Fuzzy estimates of regression parameters in linear regression models for imprecise input and output data

Hsien-Chung Wu

Department of Information Management, National Chi Nan University, Puli, Nantou 545, Taiwan

Received 1 January 2002; accepted 2 March 2002

Abstract

The method for obtaining the fuzzy estimates of regression parameters with the help of “Resolution Identity” in fuzzy sets theory is proposed. The $\alpha$-level least-squares estimates can be obtained from the usual linear regression model by using the $\alpha$-level real-valued data of the corresponding fuzzy input and output data. The membership functions of fuzzy estimates of regression parameters will be constructed according to the form of “Resolution Identity” based on the $\alpha$-level least-squares estimates. In order to obtain the membership degree of any given value taken from the fuzzy estimate, optimization problems have to be solved. Two computational procedures are also provided to solve the optimization problems. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Linear regression analysis; Fuzzy numbers; Fuzzy estimate; Optimization; Confidence degree

1. Introduction

In the real world, the data sometimes cannot be recorded or collected precisely. For instance, the water level of a river cannot be measured in an exact way because of the fluctuation, and the temperature in a room is also not able to be measured precisely because of the similar reason. Therefore, the fuzzy sets theory is naturally to be an appropriate tool in modeling the statistical models when the fuzzy data have been observed. The more appropriate way to describe the water level is to say that the water level is around 30 m. The phrase “around 30 m” can be regarded as a fuzzy number 30. This is the main concern of this paper.
Since Zadeh (1965) introduced the concept of fuzzy sets, the applications of considering fuzzy data to the regression models have been proposed in the literature. Tanaka et al. (1982) initiated this research topic. They also generalized their approaches to the more general models in Tanaka and Warada (1988), Tanaka et al. (1989), Tanaka and Ishibuchi (1991). The collection of papers edited by Kacprzyk and Fedrizzi (1992) gave an insightful survey.

indices for equalities between fuzzy numbers. From these three indices, three types of multiobjective programming problems were formulated. Tanaka and Lee (1998) used the quadratic programming approach to obtain the possibility and necessity regression models simultaneously. The advantage of adopting a quadratic programming approach is to be able to integrate both the property of central tendency in least squares and the possibilistic property in fuzzy regression.

In this paper, we will first obtain the \( \alpha \)-level least-squares estimates from the usual linear regression model by using the \( \alpha \)-level real-valued data of the corresponding fuzzy input and output data, and then construct the fuzzy estimates of regression parameters according to the form of “Resolution Identity” in fuzzy sets theory which was introduced by Zadeh (1975). In order to obtain the membership degree of any value taken from the fuzzy estimate, the optimization problems have to be solved. We also develop two computational procedures to solve the optimization problems.

In Section 2, we give some properties of fuzzy numbers. In Section 3, we obtain the \( \alpha \)-level least-squares estimates from the usual linear regression model by using the \( \alpha \)-level real-valued data of the corresponding fuzzy input and output data. The membership functions of fuzzy estimates will be constructed according to the form of “Resolution Identity” from the \( \alpha \)-level least-squares estimates obtained above. In Section 4, we develop two computational procedures to obtain the membership degree of any given value taken from the fuzzy estimates. We also provide the methodology to transact the predicted fuzzy output data. In Section 5, the numerical examples are given to clarify the theoretical results, and show the possible applications in linear regression analysis for imprecise data.

2. Fuzzy numbers

Let \( X \) be a universal set. Then a fuzzy subset \( \tilde{A} \) of \( X \) is defined by its membership function \( \tilde{\xi}_\tilde{A}:X \rightarrow [0,1] \). We denote by \( \tilde{A}_\alpha = \{ x: \tilde{\xi}_\tilde{A}(x) \geq \alpha \} \) the \( \alpha \)-level set of \( \tilde{A} \), where \( \tilde{A}_0 \) is the closure of the set \( \{ x: \tilde{\xi}_\tilde{A}(x) \neq 0 \} \). \( \tilde{A} \) is called a normal fuzzy set if there exists an \( x \) such that \( \tilde{\xi}_\tilde{A}(x) = 1 \). \( \tilde{A} \) is called a convex fuzzy set if \( \tilde{\xi}_\tilde{A}(\lambda x + (1 - \lambda)y) \geq \min\{\tilde{\xi}_\tilde{A}(x),\tilde{\xi}_\tilde{A}(y)\} \) for \( \lambda \in [0,1] \) (That is, \( \tilde{\xi}_\tilde{A} \) is a quasi-concave function.)

In this paper, the universal set \( X \) is assumed to be a real number system; that is, \( X = \mathbb{R} \). Let \( f \) be a real-valued function defined on \( \mathbb{R} \). \( f \) is said to be upper semicontinuous, if \( \{ x: f(x) \geq \alpha \} \) is a closed set for each \( \alpha \). Or equivalently, \( f \) is upper semicontinuous at \( y \) if and only if \( \forall \varepsilon > 0, \exists \delta > 0 \) such that \( |x - y| < \delta \) implies \( f(x) < f(y) + \varepsilon \).

\( \tilde{a} \) is called a fuzzy number if the following conditions are satisfied:

(i) \( \tilde{a} \) is a normal and convex fuzzy set,
(ii) Its membership function \( \tilde{\xi}_{\tilde{a}} \) is upper semicontinuous,
(iii) The \( \alpha \)-level set \( \tilde{a}_\alpha \) is bounded for each \( \alpha \in [0,1] \).

From Zadeh (1965), \( \tilde{A} \) is a convex fuzzy set if and only if its \( \alpha \)-level set \( \tilde{A}_\alpha = \{ x: \tilde{\xi}_\tilde{A}(x) \geq \alpha \} \) is a convex set for all \( \alpha \). Therefore, if \( \tilde{a} \) is a fuzzy number, then the
$z$-level set $\tilde{a}_z$ is a compact (closed and bounded in $\mathbb{R}$) and convex set; that is, $\tilde{a}$ is a closed interval. The $z$-level set of $\tilde{a}$ is then denoted by $\tilde{a}_z = [\tilde{a}_z^L, \tilde{a}_z^U]$. We also see that $\tilde{a}_z^L$ and $\tilde{a}_z^U$ are continuous with respect to $z$, since its membership function is upper semicontinuous. The following proposition is useful for further discussions.

