Linear Regression Analysis for Fuzzy Input and Output Data Using the Extension Principle

HSIEN-CHUNG WU
Department of Information Management
National Chi Nan University, Puli, Nantou 545, Taiwan, R.O.C.
hcwu@im.ncnu.edu.tw

(Received and accepted April 2002)

Abstract—The method for obtaining the fuzzy least squares estimators with the help of the extension principle in fuzzy sets theory is proposed. The membership functions of fuzzy least squares estimators will be constructed according to the usual least squares estimators. In order to obtain the membership value of any given value taken from the fuzzy least squares estimator, optimization problems have to be solved. We also provide the methodology for evaluating the predicted fuzzy output from the given fuzzy input data. © 2003 Elsevier Science Ltd. All rights reserved.

Keywords—Fuzzy numbers, Fuzzy least squares estimator, Optimization.

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 a similar reason. Therefore, the fuzzy sets theory is naturally 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 meters. The phrase “around 30 meters” can be regarded as a fuzzy number 30. This is the main concern of this paper.

Since Zadeh [1] 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. [2] initiated this research topic. They also generalized their approaches to the more general models in [3–5].

In the approach of Tanaka et al. [2], they considered the L-R fuzzy data and minimized the index of fuzziness of the fuzzy linear regression model. Redden and Woodall [6] compared various fuzzy regression models and gave the difference between the approaches of fuzzy regression analysis and usual regression analysis. They also pointed out some weaknesses of the approaches proposed by Tanaka et al. Chang and Lee [7] also pointed out another weakness of the approaches proposed by Tanaka et al. Bárđossy [8] proposed many different measures of fuzziness which must be minimized with respect to some suggested constraints. Peters [9] introduced a
new fuzzy linear regression model based on Tanaka's approach by considering the fuzzy linear programming problem. Diamond [10] introduced a metric on the set of fuzzy numbers by invoking the Hausdorff metric on the compact α-level sets, and used this metric to define a least squares criterion function as in the usual sense, which must be minimized. Ma et al. [11] generalized Diamond's approach by embedding the set of fuzzy numbers into a Banach space isometrically and isomorphically. Näther and Albrecht [12] and Körner and Näther [13] introduced the concept of random fuzzy sets (fuzzy random variables) into the linear regression model, and developed an estimation theory for the parameters. The other interesting references are also given in [14-26].

In this paper, we will construct the fuzzy least squares estimators using the extension principle in fuzzy sets theory which was introduced by Zadeh [27-29]. The membership functions of fuzzy least squares estimators will be constructed according to the usual least squares estimators with the help of the extension principle. In order to obtain the membership value of any given value taken from the fuzzy least squares estimator, optimization problems have to be solved. We also provide the methodology for evaluating the predicted fuzzy output from the given fuzzy input data.

In Section 2, we give some properties of fuzzy numbers. In Section 3, we give some useful results from the extension principle. In Section 4, we conduct the membership functions of fuzzy least squares estimators according to the usual least squares estimators with the help of the extension principle. In Section 5, we will develop the computational procedures to obtain the membership value of any given value taken from the fuzzy least squares estimators. We also provide an example to clarify the theoretical results, and show the possible applications in linear regression analysis for fuzzy data. In Section 6, the methodology for transacting the predicted fuzzy output from the given fuzzy input data is proposed.

2. FUZZY NUMBERS

Let X be a universal set. Then a fuzzy subset A of X is defined by its membership function \( \xi_A : X \to [0, 1] \). We denote by \( \tilde{A}_\alpha = \{ x : \xi_A(x) \geq \alpha \} \) the α-level set of \( \tilde{A} \), where \( \tilde{A}_0 \) is the closure of the set \( \{ x : \xi_A(x) \neq 0 \} \). \( \tilde{A} \) is called a normal fuzzy set if there exists an \( x \) such that \( \xi_A(x) = 1 \). \( \tilde{A} \) is called a convex fuzzy set if \( \xi_A(\lambda x + (1 - \lambda)y) \geq \min\{\xi_A(x), \xi_A(y)\} \) for \( \lambda \in [0, 1] \). (That is, \( \xi_A \) 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 \( \{ E : f(x) > 0 \} \) is a closed set for each \( \alpha \). Or equivalently, \( f \) is upper semicontinuous at \( y \) if and only if \( \forall \epsilon > 0 \), \( \exists \delta > 0 \) such that \( |x - y| < \delta \) implies \( f(x) < f(y) + \epsilon \).

\( \tilde{a} \) 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 \( \xi_{\tilde{a}} \) is upper semicontinuous.

(iii) The α-level set \( \tilde{a}_\alpha \) is bounded for each \( \alpha \in [0, 1] \).

From Zadeh [1], \( \tilde{A} \) is a convex fuzzy set if and only if its α-level set \( \tilde{A}_\alpha = \{ x : \xi_A(x) \geq \alpha \} \) is a convex set for all \( \alpha \). Therefore, if \( \tilde{a} \) is a fuzzy number, then the α-level set \( \tilde{a}_\alpha \) is a compact (closed and bounded in \( \mathbb{R} \)) and convex set; that is, \( \tilde{a} \) is a closed interval. The α-level set of \( \tilde{a} \) is then denoted by \( \tilde{a}_\alpha = [\tilde{a}_\alpha^L, \tilde{a}_\alpha^U] \). We also see that \( \tilde{a}_\alpha^L \) and \( \tilde{a}_\alpha^U \) are continuous with respect to \( \alpha \), since its membership function is upper semicontinuous. The following proposition is useful for further discussions.

