

Article

On Optimal and Asymptotic Properties of a Fuzzy L_2 Estimator

Jin Hee Yoon ^{1,*}  and Przemyslaw Grzegorzewski ^{2,3}¹ Department of Mathematics and Statistics, Sejong University, Seoul 05006, Korea² Faculty of Mathematics and Information Science, Warsaw University of Technology, Koszykowa 75, 00-662 Wasaw, Poland³ Systems Research Institute, Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland; pgrzeg@ibspan.waw.pl

* Correspondence: jin9135@sejong.ac.kr; Tel.: +82-10-9118-3135

Received: 31 August 2020; Accepted: 24 October 2020; Published: 4 November 2020



Abstract: A fuzzy least squares estimator in the multiple with fuzzy-input–fuzzy-output linear regression model is considered. The paper provides a formula for the L_2 estimator of the fuzzy regression model. This paper proposes several operations for fuzzy numbers and fuzzy matrices with fuzzy components and discussed some algebraic properties that are needed to use for proving theorems. Using the proposed operations, the formula for the variance, provided and this paper, proves that the estimators have several important optimal properties and asymptotic properties: they are Best Linear Unbiased Estimator (BLUE), asymptotic normality and strong consistency. The confidence regions of the coefficient parameters and the asymptotic relative efficiency (ARE) are also discussed. In addition, several examples are provided including a Monte Carlo simulation study showing the validity of the proposed theorems.

Keywords: fuzzy random variable; fuzzy least squares estimator; triangular fuzzy matrix; fuzzy-type linear estimator; unbiased estimator; BLUE; consistency; asymptotic normality

1. Introduction

Regression analysis is commonly perceived as one of the most useful tools in statistical modeling. If the data could be observed precisely, the classical regression appears usually as a sufficient solution. However, we encounter a lot of situations where the observations cannot be obtained precisely. In such a case we need a framework to handle the uncertain matters coming from the following two sources: randomness and imprecision. As far as randomness could be satisfactorily managed by probability theory, one has to adopt a suitable approach for modeling imprecise data. Applying fuzzy sets proposed by Zadeh [1], Tanaka et al. [2] introduced a fuzzy regression analysis. On the other hand, Diamond [3] generalized the main technique of the regression analysis, i.e., the least squared method, for fuzzy numbers. Further on numerous researchers considered different fuzzy regression models, known as crisp input-fuzzy output (CIFO) or fuzzy input-fuzzy output (FIFO), both with crisp or fuzzy parameters (see, e.g., [3–24]). Many fewer efforts were dedicated to exploring statistical properties of the estimators in fuzzy regression models. For situations where data has an assumed error structure Diamond [25] and Näther [26–29] discussed the fuzzy best linear unbiased estimators (FBLUES). Kim et al. [16] proved the asymptotic properties of fuzzy least squares estimators (FLSEs) for a fuzzy simple linear regression model. Due to the complexity of mathematical formulas describing fuzzy least squares estimators, some authors use α -cuts of fuzzy numbers to express the estimates (e.g., [21,30]), while others separate the estimators into some parts, e.g., corresponding to the mode and two spreads in the case of triangular fuzzy numbers (see, e.g., [6,26,31,32]). Moreover, many authors do not express

any analytic formulas for the desired estimators but they determine the estimates from the normal equations directly (see [3,19]).

To overcome these problems the Yoon et al. [33–35] redefined the mathematical model of a fuzzy linear regression using the so-called triangular fuzzy matrix and suitable operations defined both on triangular fuzzy numbers like on triangular fuzzy matrices. Even though this paper deals with the triangular fuzzy numbers, it can be extended to the general case [36]. Moreover, the importance of triangular and trapezoidal fuzzy numbers has been emphasized in [37]. This approach enables us to determine fuzzy least squares estimators of the regression parameters in a concise form which is also useful for exploring the statistical properties of the estimators. The asymptotic theory for fuzzy multiple regression model is hardly discussed in the paper so far. In this contribution we continue the examination of the fuzzy least squares estimator obtained there, focusing on its fundamental finite-sample and asymptotic properties.

The paper is organized as follows: In Section 2, we introduce basic notation related to triangular fuzzy numbers, triangular fuzzy matrices and operation defined on them which would be used later in the contribution. Some algebraic properties that are needed to prove theorems are discussed in Section 2. In Section 3, we discuss the fuzzy linear regression model based on the author’s previous studies [33,35]. Next, in Section 4 we prove that the fuzzy least squares estimator shown in the previous section is Best Linear Unbiased Estimator (BLUE). Then we present the next important results of the paper showing that under some general assumptions the aforementioned estimator is asymptotically normal (in Section 5) and strongly consistent (Section 6). In addition, the confidence regions of the coefficient parameters and the asymptotic relative efficiency (ARE) are also discussed in Section 6. Several examples are proposed including Monte Carlo simulation study. Section 7 concludes the results.