**Proposition 1** (Resolution Identity, Zadeh, 1975). Let $\tilde{A}$ be a fuzzy set with membership function $\tilde{\alpha}_\tilde{A}$ and the $z$-level set $\tilde{A}_z = \{x : \tilde{\alpha}_\tilde{A}(x) \geq z\}$ be given. Then

$$\tilde{\alpha}_\tilde{A}(x) = \sup_{z \in [0,1]} z \tilde{1}_{\tilde{A}_z}(x).$$

$\tilde{a}$ is called a crisp number with value $m$ if its membership function is

$$\tilde{\alpha}_\tilde{a}(x) = \begin{cases} 1 & \text{if } x = m, \\ 0 & \text{otherwise.} \end{cases}$$

It is denoted by $\tilde{a} \equiv \tilde{1}_{\{m\}}$. It is easy to see that $(\tilde{1}_{\{m\}})_z^L = (\tilde{1}_{\{m\}})_z^U = m$ for all $z \in [0,1]$.

$\tilde{a}$ is called a nonnegative fuzzy number if $\tilde{\alpha}_\tilde{a}(x) = 0$ for all $x < 0$ and a nonpositive fuzzy number if $\tilde{\alpha}_\tilde{a}(x) = 0$ for all $x > 0$. It is obvious that $\tilde{a}_z^L$ and $\tilde{a}_z^U$ are nonnegative real numbers for all $z \in [0,1]$ if $\tilde{a}$ is a nonnegative fuzzy number, and $\tilde{a}_z^L$ and $\tilde{a}_z^U$ are nonpositive real numbers for all $z \in [0,1]$ if $\tilde{a}$ is a nonpositive fuzzy number.

Let $\odot$ be any binary operation $\oplus$ or $\otimes$ between two fuzzy numbers $\tilde{a}$ and $\tilde{b}$. The membership function of $\tilde{a} \odot \tilde{b}$ is defined by

$$\tilde{\alpha}_{\tilde{a} \odot \tilde{b}}(z) = \sup_{x,y} \min \{\tilde{\alpha}_\tilde{a}(x), \tilde{\alpha}_\tilde{b}(y)\}$$

using the extension principle in Zadeh (1975), where the operations $\odot = \oplus$ or $\otimes$ corresponding to the operations $\circ = +$ or $\times$. Then we have the following well-known results.

**Proposition 2.** Let $\tilde{a}$ and $\tilde{b}$ be two fuzzy numbers. Then $\tilde{a} \oplus \tilde{b}$ and $\tilde{a} \otimes \tilde{b}$ are also fuzzy numbers. Furthermore, we have

$$(\tilde{a} \oplus \tilde{b})_z = [\tilde{a}_z^L + \tilde{b}_z^L, \tilde{a}_z^U + \tilde{b}_z^U]$$

and

$$(\tilde{a} \otimes \tilde{b})_z = \min \{\tilde{a}_z^L \tilde{b}_z^L, \tilde{a}_z^L \tilde{b}_z^U, \tilde{a}_z^U \tilde{b}_z^L, \tilde{a}_z^U \tilde{b}_z^U\}, \max \{\tilde{a}_z^L \tilde{b}_z^L, \tilde{a}_z^U \tilde{b}_z^L, \tilde{a}_z^L \tilde{b}_z^U, \tilde{a}_z^U \tilde{b}_z^U\}.$$
for $i = 1, \ldots, n$, where $\epsilon_i$ are the errors. Let

$$
\mathbf{X} = \begin{bmatrix}
1 & X_{11} & \cdots & X_{1,p-1} \\
1 & X_{21} & \cdots & X_{2,p-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & X_{n1} & \cdots & X_{n,p-1}
\end{bmatrix}
\quad \text{and} \quad
\mathbf{Y} = \begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_n
\end{bmatrix}.
$$

It is well known that the least-squares estimates are

$$
\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y},
$$

where $\hat{\mathbf{b}} = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_{p-1})$, for the following linear model:

$$
Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{p-1} X_{ip-1}
$$

for $i = 1, \ldots, n$. Now let us consider the following two $\alpha$-level linear models:

$$
(\tilde{Y}_i)^L = \beta^L_{0a} + \beta^L_{1a}(\tilde{X}_{i1})^L + \beta^L_{2a}(\tilde{X}_{i2})^L + \cdots + \beta^L_{p-1,a}(\tilde{X}_{ip-1})^L
$$

(2)

and

$$
(\tilde{Y}_i)^U = \beta^U_{0a} + \beta^U_{1a}(\tilde{X}_{i1})^U + \beta^U_{2a}(\tilde{X}_{i2})^U + \cdots + \beta^U_{p-1,a}(\tilde{X}_{ip-1})^U
$$

(3)

for $i = 1, \ldots, n$. Let $\tilde{Y}_i$ and $\tilde{X}_{ij}$ be given fuzzy data. Here the fuzzy data are regarded as the fuzzy numbers. Therefore, we have the corresponding real-valued data $(\tilde{Y}_i)^L$ and $(\tilde{X}_{ij})^L$ for any given $a \in [0, 1]$. Using model (2) and Eq. (1), the least-squares estimates can be obtained using the real-valued data $(\tilde{Y}_i)^L$ and $(\tilde{X}_{ij})^L$. These least-squares estimates are denoted by $\hat{\beta}^L_{ja}$ for $j = 0, 1, 2, \ldots, p - 1$. Similarly, we also have the corresponding least-squares estimates $\hat{\beta}^U_{ja}$ for $j = 0, 1, 2, \ldots, p - 1$ using the real-valued data $(\tilde{Y}_i)^U$ and $(\tilde{X}_{ij})^U$ based on model (3).