**Proposition 2.1. Resolution Identity.** (See [27-29].) Let \( \tilde{A} \) be a fuzzy set with membership function \( \xi_{\tilde{A}} \) and the α-level set \( \tilde{A}_\alpha = \{ x : \xi_{\tilde{A}}(x) \geq \alpha \} \) be given. Then

\[
\xi_{\tilde{A}}(x) = \sup_{\alpha \in [0, 1]} \alpha \cdot 1_{\tilde{A}_\alpha}(x).
\]
\( \tilde{a} \) is called a crisp number with value \( m \) if its membership function is
\[
\xi_{\tilde{a}}(x) = \begin{cases} 
1, & \text{if } x = m, \\
0, & \text{otherwise}.
\end{cases}
\]

It is denoted by \( \tilde{a} \equiv \tilde{1}_m \). It is easy to see that \( \tilde{1}_m \) is a crisp number if \( \xi_{\tilde{a}}(x) = \begin{cases} 
1, & \text{if } x = m, \\
0, & \text{otherwise}.
\end{cases} \) for all \( \alpha \in [0, 1] \).

\( \tilde{a} \) is called a nonnegative fuzzy number if \( \xi_{\tilde{a}}(x) = 0 \) for all \( x < 0 \) and a nonpositive fuzzy number if \( \xi_{\tilde{a}}(x) = 0 \) for all \( x > 0 \). It is obvious that \( \tilde{a}_L^L \) and \( \tilde{a}_U^U \) are nonnegative real numbers for all \( \alpha \in [0, 1] \) if \( \tilde{a} \) is a nonnegative fuzzy number, and \( \tilde{a}_L^L \) and \( \tilde{a}_U^U \) are nonpositive real numbers for all \( \alpha \in [0, 1] \) 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
\[
\xi_{\tilde{a} \odot \tilde{b}}(z) = \sup_{xoy=z} \min \{ \xi_{\tilde{a}}(x), \xi_{\tilde{b}}(y) \}
\]

using the extension principle in [27-29], where the operations \( \odot = \oplus \) or \( \odot = \otimes \) correspond to the operations \( \circ = + \) or \( \times \). Then we have the following well-known results.

**Proposition 2.2.** 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
\[
(\tilde{a} \oplus \tilde{b})_\alpha = [\tilde{a}_L^L + \tilde{b}_L^L, \tilde{a}_U^U + \tilde{b}_U^U]
\]
and
\[
(\tilde{a} \otimes \tilde{b})_\alpha = \left[ \min \left\{ a_{\alpha}^L b_{\alpha}^L, a_{\alpha}^U b_{\alpha}^U, a_{\alpha}^L b_{\alpha}^U, a_{\alpha}^U b_{\alpha}^U \right\}, \max \left\{ a_{\alpha}^L b_{\alpha}^L, a_{\alpha}^U b_{\alpha}^U, a_{\alpha}^L b_{\alpha}^U, a_{\alpha}^U b_{\alpha}^U \right\} \right].
\]

### 3. Extension Principle

We denote by \( \mathcal{F} \) the set of all fuzzy subsets of \( \mathbb{R} \). Let \( f(x_1, x_2, \ldots, x_n) \) be a nonfuzzy function from \( \mathbb{R}^n \) into \( \mathbb{R} \) and \( \tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n \) be \( n \) fuzzy subsets of \( \mathbb{R} \). By the extension principle in [27-29], we can induce a fuzzy-valued function \( \tilde{f} : \mathcal{F}^n \to \mathcal{F} \) from the nonfuzzy function \( f(x_1, x_2, \ldots, x_n) \). That is to say, \( \tilde{f}(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) \) is a fuzzy subset of \( \mathbb{R} \). The membership function of \( \tilde{f}(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) \) is defined by
\[
\mu_{\tilde{f}(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n)}(r) = \sup \left\{ \mu_{\tilde{A}_i}(x_i) : (x_1, \ldots, x_n) : f(x_1, \ldots, x_n) = r \right\}.
\]

The following propositions are very useful for further discussions.

**Proposition 3.1.** (See [30].) Let \( S \) be a compact set in \( \mathbb{R}^n \). If \( f \) is upper semicontinuous on \( S \), then \( f \) attains maximum over \( S \), and if \( f \) is lower semicontinuous on \( S \), then \( f \) attains minimum over \( S \).

