = 20.3559 \oplus 0.8197X_1 \oplus (-0.1661)X_2 \oplus (-0.2144)X_3 \oplus 0.0351X_4. \]

For applying our proposed method to this data set, we first performed a least absolute regression for the modes of input and output data. The estimated intercept and coefficients were obtained as 21.14, 0.82, -0.18, -0.21, and 0.036, respectively. Therefore, by considering the signs of the coefficients in program (6), the optimal model (TK) is obtained as

\[ \hat{Y}_{TK} = 16.7295 \oplus 0.8464X_1 \oplus (-0.1679)X_2 \oplus (-0.1947)X_3 \oplus 0.0517X_4 \oplus (0, 0.8320, 0.5490)^T. \]

The values of goodness of fit criteria are summarized in Table 14. The amounts of MD and ME for the proposed model (TK) are smaller than those of the models (CB), (CH), and (HMY). Also, the amount of MSM for the proposed model is very close to that of the model (CB). Therefore, the TK’s method performs a better model than the other methods for this data set.

7. Conclusion

A new method, based on least absolutes deviations, is proposed for fuzzy regression modeling, using a new metric on the space of fuzzy numbers. In regard to three criteria for goodness of fit, it is shown that the proposed method performs more convenient models with respect to some well-known methods in some data sets, especially when the data set includes some outlier data point(s).

Concerning the proposed approach, the research on the other robust approaches to fuzzy regression and also investigating other methods for the outliers detection may be some more topics for the future research.

Acknowledgements. The authors would like to thank the referees for their helpful suggestions and valuable comments.

References

[27] H. Rezaei, M. Emoto and M. Mukaidono, New similarity measure between two fuzzy sets, Journal of Advanced Computational Intelligence and Intelligent Informatics, 10 (2006), 946-953.

S. M. Taheri*, Department of Mathematical Sciences, Isfahan University of Technology, Isfahan 84156-83111, Iran and Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

E-mail address: taheri@cc.iut.ac.ir

M. Kelkinnama, Department of Mathematical Sciences, Isfahan University of Technology, Isfahan 84156-83111, Iran

E-mail address: mkelkinnama@yahoo.com

*Corresponding author