We now denote the closed interval $A_{ja}$ for $j = 0, 1, 2, \ldots, p - 1$ by

$$
A_{ja} = [\min\{\hat{\beta}^L_{ja}, \hat{\beta}^U_{ja}\}, \max\{\hat{\beta}^L_{ja}, \hat{\beta}^U_{ja}\}] 
$$

(4)

According to the form of “Resolution Identity” in Proposition 1, the fuzzy estimate of regression parameter $\tilde{\beta}_j$, denoted by $\tilde{\beta}_j$, can be induced by the family of closed interval $\{A_{ja}: 0 \leq a \leq 1\}$, and its membership function is defined by

$$
\bar{\zeta}_{\tilde{\beta}_j}(r) = \sup_{0 \leq a \leq 1} a \mathbb{1}_{A_{ja}}(r).
$$

(5)

The fuzzy estimate given in (5) is not an extended least-squares estimators (ELSE), i.e., an estimator obtained from (1) using extension principle.

Let us also denote $A_{ja}$ by $A_{ja} \equiv [l_j(a), u_j(a)]$, where $l_j(a) = \min\{\hat{\beta}^L_{ja}, \hat{\beta}^U_{ja}\}$ and $u_j(a) = \max\{\hat{\beta}^L_{ja}, \hat{\beta}^U_{ja}\}$. If $l_j(a)$ and $u_j(a)$ are continuous functions of $a$, then it is not hard to see that the $\alpha$-level set $(\tilde{\beta}_j)^L$ of the fuzzy estimate $\tilde{\beta}_j$ is

$$
[(\tilde{\beta}_j)^L] = (\tilde{\beta}_j)^L = \{r: \zeta_{\tilde{\beta}_j}(r) \geq a\} = \left[ \min_{x \leq \gamma \leq 1} l_j(\gamma), \max_{x \leq \gamma \leq 1} u_j(\gamma) \right].
$$

(6)
It says that the fuzzy estimate \( \tilde{\hat{\beta}}_j \) is also a fuzzy number, since the closed interval \( A_{j1} \) is not an empty set and the membership function of \( \tilde{\hat{\beta}}_j \) is upper semicontinuous (the \( \alpha \)-level set \( (\tilde{\hat{\beta}}_j)_{\alpha} \) in (6) is a closed set).

We have another viewpoint to focus on the \( \alpha \)-level interval of \( \tilde{\hat{\beta}}_j \). For any given value \( r \) in the \( \alpha \)-level interval \( (\tilde{\hat{\beta}}_j)_{\alpha} \) of \( \tilde{\hat{\beta}}_j \), we can then say that \( r \) is the estimate of \( \hat{\beta}_j \) with confidence degree \( \alpha \). It is easy to see that if \( r \) is the estimate of \( \hat{\beta}_j \) with confidence degree \( \alpha \), then \( r \) is also an estimate of \( \hat{\beta}_j \) with confidence degree \( \gamma \) for \( \gamma < \alpha \), since \( (\tilde{\hat{\beta}}_j)_{\alpha} \subset (\tilde{\hat{\beta}}_j)_{\gamma} \) for \( \gamma < \alpha \). Suppose that the decision-makers can tolerate confidence degree \( \alpha \). Then they can pick up any value \( r \) from the \( \alpha \)-level interval \( (\tilde{\hat{\beta}}_j)_{\alpha} \) of \( \tilde{\hat{\beta}}_j \) as the estimate of \( \hat{\beta}_j \) for their later use in the further inference.

### 4. Computational procedures

Given an estimate \( r \) of the fuzzy estimate \( \tilde{\hat{\beta}}_j \), we plan to know its membership degree \( \alpha \). If the decision-makers are comfortable with this membership degree \( \alpha \), then it will be reasonable to take the value \( r \) as the estimate of \( \hat{\beta}_j \). In this case, the decision-makers can accept the value \( r \) as the estimate of \( \hat{\beta}_j \) with confidence degree \( \alpha \).

Now from (5), the membership value of any given value \( r \) of \( \tilde{\hat{\beta}}_j \) can be obtained by solving the following optimization problem:

\[
\text{(MP1)} \quad \max \quad \alpha \\
\text{s.t.} \quad \min \left\{ \hat{\beta}_{j2}, \hat{\beta}_{j2} \right\} \leq r \leq \max \left\{ \hat{\beta}_{j2}, \hat{\beta}_{j2} \right\}, \\
0 \leq \alpha \leq 1.
\]

Now we see that the least-squares estimates \( \hat{\beta}_{j2}^L \) and \( \hat{\beta}_{j2}^U \) are obtained from (1). Therefore, \( \hat{\beta}_{j2}^L \) and \( \hat{\beta}_{j2}^U \) are continuous with respect to \( \alpha \), since \( (\tilde{X}_{ij})^L \), \( (\tilde{X}_{ij})^U \), \( (\tilde{Y}_i)^L \) and \( (\tilde{Y}_i)^U \) are all continuous with respect to \( \alpha \). It says that \( l_j(\alpha) \) and \( u_j(\alpha) \) are continuous functions of \( \alpha \). Then from (6), we have that

\[
(\tilde{\hat{\beta}}_j)^L_\alpha = \min \left\{ \hat{\beta}_{j2}, \hat{\beta}_{j2} \right\} = \min \min \left\{ \hat{\beta}_{j2}^L, \hat{\beta}_{j2}^U \right\},
\]

and

\[
(\tilde{\hat{\beta}}_j)^U_\alpha = \max \left\{ \hat{\beta}_{j2}, \hat{\beta}_{j2} \right\} = \max \max \left\{ \hat{\beta}_{j2}^L, \hat{\beta}_{j2}^U \right\}.
\]

From Proposition 1, the membership function of \( \tilde{\beta}_j \) is also given by

\[
\xi_{\tilde{\beta}_j}(r) = \sup_{0 \leq \alpha \leq 1} \alpha \frac{1}{(\tilde{\beta}_j)_\alpha}(r).
\]
Therefore, given any value $r$ of $\tilde{\beta}_j$, its membership degree can also be obtained by solving the following optimization problem:

\begin{align*}
\text{(MP2)} & \quad \max \quad \alpha \\
\text{s.t.} & \quad (\tilde{\beta}_j)_L^L \leq r \leq (\tilde{\beta}_j)_U^U, \\
& \quad 0 \leq \alpha \leq 1,
\end{align*}

where $(\tilde{\beta}_j)_L^L$ and $(\tilde{\beta}_j)_U^U$ are from Eqs. (7) and (8). Two procedures will be proposed to solve problems (MP1) and (MP2), respectively.