**Proposition 3.2.** Let \( f(x_1, \ldots, x_n) \) be a real-valued function and \( \tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n \) be \( n \) fuzzy subsets of \( \mathbb{R} \). Let \( \tilde{f} : \mathcal{F}^n \to \mathcal{F} \) be a fuzzy-valued function induced by \( f(x_1, \ldots, x_n) \) via the extension principle. Suppose that each membership function \( \mu_{\tilde{A}_i} \) is upper semicontinuous for all \( i = 1, \ldots, n \) and \( \{(x_1, \ldots, x_n) : f(x_1, \ldots, x_n) \} \) is a compact set (it will be a closed and bounded set in \( \mathbb{R}^n \) for all \( r \)). Then the \( \alpha \)-level set of \( \tilde{f}(\tilde{A}_1, \ldots, \tilde{A}_n) \) is
\[
(\tilde{f}(\tilde{A}_1, \ldots, \tilde{A}_n))_\alpha = \left\{ f(x_1, \ldots, x_n) : x_1 \in (\tilde{A}_1)_\alpha, \ldots, x_n \in (\tilde{A}_n)_\alpha \right\}.
\]

**Proof.** If
\[
r \in \left\{ f(x_1, \ldots, x_n) : x_1 \in (\tilde{A}_1)_\alpha, \ldots, x_n \in (\tilde{A}_n)_\alpha \right\},
\]

then there exist \( (x_1, \ldots, x_n) \) such that \( f(x_1, \ldots, x_n) = r \) and \( x_i \in (\tilde{A}_i)_\alpha \) for all \( i = 1, \ldots, n \). Therefore, \( \mu_{(\tilde{A}_1, \ldots, \tilde{A}_n)}(r) \) is well-defined.
then there exists an \((x_1, \ldots, x_n)\) such that \(r = f(x_1, \ldots, x_n)\) and \(x_i \in \tilde{A}_i\) for all \(i = 1, \ldots, n\); that is, \(\mu_{\tilde{A}_i}(x_i) \geq \alpha\) for all \(i = 1, \ldots, n\). Thus, \(\min_{1 \leq i \leq n} \{\mu_{\tilde{A}_i}(x_i)\} \geq \alpha\). It says that

\[
\mu_{\tilde{f}(\tilde{A}_1, \ldots, \tilde{A}_n)}(r) = \sup_{\{x_1, \ldots, x_n, r = f(x_1, \ldots, x_n)\}} \min_{1 \leq i \leq n} \{\mu_{\tilde{A}_i}(x_i)\} \geq \alpha.
\]

Therefore,

\[
\left\{ f(x_1, \ldots, x_n) : x_1 \in \left(\tilde{A}_1\right)_\alpha, \ldots, x_n \in \left(\tilde{A}_n\right)_\alpha \right\} \subseteq \left(\tilde{f}(\tilde{A}_1, \ldots, \tilde{A}_n)\right)_\alpha.
\]

On the other hand, if \(r \in (\tilde{f}(\tilde{A}_1, \ldots, \tilde{A}_n))_\alpha\), then

\[
\sup_{\{x_1, \ldots, x_n, r = f(x_1, \ldots, x_n)\}} \min_{1 \leq i \leq n} \{\mu_{\tilde{A}_i}(x_i)\} \geq \alpha;
\]

that is, there exists an \((x_1, \ldots, x_n)\) such that \(\min_{1 \leq i \leq n} \{\mu_{\tilde{A}_i}(x_i)\} \geq \alpha\) and \(r = f(x_1, \ldots, x_n)\) by using Proposition 3.1, since \(\{x_1, \ldots, x_n : r = f(x_1, \ldots, x_n)\}\) is a compact set and \(\min_{1 \leq i \leq n} \{\mu_{\tilde{A}_i}(x_i)\}\) is upper semicontinuous on \(\mathbb{R}^n\). (From [31, p. 381], the infimum of any collection of upper semicontinuous function is upper semicontinuous.) Therefore, \(\mu_{\tilde{A}_i}(x_i) \geq \alpha\) for all \(i = 1, \ldots, n\); that is, \(x_i \in \tilde{A}_i\) for all \(i = 1, \ldots, n\) and

\[
r \in \left\{ f(x_1, \ldots, x_n) : x_1 \in (\tilde{A}_1)\alpha, \ldots, x_n \in (\tilde{A}_n)\alpha \right\}.
\]

This completes the proof. \[\Box\]

PROPOSITION 3.3. Let \(f(x_1, \ldots, x_n)\) be a continuous real-valued function and \(\tilde{a}_1, \ldots, \tilde{a}_n\) be \(n\) fuzzy numbers. Let \(\tilde{f} : \mathcal{F}^n \rightarrow \mathcal{F}\) be a fuzzy-valued function induced by \(f(x_1, \ldots, x_n)\) via the extension principle. Suppose that \(\{(x_1, \ldots, x_n) : r = f(x_1, \ldots, x_n)\}\) is a compact set for all \(r\). Then \(\tilde{f}(\tilde{a}_1, \ldots, \tilde{a}_n)\) is a fuzzy number and its \(\alpha\)-level set is

\[
\left(\tilde{f}(\tilde{a}_1, \ldots, \tilde{a}_n)\right)_\alpha = \left\{ f(x_1, \ldots, x_n) : x_1 \in (\tilde{a}_1)\alpha, \ldots, x_n \in (\tilde{a}_n)\alpha \right\}.
\]

PROOF. By Proposition 3.2, [32, p. 82, Theorem 4.25], and [32, p. 87, Theorem 4.37],

\[
\{\mu_{\tilde{f}(\tilde{a}_1, \ldots, \tilde{a}_n)}(r) \geq \alpha\} = \left(\tilde{f}(\tilde{a}_1, \ldots, \tilde{a}_n)\right)_\alpha = \left\{ f(x_1, \ldots, x_n) : x_1 \in (\tilde{a}_1)\alpha, \ldots, x_n \in (\tilde{a}_n)\alpha \right\}
\]

is a closed interval, since \((\tilde{a}_i)\alpha\) is a closed interval for all \(i = 1, \ldots, n\). This also says that the membership function \(\mu_{\tilde{f}(\tilde{a}_1, \ldots, \tilde{a}_n)}(r)\) is upper semicontinuous. This completes the proof. \[\Box\]

4. FUZZY LEAST SQUARES ESTIMATORS

The linear regression model is displayed as follows:

\[
Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{p-1} X_{ip-1} + \varepsilon_i,
\]

for \(i = 1, \ldots, n\), where \(\varepsilon_i\) are independent normal random variables with expectation \(E(\varepsilon_i) = 0\) and variance \(V(\varepsilon_i) = \sigma^2\). Let

\[
X = \begin{bmatrix}
1 & X_{11} & \cdots & X_{1,p-1} \\
1 & X_{21} & \cdots & X_{2,p-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & X_{n1} & \cdots & X_{n,p-1}
\end{bmatrix}
\quad \text{and} \quad
Y = \begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_n
\end{bmatrix}.
\]
It is well known that the least squares estimators are

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

(1)

where \(\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_{p-1})\). Alternatively, the least squares estimators \(\hat{\beta}_j\), for \(j = 0, 1, 2, \ldots, p - 1\), can be regarded as functions of \(X\) and \(Y\). They are denoted by \(\hat{\beta}_j = f_j(X, Y)\) for \(j = 0, 1, 2, \ldots, p - 1\).

Under the consideration of fuzzy data \(\tilde{X}\) and \(\tilde{Y}\), where the components of \(\tilde{X}\) and \(\tilde{Y}\) are \(\tilde{X}_{ij}\) and \(\tilde{Y}_i\), respectively, for \(i = 1, \ldots, n\) and \(j = 1, \ldots, p - 1\) (here the fuzzy data are regarded as the fuzzy numbers), we can obtain the fuzzy least squares estimators \(\tilde{\beta}_j\), for \(j = 0, 1, 2, \ldots, p - 1\), according to equation (1) and the extension principle. The membership function of \(\tilde{\beta}_j\) is then given by

$$\xi_{\tilde{\beta}_j}(r) = \left\{ \begin{array}{l}
\sup_{(x, y): r = f_j(x, y)} \min_{i=1, \ldots, n; j=1, \ldots, p-1} \left\{ \xi_{\tilde{X}_{ij}}(x_{ij}), \xi_{\tilde{Y}_i}(y_i) \right\}.
\end{array} \right.$$  

(2)

From Proposition 3.3, the fuzzy least squares estimators \(\tilde{\beta}_j\) are all fuzzy numbers for \(j = 0, 1, 2, \ldots, p - 1\). From Proposition 2.1, the membership function of \(\tilde{\beta}_j\) can also be written as

$$\xi_{\tilde{\beta}_j}(r) = \sup_{0 \leq \alpha \leq 1} \alpha \cdot \mathbb{1}_{(\tilde{\beta}_j)_\alpha}(r),$$

(3)

where \((\tilde{\beta}_j)_\alpha\) is the \(\alpha\)-level set of \(\tilde{\beta}_j\). From Proposition 3.3 and equation (2), \(\tilde{\beta}_j\) is a fuzzy number and its \(\alpha\)-level set \((\tilde{\beta}_j)_\alpha\) is then given by

$$(\tilde{\beta}_j)_\alpha = \left\{ f_j(x, y) : x_{ij} \in (\tilde{X}_{ij})_\alpha, y_i \in (\tilde{Y}_i)_\alpha, \text{ for } i = 1, \ldots, n; j = 1, \ldots, p - 1 \right\}.$$  

(4)

Since \(\tilde{\beta}_j\) is a fuzzy number, its \(\alpha\)-level set \((\tilde{\beta}_j)_\alpha\) is a closed interval

$$(\tilde{\beta}_j)_\alpha = \left[ (\tilde{\beta}_j)^L_\alpha, (\tilde{\beta}_j)^U_\alpha \right].$$

From equation (4), the left-endpoint \((\tilde{\beta}_j)^L_\alpha\) and right-endpoint \((\tilde{\beta}_j)^U_\alpha\) can be displayed as

$$\left(\tilde{\beta}_j\right)^L_\alpha = \left\{ (x, y) : (\tilde{X}_{ij})_\alpha \leq x_{ij} \leq (\tilde{X}_{ij})_\alpha, (\tilde{V}_i)_\alpha \leq y_i \leq (\tilde{V}_i)_\alpha, \text{ for } i = 1, \ldots, n; j = 1, \ldots, p - 1 \right\} f_j(x, y)$$

(5)

and

$$\left(\tilde{\beta}_j\right)^U_\alpha = \left\{ (x, y) : (\tilde{X}_{ij})_\alpha \leq x_{ij} \leq (\tilde{X}_{ij})_\alpha, (\tilde{V}_i)_\alpha \leq y_i \leq (\tilde{V}_i)_\alpha, \text{ for } i = 1, \ldots, n; j = 1, \ldots, p - 1 \right\} f_j(x, y).$$

(6)

Recall that \(f_j(x, y)\) is the \(j\)th element of the vector \((X'X)^{-1}X'Y\) (ref. equation (1)).