**Procedure I.** This procedure will solve problem (MP1). We consider the following two subproblems of (MP1).

\begin{align*}
\text{(MP1-I)} & \quad \max \quad \alpha \\
\text{s.t.} & \quad \hat{\beta}_{jx}^L \leq r \leq \hat{\beta}_{jx}^U, \\
& \quad 0 \leq \alpha \leq 1
\end{align*}

and

\begin{align*}
\text{(MP1-II)} & \quad \max \quad \alpha \\
\text{s.t.} & \quad \hat{\beta}_{jx}^U \leq r \leq \hat{\beta}_{jx}^L, \\
& \quad 0 \leq \alpha \leq 1
\end{align*}

Let $x^*_I$ and $x^*_II$ be the optimums of subproblems (MP1-I) and (MP1-II), respectively. Let $x^* = \max \{x^*_I, x^*_II\}$ and $x^*_0$ be the optimum of original problem (MP1). Then we want to claim $x^* = x^*_0$.

(i) We see that $x^*_I$ and $x^*_II$ are the optimal solutions of subproblems (MP1-I) and (MP1-II), respectively. Since $x^*_I$ is the feasible solution of subproblem (MP1-I), we have

\begin{align*}
\min \{\hat{\beta}_{jx_I}^L, \hat{\beta}_{jx_I}^U\} \leq \tilde{\beta}_{jx_I}^L \leq r \leq \tilde{\beta}_{jx_I}^U \leq \max \{\hat{\beta}_{jx_I}^L, \hat{\beta}_{jx_I}^U\}.
\end{align*}

This says that $x^*_I$ is the feasible solution of problem (MP1), i.e., $x^*_I \leq x^*_0$. Similarly, $x^*_II$ is also the feasible solution of problem (MP1). Thus $x^*_II \leq x^*_0$. We then conclude that $x^* = \max \{x^*_I, x^*_II\} \leq x^*_0$.

(ii) Since $x^*_0$ is the feasible solution of problem (MP1), $x^*_0$ satisfies

\begin{align*}
\min \{\hat{\beta}_{jx_0}^L, \hat{\beta}_{jx_0}^U\} \leq r \leq \max \{\hat{\beta}_{jx_0}^L, \hat{\beta}_{jx_0}^U\}. \tag{9}
\end{align*}

We consider the following cases.

(a) If $\min \{\hat{\beta}_{jx_0}^L, \hat{\beta}_{jx_0}^U\} = \hat{\beta}_{jx_0}^L$ then $\max \{\hat{\beta}_{jx_0}^L, \hat{\beta}_{jx_0}^U\} = \hat{\beta}_{jx_0}^U$. Thus from (9), we see that $x^*_0$ is a feasible solution of subproblem (MP1-I). It says that $x^*_0 \leq x^*_I \leq x^*$.

(b) If $\min \{\hat{\beta}_{jx_0}^L, \hat{\beta}_{jx_0}^U\} = \hat{\beta}_{jx_0}^U$ then $\max \{\hat{\beta}_{jx_0}^L, \hat{\beta}_{jx_0}^U\} = \hat{\beta}_{jx_0}^L$. Thus from (9), we see that $x^*_0$ is a feasible solution of subproblem (MP1-II). It says that $x^*_0 \leq x^*_II \leq x^*$.
From (a) and (b), we have $z_0^* \leq \max \{ z_I^*, z_{II}^* \} = z^*$.

Now from (i) and (ii), we conclude that $z_0^* = z^* = \max \{ z_I^*, z_{II}^* \}$. Next we want to focus on the infeasibility of the subproblems. It is obvious that problem (MP1) is infeasible if all of the two subproblems are infeasible, and problem (MP1) is feasible if one of the subproblems is feasible. If problem (MP1) is infeasible for a fixed $r$, then it means that $r \notin A_{f_i}$ for all $0 \leq x \leq 1$. That is to say, $\tilde{\zeta}_{\hat{\beta}_j}(r) = 0$. If problem (MP1) is feasible, then at least one of the subproblems is feasible and the other one is infeasible. If the subproblem $i$, $i=I$ or II, is infeasible, we assume $x^* = 0$. Then we can also see that $z_0^* = z^* = \max \{ z_I^*, z_{II}^* \}$. Therefore, it is enough to just solve subproblems (MP1-I) and (MP1-II) in order to solve the original problem (MP1).

Procedure II. This procedure will solve problem (MP2). From Eqs. (7) and (8), let

$$\eta(x) \equiv (\hat{\beta}_j)_L^x = \min_{x \leq \gamma \leq 1} \{ \hat{\beta}_j^L, \hat{\beta}_j^U \}$$

and

$$\zeta(x) \equiv (\hat{\beta}_j)_U^x = \max_{x \leq \gamma \leq 1} \{ \hat{\beta}_j^L, \hat{\beta}_j^U \}.$$  

Then $\eta(x)$ is an increasing function of $x$ and $\zeta(x)$ is a decreasing function of $x$. It is not hard to see that $\eta(x)$ and $\zeta(x)$ can be rewritten as

$$\eta(x) = \min \left\{ \min_{x \leq \gamma \leq 1} \hat{\beta}_j^L, \min_{x \leq \gamma \leq 1} \hat{\beta}_j^U \right\} \quad \text{and} \quad \zeta(x) = \max \left\{ \max_{x \leq \gamma \leq 1} \hat{\beta}_j^L, \max_{x \leq \gamma \leq 1} \hat{\beta}_j^U \right\}. \quad (12)$$

The optimization problem (MP2) can now be rewritten as

\[
\begin{align*}
\text{(MP3)} & \quad \max \quad x, \\
\text{subject to} & \quad \eta(x) \leq r, \\
& \quad \zeta(x) \geq r, \\
& \quad 0 \leq x \leq 1.
\end{align*}
\]

Since $\eta(x) \leq \zeta(x)$ as displayed in (10) and (11), we can discard one of the constraints $\eta(x) \leq r$ or $\zeta(x) \geq r$ in the ways described as follows:

(i) If $\eta(1) \leq r \leq \zeta(1)$ then $\tilde{\zeta}_{\hat{\beta}_j}(r) = 1$.