We have another viewpoint to focus on the \(\alpha\)-level interval of \(\tilde{\beta}_j\). For any given value \(r\) in the \(\alpha\)-level interval \((\tilde{\beta}_j)_\alpha\) of \(\tilde{\beta}_j\), we can then say that \(r\) is the estimate of \(\beta_j\) with confidence degree \(\alpha\). It is easy to see that if \(r\) is the estimate of \(\beta_j\) with confidence degree \(\alpha\), then \(r\) is also an estimate of \(\beta_j\) with confidence degree \(\gamma\) for \(\gamma < \alpha\), since \((\tilde{\beta}_j)_\alpha \subset (\tilde{\beta}_j)_\gamma\) 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 \((\tilde{\beta}_j)_\alpha\) of \(\tilde{\beta}_j\) as the estimate of \(\beta_j\) for their later use in the statistical inference.
5. COMPUTATIONAL METHODS AND EXAMPLE

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

Now from equation (3), given a least squares estimate \( r \) of \( \hat{\beta}_j \), its membership value can be obtained by solving the following optimization problem:

\[
\max \quad \alpha, \\
\text{subject to} \quad (\hat{\beta}_j)_\alpha^L \leq t \leq (\hat{\beta}_j)_\alpha^U, \\
\quad 0 \leq \alpha \leq 1.
\]  

Let \( \eta_j(\alpha) = (\hat{\beta}_j)_\alpha^L \) and \( \zeta_j(\alpha) = (\hat{\beta}_j)_\alpha^U \). The optimization problem (MP1) can now be rewritten as

\[
\max \quad \alpha, \\
\text{subject to} \quad \eta_j(\alpha) \leq r, \\
\quad \zeta_j(\alpha) \geq r, \\
\quad 0 \leq \alpha \leq 1.
\]  

Since \( \eta_j(\alpha) \leq \zeta_j(\alpha) \) for all \( \alpha \in [0, 1] \), \( \eta_j(\alpha) \) is an increasing function, and \( \zeta_j(\alpha) \) is a decreasing function, we can discard one of the constraints \( \eta_j(\alpha) \leq r \) or \( \zeta_j(\alpha) \geq r \) in the ways described as follows.

(i) If \( \eta_j(1) \leq r \leq \zeta_j(1) \), then \( \mu_{\hat{\beta}_j}(r) = 1 \).

(ii) If \( r < \eta_j(1) \), then the constraint \( \zeta_j(\alpha) \geq r \) is redundant since \( \zeta_j(\alpha) \geq \zeta_j(1) \geq \eta_j(1) \geq r \) for all \( \alpha \in [0, 1] \) using the fact that \( \zeta_j(\alpha) \) is decreasing and \( \eta_j(\alpha) \leq \zeta_j(\alpha) \) for all \( \alpha \in [0, 1] \).

Thus, the following easier optimization problem will be solved:

\[
\max \quad \alpha, \\
\text{subject to} \quad \eta_j(\alpha) \leq r, \\
\quad 0 \leq \alpha \leq 1.
\]  

(iii) If \( r > \zeta_j(1) \), then the constraint \( \eta_j(\alpha) \leq r \) is redundant since \( \eta_j(\alpha) \leq \eta_j(1) \leq \zeta_j(1) \leq r \) for all \( \alpha \in [0, 1] \) using the fact that \( \eta_j(\alpha) \) is increasing and \( \eta_j(\alpha) \leq \zeta_j(\alpha) \) for all \( \alpha \in [0, 1] \).

Thus, the following easier optimization problem will be solved:

\[
\max \quad \alpha, \\
\text{subject to} \quad \zeta_j(\alpha) \geq r, \\
\quad 0 \leq \alpha \leq 1.
\]  

Since \( \eta_j(\alpha) \) is increasing, problem (MP3) can be solved using the following algorithm (bisection search).

**STEP 1.** Let \( \epsilon \) be the tolerance and \( \alpha_0 \) be the initial value. Set \( \alpha \leftarrow \alpha_0 \), \( \text{low} \leftarrow 0 \), and \( \text{up} \leftarrow 1 \).

**STEP 2.** Find \( \eta_j(\alpha) \). If \( \eta_j(\alpha) \leq r \) then go to Step 3; otherwise go to Step 4.

**STEP 3.** If \( r - \eta_j(\alpha) < \epsilon \) then EXIT and the maximum is \( \alpha \); otherwise set \( \text{low} \leftarrow \alpha \), \( \alpha \leftarrow (\text{low} + \text{up})/2 \) and go to Step 2.

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

For problem (MP4), it is enough to consider the equivalent constraint

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

since \( \zeta_j(\alpha) \) is decreasing, i.e., \(-\zeta_j(\alpha) \) is increasing. Thus, the above algorithm is still applicable.
In Step 2, we will encounter a difficult problem for finding \( \eta_j(\alpha) \) or \( \zeta_j(\alpha) \) as in equations (5) and (6). Next, we are going to provide a more convenient computational procedure. Let \( \tilde{X} \) and \( \tilde{Y} \) be the given fuzzy data. We adopt the following notations:

\[
I_j(\alpha) = \min \left\{ f_1\left( \left( \tilde{X} \right)^L, \left( \tilde{Y} \right)^L \right) , f_1\left( \left( \tilde{X} \right)^U, \left( \tilde{Y} \right)^U \right) \right\}
\]

and

\[
u_j(\alpha) = \max \left\{ f_1\left( \left( \tilde{X} \right)^L, \left( \tilde{Y} \right)^L \right) , f_1\left( \left( \tilde{X} \right)^U, \left( \tilde{Y} \right)^U \right) \right\},
\]

where the vectors \( \left( \tilde{X} \right)^L, \left( \tilde{Y} \right)^L \), and \( \left( \tilde{Y} \right)^U \) have the components \( \left( \tilde{X} \right)^{ij}, \left( \tilde{Y} \right)^{ij} \), and \( \left( \tilde{Y} \right)^{ij} \), respectively, for \( i = 1, \ldots, n \) and \( j = 1, \ldots, p - 1 \). Then we see that

\[
\left( \tilde{\beta}_j \right)^L_\alpha \leq I_j(\alpha) \quad \text{and} \quad \left( \tilde{\beta}_j \right)^U_\alpha \geq \nu_j(\alpha),
\]

or equivalently, the closed interval \( A_{j\alpha} = [I_j(\alpha), \nu_j(\alpha)] \) satisfies

\[
A_{j\alpha} \subseteq \left( \tilde{\beta}_j \right)^L_\alpha = \left( \tilde{\beta}_j \right)^U_\alpha.
\]

Now, given a least squares estimate \( r \) of \( \tilde{\beta}_j \), the above bisection search algorithm will be invoked by setting \( \eta_j(\alpha) = I_j(\alpha) \) and \( \zeta_j(\alpha) = \nu_j(\alpha) \) if there exists an \( \alpha_0 \) such that the family of closed intervals

\[
\{ A_{j\alpha} = [I_j(\alpha), \nu_j(\alpha)] : \alpha_0 \leq \alpha \leq 1 \}
\]

is decreasing with respect to \( \alpha \) (i.e., \( A_{j\alpha} \subseteq A_{j\beta} \) for \( \alpha_0 \leq \beta < \alpha \leq 1 \)). If the final value \( \alpha^* \) is obtained from the above algorithm, then, from equation (7), the membership value of \( r \) will be greater than \( \alpha^* \), since \( \left( \tilde{\beta}_j \right)^L_\alpha \) is increasing and \( \left( \tilde{\beta}_j \right)^U_\alpha \) is decreasing with respect to \( \alpha \). In this case, if the decision-makers are comfortable with this value \( \alpha^* \), then they can take this value \( r \) as the least squares estimate of \( \tilde{\beta}_j \) for their later use in the statistical inference, since its confidence degree is greater than \( \alpha^* \).

Sometimes, it is really hard to show that the family of closed intervals in (9) is decreasing with respect to \( \alpha \) for some \( \alpha_0 \). However, (9) can be checked numerically by evaluating many different \( \alpha \) values. Although this technique is not so rigorous, it is really helpful in solving \( \eta_j(\alpha) \) and \( \zeta_j(\alpha) \). In many cases, (9) is satisfied for \( \alpha_0 \geq 0.9 \). In this case, it will be reasonable and comfortable, since we are really interested in the higher \( \alpha \)-level cases.

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

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

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,

\[
\tilde{a}^L_\alpha = (1 - \alpha)a_1 + \alpha a_2 \quad \text{and} \quad \tilde{a}^U_\alpha = (1 - \alpha)a_3 + \alpha a_2.
\]
Table 1.

<table>
<thead>
<tr>
<th>i</th>
<th>( \hat{Y}_i )</th>
<th>( \hat{X}_{i1} )</th>
<th>( \hat{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>(1444, 2838, 3202)</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, 266, 362)</td>
<td>(2163, 3782, 6091)</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, 539)</td>
<td>(2592, 3802, 5116)</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>

Example 5.1. Suppose that we have the following triangular fuzzy data \( \hat{Y}_i = (Y_i^L, Y_i, Y_i^U) \) and \( \hat{X}_{ij} = (X_{ij}^L, X_{ij}, X_{ij}^U) \), listed in Table 1.

The fuzzy data \( \hat{X}_{ij} \) and \( \hat{Y}_i \) can be interpreted as "around \( X_{ij} \) and \( Y_i \)", respectively. Since \( f_i = (Y_i^L, Y_i, Y_i^U) \) and \( Z_{ij} = (X_{ij}^L, X_{ij}, X_{ij}^U) \) are triangular fuzzy data, we see that the core values of \( \hat{Y}_i \) and \( \hat{X}_{ij} \) are \( Y_i \) and \( X_{ij} \), respectively. Then the least squares estimates using the real-valued data \( Y_i \) and \( X_{ij} \) are

\[
\hat{\beta}_0 = 3.452612790, \quad \hat{\beta}_1 = 0.4960049761, \quad \text{and} \quad \hat{\beta}_2 = 0.009199081. \tag{10}
\]

We then have the \( A_{j\alpha} \) described in Table 2 for \( j = 0, 1, 2 \).