(ii) If $r \leq \eta(1)$ then the constraint $\zeta(x) \geq r$ is redundant since $\zeta(x) \geq \zeta(1) \geq \eta(1) \geq r$ for all $x \in [0, 1]$ using the fact that $\zeta(x)$ is decreasing and $\eta(x) \leq \zeta(x)$ for all $x \in [0, 1]$. 

Thus the following easier optimization problem will be solved

\[(\text{MP4}) \quad \max \ \alpha \]

\[\text{s.t.} \quad \eta(\alpha) \leq r, \quad 0 \leq \alpha \leq 1. \]

(iii) If \( r > \zeta(1) \) then the constraint \( \eta(\alpha) \leq r \) is redundant since \( \eta(\alpha) \leq \eta(1) \leq \zeta(1) \leq r \) for all \( \alpha \in [0, 1] \) using the fact that \( \eta(\alpha) \) is increasing and \( \eta(\alpha) \leq \zeta(\alpha) \) for all \( \alpha \in [0, 1] \). Thus the following easier optimization problem will be solved

\[(\text{MP5}) \quad \max \ \alpha \]

\[\text{s.t.} \quad \zeta(\alpha) \geq r, \quad 0 \leq \alpha \leq 1. \]

Since \( \eta(\alpha) \) is increasing, problem (MP4) can be solved using the following algorithm

\[\text{Step 1: Let } \varepsilon \text{ be the tolerance and } \alpha_0 \text{ be the initial value. Set } \alpha \leftarrow \alpha_0, \ low \leftarrow 0 \text{ and } up \leftarrow 1. \]

\[\text{Step 2: Find } \eta(\alpha) \text{ from (12). If } \eta(\alpha) \leq r \text{ then go to Step 3 otherwise go to Step 4.} \]

\[\text{Step 3: If } r - \eta(\alpha) < \varepsilon \text{ then EXIT and the maximum is } \alpha, \text{ otherwise set } low \leftarrow \alpha, \]

\[\alpha \leftarrow (low + up)/2 \text{ and go to Step 2.} \]

\[\text{Step 4: Set } up \leftarrow \alpha, \ low \leftarrow (low + up)/2 \text{ and go to Step 2.} \]

For problem (MP5), it is enough to consider the equivalent constraint

\[-\zeta(\alpha) \leq -r \]

since \( \zeta(\alpha) \) is decreasing, i.e., \( -\zeta(\alpha) \) is increasing. Thus the above algorithm is still applicable.

Procedure I needs to solve the constrained optimization problems (MP1-I) and (MP1-II). However, Procedure II just needs to solve the unconstrained optimization problems with bounded decision variable in Step 2 for finding \( \eta(\alpha) \). Therefore, Procedure II is preferred to be invoked to obtain the confidence degree of any given \( r \) of \( \hat{\beta}_j \).

Now it is the right time to discuss the predicted fuzzy output data. Given the particular fuzzy data \( \hat{X}_j \) for \( j = 1, 2, \ldots, p - 1 \), the predicted fuzzy output data can be obtained according to the following formula:

\[ \widetilde{Y} = \hat{\beta}_0 \oplus (\hat{\beta}_1 \otimes \hat{X}_1) \oplus (\hat{\beta}_2 \otimes \hat{X}_2) \oplus \cdots \oplus (\hat{\beta}_{p-1} \otimes \hat{X}_{p-1}). \]

The \( \alpha \)-level interval \((\widetilde{Y})_\alpha \) of \( \widetilde{Y} \) can be obtained using Proposition 2. For any given predicted output \( y \) of \( \widetilde{Y} \), its membership degree can be obtained by solving the following optimization problem

\[(\text{MP6}) \quad \max \ \alpha \]

\[\text{s.t.} \quad (\widetilde{Y})_\alpha^L \leq r \leq (\widetilde{Y})_\alpha^U, \quad 0 \leq \alpha \leq 1. \]

The above computational procedure is still applicable for solving problem (MP6).
5. Numerical examples

The membership function of a triangular fuzzy number \( \tilde{a} \) is defined by

\[
\tilde{\xi}_a(r) = \begin{cases} 
(r - a_1)/(a_2 - a_1) & \text{if } a_1 \leq r \leq a_2, \\
(a_3 - r)/(a_3 - a_2) & \text{if } a_2 < r \leq a_3, \\
0 & \text{otherwise,}
\end{cases}
\]

which is denoted by \( \tilde{a} = (a_1,a_2,a_3) \). The triangular fuzzy number \( \tilde{a} \) can be expressed as “around \( a_2 \)” or “being approximately equal to \( a_2 \)” \( a_2 \) is called the core value of \( \tilde{a} \), and \( a_1 \) and \( a_3 \) are called the left and right spread values of \( \tilde{a} \), respectively. The \( \alpha \)-level set (a closed interval) of \( \tilde{a} \) is then \( \tilde{a}_\alpha = [(1 - \alpha)a_1 + \alpha a_2, (1 - \alpha)a_3 + \alpha a_2] \); that is, \( \tilde{a}_\alpha^L = (1 - \alpha)a_1 + \alpha a_2 \) and \( \tilde{a}_\alpha^U = (1 - \alpha)a_3 + \alpha a_2 \).