Table 2.

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( A_{0\alpha} )</th>
<th>( A_{1\alpha} )</th>
<th>( A_{2\alpha} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>[38.5296, 24.0362]</td>
<td>[0.4716, 0.5099]</td>
<td>[0.0056, 0.0192]</td>
</tr>
<tr>
<td>0.15</td>
<td>[34.2640, 23.4170]</td>
<td>[0.4739, 0.5237]</td>
<td>[0.0041, 0.0185]</td>
</tr>
<tr>
<td>0.25</td>
<td>[29.9040, 22.3372]</td>
<td>[0.4762, 0.5309]</td>
<td>[0.0034, 0.0176]</td>
</tr>
<tr>
<td>0.35</td>
<td>[25.4659, 20.7820]</td>
<td>[0.4787, 0.5328]</td>
<td>[0.0033, 0.0167]</td>
</tr>
<tr>
<td>0.45</td>
<td>[20.9708, 18.7952]</td>
<td>[0.4813, 0.5310]</td>
<td>[0.0038, 0.0158]</td>
</tr>
<tr>
<td>0.55</td>
<td>[16.4439, 16.4499]</td>
<td>[0.4840, 0.5266]</td>
<td>[0.0045, 0.0147]</td>
</tr>
<tr>
<td>0.65</td>
<td>[11.9149, 13.8269]</td>
<td>[0.4868, 0.5208]</td>
<td>[0.0054, 0.0136]</td>
</tr>
<tr>
<td>0.75</td>
<td>[7.4172, 11.0021]</td>
<td>[0.4896, 0.5140]</td>
<td>[0.0065, 0.0124]</td>
</tr>
<tr>
<td>0.85</td>
<td>[2.9866, 8.0402]</td>
<td>[0.4923, 0.5068]</td>
<td>[0.0076, 0.0111]</td>
</tr>
<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>[1.7554, 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>

Invoking the above computational procedure, we let \( \epsilon = 10^{-6} \) and the initial value \( \alpha_0 = 0.9 \). Then the membership value of any given value \( r \) taken from the different fuzzy least squares estimators \( \hat{\beta}_0, \hat{\beta}_1, \) and \( \hat{\beta}_2 \) are obtained and described in Table 3.
We see that if \( r \) is taken to close \( \hat{\beta}_j \) in (10) for \( j = 0, 1, 2 \), then its membership value will close to one. This is reasonable in intuitive viewpoint. For the fuzzy least squares estimator \( \hat{\beta}_0 \), the estimate \( r = 3.6 \) has membership value greater than 0.9952. If the decision-maker is comfortable with this membership value 0.9952, then he (she) can take the value \( r = 3.6 \) as the least squares estimate of parameter \( \beta_0 \) under the circumstances of fuzzy (imprecise) input and output data. On the other hand, we see that \( A_{j\alpha} \subseteq (\hat{\beta}_j)_\alpha \) for \( j = 1, \ldots, p-1 \) from (8), and the closed intervals \( A_{0\alpha}, A_{1\alpha}, \) and \( A_{2\alpha} \) for \( \alpha = 0.98 \) are \([2.6120, 4.0702], [0.4956, 0.4974], \) and \([0.0090, 0.0095], \) respectively. Therefore, the decision-makers can pick up any values \( r \) from those intervals as the estimates of \( \beta_0, \beta_1, \) and \( \beta_2, \) respectively, with confidence degree at least 0.98 if they are comfortable with this confidence degree.

### 6. PREDICTED FUZZY OUTPUT DATA

Given the particular fuzzy data \( \tilde{X}_j \) for \( j = 1, 2, \ldots, p-1 \), the predicted fuzzy output data can be obtained according to the following formula:

\[
\tilde{Y} = \tilde{\beta}_0 \ominus (\tilde{\beta}_1 \ominus \tilde{X}_1) \ominus (\tilde{\beta}_2 \ominus \tilde{X}_2) \ominus \cdots \ominus (\tilde{\beta}_{p-1} \ominus \tilde{X}_{p-1}).
\]

The \( \alpha \)-level interval \((\tilde{Y})_\alpha\) of \( \tilde{Y} \) can be obtained using Proposition 2.2.

Let \( \ominus \) be any binary operation \( \ominus \) or \( \ominus \) between two closed intervals \([a, b]\) and \([c, d]\). Then \([a, b] \ominus \ominus [c, d]\) is defined by

\[
[a, b] \ominus \ominus [c, d] = \{ z \in \mathbb{R} \mid z = x \circ y, \forall x \in [a, b], \forall y \in [c, d], \text{where } \circ \text{ is an usual binary operation } + \text{ or } \times \}.
\]

Then it is not hard to see that

\[
[a, b] \ominus \ominus [c, d] = [a + c, b + d] \text{ and } [a, b] \ominus \ominus [c, d] = [\min\{ac, ad, bc, bd\}, \max\{ac, ad, bc, bd\}].
\]

Now we denote the closed interval \( B_\alpha = [B^L_\alpha, B^U_\alpha] \) as

\[
B_\alpha = A_{0\alpha} \ominus \ominus \left( A_{1\alpha} \ominus \ominus \left( A_{2\alpha} \ominus \ominus \ldots \ominus \ominus A_{p-1,\alpha} \ominus \ominus \left( \tilde{X}_{p-1} \circ \cdots \circ \tilde{X}_2 \circ \cdots \circ (\tilde{\beta}_{p-1} \ominus \tilde{X}_{p-1})_\alpha \right) \cdots \circ \tilde{X}_2 \circ \cdots \circ \tilde{\beta}_{p-1} \right) \right) \circ \cdots \circ \tilde{\beta}_2 \circ \cdots \circ \tilde{\beta}_1 \circ \tilde{\beta}_0 \right).
\]

for \( \alpha \in [0, 1] \). Since \( A_{j\alpha} \subseteq (\tilde{\beta}_j)_\alpha \) for \( j = 1, \ldots, p-1 \) from (8), we see that

\[
B_\alpha \subseteq (\tilde{Y})_\alpha.
\]
using Proposition 2.2. We also see that the family of closed intervals \( \{B_\alpha : \alpha_0 \leq \alpha \leq 1\} \) is decreasing with respect to \( \alpha \) if there exists an \( \alpha_0 \) such that the families of closed intervals \( \{A_{j\alpha} : \alpha_0 \leq \alpha \leq 1\} \) are decreasing for all \( j = 1, \ldots, p - 1 \).

For any given predicted output \( y \) of \( \tilde{Y} \), the above bisection search algorithm can also be invoked by setting \( \eta_j(\alpha) = B^L_\alpha \) and \( \zeta_j(\alpha) = B^U_\alpha \) if there exists an \( \alpha_0 \) such that the families of closed intervals \( \{A_{j\alpha} : \alpha_0 \leq \alpha \leq 1\} \) are decreasing for all \( j = 1, \ldots, p - 1 \). If the final value \( \alpha^* \) is obtained from the above algorithm, then, from equation (11), the membership value of \( y \) will be greater than \( \alpha^* \), since \( (\tilde{Y})^L_\alpha \) is increasing and \( (\tilde{Y})^U_\alpha \) is decreasing with respect to \( \alpha \). In this case, if the decision-makers are comfortable with this value \( \alpha^* \), then they can take this value \( y \) as the predicted output for their later use in the statistical inference, since its confidence degree is greater than \( \alpha^* \).

Using the data in Example 5.1, the predicted fuzzy output data will be obtained from the following formula:

\[
\tilde{Y} = \tilde{\beta}_0 \odot (\tilde{\beta}_1 \otimes \tilde{X}_1) \odot (\tilde{\beta}_2 \otimes \tilde{X}_2).
\]

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 (10), the predicted output is 179.3839. Table 4 gives the membership value \( \zeta_\tilde{Y}(y) \) for different predicted output \( y \) using the computational procedure proposed above.

<table>
<thead>
<tr>
<th>( y )</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>
</tr>
</thead>
<tbody>
<tr>
<td>( \zeta_\tilde{Y}(y) )</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>
</tr>
</tbody>
</table>

We see that \( y = 180 \) has membership value greater than 0.9966. If the decision-maker is comfortable with this membership value 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 closed intervals \( B_\alpha \) are obtained and displayed in Table 5.

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( B_\alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>[163.6926, 198.0592]</td>
</tr>
<tr>
<td>0.91</td>
<td>[165.2630, 196.1060]</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>

The 0.98-level closed interval is \( [176.2508, 183.0715] \). Since \( B_\alpha \subseteq (\tilde{Y})_\alpha \) in (11), this means that the decision-makers can pick up any values \( y \) from this interval as the predicted output with confidence degree at least 0.98 if they can tolerate this confidence degree.
REFERENCES

3. H. Tanaka and J. Watada, Possibilistic linear system and their application to the linear regression model,
5. H. Tanaka and H. Ishibuchi, Identification of possibilistic linear systems by quadratic membership functions
6. D.C. Redden and W.H. Woodall, Properties of certain fuzzy linear regression methods, *Fuzzy Sets and
(1997).
12. W. Næther and M. Albrecht, Linear regression with random fuzzy observations, *Statistics* 21, 521–531,
13. R. Körner and W. Næther, Linear regression with random fuzzy variables: Extended classical estimates, best
16. P.-T. Chang and E.S. Lee, A generalized fuzzy weighted least-squares regression, *Fuzzy Sets and Systems* 82,
17. J.F. Dunyak and D. Wunsch, Fuzzy regression by fuzzy number neural networks, *Fuzzy Sets and Systems*
21. B. Kim and R.R. Bishu, Evaluation of fuzzy linear regression models by comparing membership functions,
22. H. Moskowitz and K. Kim, On assessing the H value in fuzzy linear regression, *Fuzzy Sets and Systems* 58,
27. L.A. Zadeh, The concept of linguistic variable and its application to approximate reasoning I, *Information