Now we consider the triangular fuzzy data \( \tilde{Y}_i = (Y_i^L, Y_i, Y_i^U) \) and \( \tilde{X}_{ij} = (X_{ij}^L, X_{ij}, X_{ij}^U) \). The fuzzy data \( \tilde{X}_{ij} \) and \( \tilde{Y}_i \) can be interpreted as “around \( X_{ij} \) and \( Y_i \)”, respectively. Since \( \tilde{Y}_i = (Y_i^L, Y_i, Y_i^U) \) and \( \tilde{X}_{ij} = (X_{ij}^L, X_{ij}, X_{ij}^U) \) are triangular fuzzy data, we see that the core values of \( \tilde{Y}_i \) and \( \tilde{X}_{ij} \) are \( Y_i \) and \( X_{ij} \), respectively. In other words, \( (\tilde{X}_{ij})_1 = X_{ij} = (\tilde{X}_{ij})_1^L \) and \( (\tilde{Y}_i)_1 = Y_i = (\tilde{Y}_i)_1^U \). It also says that the least-squares estimates \( \tilde{\beta}_{j1}^L \) and \( \tilde{\beta}_{j1}^U \) are identical with each other. Therefore, the closed interval \( A_{j1} \) in (4) is degenerated as a singleton set \( A_{j1} = \{\tilde{\beta}_{j1}^L\} = \{\tilde{\beta}_{j1}^U\} \). In this case, the least-squares estimates obtained by using the core values \( \tilde{Y}_i \) and \( \tilde{X}_{ij} \), denoted by \( \tilde{\beta}_{j} \), satisfies \( \tilde{\beta}_{j} = \tilde{\beta}_{j1}^L = \tilde{\beta}_{j1}^U \). It also says that the estimate \( \tilde{\beta}_{j} \) has membership degree 1 based on the fuzzy estimate in (5). We conclude that the least-squares estimates obtained by using the crisp representatives out of the fuzzy data have membership degree 1. However, it is possible for some real-valued data taken from the support of the corresponding fuzzy data with positive membership degrees such that the corresponding least-squares estimates are not in the support of the corresponding fuzzy estimates. In this case, we may regard the corresponding least-squares estimates as having membership degrees 0 with respect to the corresponding fuzzy estimates. Now we give an example to clarify the theoretical discussions so far.

Example 3. Suppose that we have the following triangular fuzzy data \( \tilde{Y}_i = (Y_i^L, Y_i, Y_i^U) \) and \( \tilde{X}_{ij} = (X_{ij}^L, X_{ij}, X_{ij}^U) \) which are described in Table 1.

Since \( \tilde{Y}_i = (Y_i^L, Y_i, Y_i^U) \) and \( \tilde{X}_{ij} = (X_{ij}^L, X_{ij}, X_{ij}^U) \) are triangular fuzzy data, we see that the core values of \( \tilde{Y}_i \) and \( \tilde{X}_{ij} \) are \( Y_i \) and \( X_{ij} \), respectively. Then the least-squares estimates using the real-valued data \( Y_i \) and \( X_{ij} \) are

\[
\tilde{\beta}_0 = 3.452612790, \quad \tilde{\beta}_1 = 0.4960049761 \quad \text{and} \quad \tilde{\beta}_2 = 0.009199081. \quad (13)
\]

Invoking the above computational procedure, we let \( \varepsilon = 10^{-6} \) and the initial value \( x_0 = 0.9 \). Then the membership degree of any given value \( r \) taken from the different fuzzy estimates \( \tilde{\beta}_0, \tilde{\beta}_1 \) and \( \tilde{\beta}_2 \) are obtained and described in Table 2.

We see that if \( r \) is taken to close \( \tilde{\beta}_j \) in (13) for \( j=0,1,2 \), then its membership degree will close to one. This is reasonable in intuitive viewpoint. For the fuzzy estimate \( \tilde{\beta}_0 \),...
Table 1
The triangular fuzzy input and output data

<table>
<thead>
<tr>
<th>i</th>
<th>( y_i )</th>
<th>( x_{i1} )</th>
<th>( x_{i2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(111, 162, 194)</td>
<td>(151, 274, 322)</td>
<td>(1432, 2450, 3461)</td>
</tr>
<tr>
<td>2</td>
<td>(88, 120, 161)</td>
<td>(101, 180, 291)</td>
<td>(2448, 3254, 4463)</td>
</tr>
<tr>
<td>3</td>
<td>(161, 223, 288)</td>
<td>(221, 375, 539)</td>
<td>(2592, 3802, 5116)</td>
</tr>
<tr>
<td>4</td>
<td>(83, 131, 194)</td>
<td>(128, 205, 313)</td>
<td>(1414, 2838, 3252)</td>
</tr>
<tr>
<td>5</td>
<td>(51, 67, 83)</td>
<td>(62, 86, 112)</td>
<td>(1024, 2347, 3766)</td>
</tr>
<tr>
<td>6</td>
<td>(124, 169, 213)</td>
<td>(132, 265, 362)</td>
<td>(2163, 3782, 5091)</td>
</tr>
<tr>
<td>7</td>
<td>(62, 81, 102)</td>
<td>(66, 98, 152)</td>
<td>(1687, 3008, 4325)</td>
</tr>
<tr>
<td>8</td>
<td>(138, 192, 241)</td>
<td>(151, 330, 463)</td>
<td>(1524, 2450, 3864)</td>
</tr>
<tr>
<td>9</td>
<td>(82, 116, 159)</td>
<td>(115, 195, 291)</td>
<td>(1216, 2137, 3161)</td>
</tr>
<tr>
<td>10</td>
<td>(41, 55, 71)</td>
<td>(35, 53, 71)</td>
<td>(1432, 2560, 3782)</td>
</tr>
<tr>
<td>11</td>
<td>(168, 252, 367)</td>
<td>(307, 430, 584)</td>
<td>(2592, 4020, 5562)</td>
</tr>
<tr>
<td>12</td>
<td>(178, 232, 346)</td>
<td>(284, 372, 498)</td>
<td>(2792, 4427, 6163)</td>
</tr>
<tr>
<td>13</td>
<td>(111, 144, 198)</td>
<td>(121, 236, 370)</td>
<td>(1734, 2660, 4094)</td>
</tr>
<tr>
<td>14</td>
<td>(78, 103, 148)</td>
<td>(103, 157, 211)</td>
<td>(1426, 2088, 3312)</td>
</tr>
<tr>
<td>15</td>
<td>(167, 212, 267)</td>
<td>(216, 370, 516)</td>
<td>(1785, 2605, 4042)</td>
</tr>
</tbody>
</table>

Table 2
Membership degrees of \( \tilde{\beta}_j \)

<table>
<thead>
<tr>
<th>r</th>
<th>Membership ( \tilde{\beta}_0 )</th>
<th>r</th>
<th>Membership ( \tilde{\beta}_1 )</th>
<th>r</th>
<th>Membership ( \tilde{\beta}_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4</td>
<td>0.9750</td>
<td>0.491</td>
<td>0.7997</td>
<td>0.0087</td>
<td>0.9531</td>
</tr>
<tr>
<td>2.6</td>
<td>0.9797</td>
<td>0.492</td>
<td>0.8369</td>
<td>0.0088</td>
<td>0.9624</td>
</tr>
<tr>
<td>2.8</td>
<td>0.9845</td>
<td>0.493</td>
<td>0.8751</td>
<td>0.0089</td>
<td>0.9718</td>
</tr>
<tr>
<td>3.0</td>
<td>0.9892</td>
<td>0.494</td>
<td>0.9145</td>
<td>0.0090</td>
<td>0.9812</td>
</tr>
<tr>
<td>3.2</td>
<td>0.9940</td>
<td>0.495</td>
<td>0.9558</td>
<td>0.0091</td>
<td>0.9906</td>
</tr>
<tr>
<td>3.4</td>
<td>0.9987</td>
<td>0.496</td>
<td>0.9998</td>
<td>0.0092</td>
<td>0.9999</td>
</tr>
<tr>
<td>3.6</td>
<td>0.9952</td>
<td>0.497</td>
<td>0.9860</td>
<td>0.0093</td>
<td>0.9922</td>
</tr>
<tr>
<td>3.8</td>
<td>0.9888</td>
<td>0.498</td>
<td>0.9721</td>
<td>0.0094</td>
<td>0.9845</td>
</tr>
<tr>
<td>4.0</td>
<td>0.9823</td>
<td>0.499</td>
<td>0.9582</td>
<td>0.0095</td>
<td>0.9768</td>
</tr>
<tr>
<td>4.2</td>
<td>0.9758</td>
<td>0.500</td>
<td>0.9443</td>
<td>0.0096</td>
<td>0.9690</td>
</tr>
<tr>
<td>4.4</td>
<td>0.9693</td>
<td>0.501</td>
<td>0.9304</td>
<td>0.0097</td>
<td>0.9613</td>
</tr>
</tbody>
</table>

the estimate \( r=3.6 \) has membership degree 0.9952. If the decision-maker is comfortable with this membership degree 0.9952, then he/she can take the value \( r = 3.6 \) as the estimate of parameter \( \beta_0 \) under the circumstances of fuzzy (imprecise) input and output data.

We also have the \( \alpha \)-level closed intervals \( (\tilde{\beta}_j)_\alpha \) described in Table 3 for \( j=0,1,2 \).

We see that the 1.00-level closed intervals are equal to the least-squares estimates in (13). The 0.98-level closed intervals of \( \tilde{\beta}_0, \tilde{\beta}_1 \) and \( \tilde{\beta}_2 \) are \([2.6120, 4.0702], [0.4956, 0.4974] \) and \([0.0090, 0.0095] \), respectively. This means that the decision-makers can pick up any values \( r \) from those intervals as the estimates of \( \beta_0, \beta_1 \) and \( \beta_2 \) with confidence degree 0.98 if they can tolerate this confidence degree.
Table 3
The $\tilde{x}$-level closed intervals $(\tilde{\beta}_j)_x$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$(\tilde{\beta}_0)_x$</th>
<th>$(\tilde{\beta}_1)_x$</th>
<th>$(\tilde{\beta}_2)_x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>[1.3398, 4.9941]</td>
<td>[0.4949, 0.4996]</td>
<td>[0.0087, 0.0098]</td>
</tr>
<tr>
<td>0.96</td>
<td>[1.7654, 4.6865]</td>
<td>[0.4951, 0.4989]</td>
<td>[0.0088, 0.0097]</td>
</tr>
<tr>
<td>0.97</td>
<td>[2.1894, 4.3785]</td>
<td>[0.4953, 0.4982]</td>
<td>[0.0089, 0.0096]</td>
</tr>
<tr>
<td>0.98</td>
<td>[2.6120, 4.0702]</td>
<td>[0.4956, 0.4974]</td>
<td>[0.0090, 0.0095]</td>
</tr>
<tr>
<td>0.99</td>
<td>[3.0331, 3.7615]</td>
<td>[0.4958, 0.4967]</td>
<td>[0.0091, 0.0093]</td>
</tr>
<tr>
<td>1.00</td>
<td>[3.4526, 3.4526]</td>
<td>[0.4960, 0.4960]</td>
<td>[0.0092, 0.0092]</td>
</tr>
</tbody>
</table>

Table 4
Membership degrees $\xi_{\tilde{x}}(y)$

<table>
<thead>
<tr>
<th>$y$</th>
<th>139</th>
<th>149</th>
<th>159</th>
<th>169</th>
<th>179</th>
<th>180</th>
<th>190</th>
<th>200</th>
<th>210</th>
<th>220</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi_{\tilde{x}}(y)$</td>
<td>0.7418</td>
<td>0.8062</td>
<td>0.8701</td>
<td>0.9338</td>
<td>0.9975</td>
<td>0.9966</td>
<td>0.9428</td>
<td>0.8898</td>
<td>0.8375</td>
<td>0.7858</td>
</tr>
</tbody>
</table>

Table 5
The $\tilde{x}$-level closed intervals of $\tilde{y}$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$(\tilde{Y})_x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>[163.6926, 198.0592]</td>
</tr>
<tr>
<td>0.91</td>
<td>[165.2630, 196.1663]</td>
</tr>
<tr>
<td>0.92</td>
<td>[166.8335, 194.2787]</td>
</tr>
<tr>
<td>0.93</td>
<td>[168.4040, 192.3965]</td>
</tr>
<tr>
<td>0.94</td>
<td>[169.9744, 190.5199]</td>
</tr>
<tr>
<td>0.95</td>
<td>[171.5444, 188.6489]</td>
</tr>
<tr>
<td>0.96</td>
<td>[173.1139, 186.7838]</td>
</tr>
<tr>
<td>0.97</td>
<td>[174.6828, 184.9246]</td>
</tr>
<tr>
<td>0.98</td>
<td>[176.2508, 183.0715]</td>
</tr>
<tr>
<td>0.99</td>
<td>[177.8179, 181.2246]</td>
</tr>
<tr>
<td>1.00</td>
<td>[179.3839, 179.3839]</td>
</tr>
</tbody>
</table>

Example 4. Using the data in Example 3, the predicted fuzzy output data will be obtained from the following formula:

$$\tilde{Y} = \tilde{\beta}_0 \oplus (\tilde{\beta}_1 \otimes \tilde{X}_1) \oplus (\tilde{\beta}_2 \otimes \tilde{X}_2).$$

Suppose that the triangular fuzzy data $\tilde{X}_1 = (161, 310, 483)$ and $\tilde{X}_2 = (1484, 2410, 3894)$ are given. The core values of $\tilde{X}_1$ and $\tilde{X}_2$ are 310 and 2410, respectively. If we use these two real-valued data 310 and 2410 and the least-squares estimates in (13), the predicted output is 179.3839. Table 4 gives the membership degree $\xi_{\tilde{x}}(y)$ for different predicted output $y$ using the computational procedure proposed above.

We see that $y = 180$ has membership degree 0.9966. If the decision-maker is comfortable with this membership degree 0.9966, then he/she can take the value $y = 180$ as the predicted output under the circumstances of fuzzy (imprecise) input.
data $\tilde{X}_1 = (161, 310, 483)$ and $\tilde{X}_2 = (1484, 2410, 3894)$. The $\alpha$-level closed intervals of $\tilde{Y}$ are obtained and displayed in Table 5.

The 0.98-level closed interval of $\tilde{Y}$ is $[176.2508, 183.0715]$. This means that the decision-makers can pick up any values $y$ from this interval as the predicted output with confidence degree 0.98 if they can tolerate this confidence degree.

6. Conclusions

We have obtained the fuzzy estimates of regression parameters with the help of “Resolution Identity”. That is to say, the fuzzy estimates are constructed from the $\alpha$-level least-squares estimates using the $\alpha$-level real-valued data of the corresponding fuzzy input and output data. In order to obtain the membership degree of any given value taken from the fuzzy estimates of regression parameters, we have to solve the optimization problems. We also propose two computational procedures to solve the optimization problems. Finally, a numerical example is provided for clarifying the theoretical discussions.

There are some weak points that were pointed out by one of the referees. Let us consider the simplest regression problem $Y = \beta X$. Given a single pair of triangular fuzzy data $(\tilde{X}, \tilde{Y})$ with $\tilde{X} = (5 - \delta, 5, 5 + \delta) = \tilde{Y}$. Then the fuzzy estimate $\tilde{\beta}$ is degenerated as a real number 1. This situation will rarely occur. The reason is as follows. From Eq. (4), the fuzzy estimates are degenerated as the crisp numbers (real numbers) if $\hat{\beta}_{jx} = \hat{\beta}_{jx}$ for all $x \in [0, 1]$. We denote the matrices by $X_L = [(\tilde{X}_{ij})_L]$, $X_U = [(\tilde{X}_{ij})_U]$, $Y_L = [(\tilde{Y}_i)_L]$ and $Y_U = [(\tilde{Y}_i)_U]$. Then from (1), the estimate $\hat{\beta}_L$ is the $j$th element of the vector $\hat{\beta}_L = (X_L^t X_L)^{-1}X_L^t Y_L$ and $\hat{\beta}_U$ is the $j$th element of the vector $\hat{\beta}_U = (X_U^t X_U)^{-1}X_U^t Y_U$. Therefore, $\hat{\beta}_L = \hat{\beta}_U$ for all $x \in [0, 1]$ will rarely occur. However if it, unfortunately, happened, then we still can take this crisp number as a fuzzy estimate since a real number is just a degenerate case of a fuzzy number. Or, alternatively, we can change some of the fuzzy data a little bit without changing their core values to avoid this situation. For example, we may change their spread values a little bit.

On the other hand, if we consider the same problem $Y = \beta X$ then the estimate $\tilde{\beta}$ is really a fuzzy number when the single pair triangular fuzzy data $(\tilde{X}, \tilde{Y})$ is taken from $\tilde{X} = (6 - s, 6, 6 + s)$ and $\tilde{Y} = (3 - s, 3, 3 + s)$. This example elicits a situation that the different fuzzy data sets (with the same fuzziness) for the same problem may obtain the extremely different fuzzy estimates, the crisp case and the fuzzy case, respectively, as the above examples described. Hence the fuzziness of the fuzzy estimates depends not only on the fuzziness of the fuzzy data but also on the core values (the position) of the data. This strange situation can be explained as follows. The reason is that the fuzzy estimate is a fuzzy number. When we talk about number, we should regard the number as a whole item (a united item). Therefore, it would be natural to treat the fuzzy number (i.e. the fuzzy estimate) as a whole item rather than considering its
core value and fuzziness separately. Under this concern, the different fuzzy data sets will give the different fuzzy estimates (fuzzy numbers) including the crisp case since the crisp number is just a special case of a fuzzy number.

Acknowledgements

The author would like to thank the anonymous referees for their valuable comments and suggestions.

References