2. Preliminaries

Each triangular fuzzy number A can be represented by an ordered triple, i.e., $A = (l_a, a, r_a)$, where a is the mode of A , while l_a and r_a denote the lower and the upper bound of the support of A , respectively. Further on let us assume that \mathcal{F}_T denote a family of all triangular fuzzy numbers defined on the real numbers \mathbb{R} .

Besides well-known basic operations on fuzzy numbers, like addition

$$X \oplus Y = (l_x + l_y, x + y, r_x + r_y), \tag{1}$$

where $X = (l_x, x, r_x), Y = (l_y, y, r_y) \in \mathcal{F}_T$, or scalar multiplication

$$kX = \begin{cases} (kl_x, kx, kr_x) & \text{if } k \geq 0, \\ (kr_x, kx, kl_x) & \text{if } k < 0. \end{cases} \tag{2}$$

Some other operations defined in \mathcal{F}_T are sometimes useful. Let us recall here some concepts proposed in [33]:

$$X \odot Y = l_x l_y + xy + r_x r_y, \tag{3}$$

$$X \otimes Y = (l_x l_y, xy, r_x r_y). \tag{4}$$

Clearly, $X \odot Y, X \otimes Y \in \mathcal{F}_T$, but the output of (3) is a crisp number (i.e., it is isomorphic with the corresponding real value).

Further on we’ll also need some operations defined on the matrices of fuzzy numbers. We denote by $M_{\mathbb{R}^*}$ a set of all $n \times n$ real crisp matrices with nonnegative elements and let $M_{\mathcal{F}_T}$ be a set of all triangular fuzzy matrices matrix (t.f.m.), i.e., matrices whose elements belong to \mathcal{F}_T . For any two

$n \times n$ triangular fuzzy matrices $\tilde{\Gamma} = [X_{ij}]$ and $\tilde{\Lambda} = [Y_{ij}]$, a crisp matrix $A = [a_{ij}]$ and a constant $k > 0$ we have

$$\tilde{\Gamma} \oplus \tilde{\Lambda} = [X_{ij} \oplus Y_{ij}], \tag{5}$$

$$\tilde{\Gamma} \odot \tilde{\Lambda} = \left[\sum_{k=1}^n X_{ik} \odot Y_{kj} \right], \tag{6}$$

$$\tilde{\Gamma} \otimes \tilde{\Lambda} = \left[\bigoplus_{k=1}^n X_{ik} \otimes Y_{kj} \right], \tag{7}$$

$$\tilde{A}\tilde{\Gamma} = \left[\bigoplus_{k=1}^n a_{ik} X_{kj} \right], \tag{8}$$

$$k\tilde{\Gamma} = [kX_{ij}]. \tag{9}$$

Of course, $\tilde{\Gamma} \oplus \tilde{\Lambda}, \tilde{\Gamma} \otimes \tilde{\Lambda}, \tilde{A}\tilde{\Gamma} \in M_{\mathcal{F}_T}$ and $\tilde{\Gamma} \odot \tilde{\Lambda} \in M_{\mathbb{R}^*}$.

We can also define the following three types of fuzzy scalar multiplications of a crisp matrix:

$$X\tilde{A} = [a_{ij}X], \tag{10}$$

$$X \odot \tilde{\Gamma} = [X \odot X_{ij}], \tag{11}$$

$$X \otimes \tilde{\Gamma} = [X \otimes X_{ij}], \tag{12}$$

where $X \in \mathcal{F}_T, \tilde{A} = [a_{ij}] \in M_{\mathbb{R}^*}$ and $\tilde{\Gamma} = [X_{ij}] \in M_{\mathcal{F}_T}$.

The proofs of the following properties of the foregoing operations in \mathcal{F}_T and in \mathcal{F}_T are straightforward.

Proposition 1. For any $X, Y, Z \in \mathcal{F}_T$ and $k \in \mathbb{R}^*$ the following properties hold:

1. $X \odot Y = Y \odot X$ and $X \otimes Y = Y \otimes X$
2. $(X \otimes Y) \otimes Z = X \otimes (Y \otimes Z)$
3. $(X \otimes Y) \odot Z = (X \otimes Z) \odot Y = X \odot (Y \otimes Z)$
4. $(kX) \odot Y = k(X \odot Y) = X \odot (kY)$
5. $(kX) \otimes Y = k(X \otimes Y) = X \otimes (kY)$.

Proposition 2. Let $\tilde{\Gamma}, \tilde{\Lambda}, \tilde{\Omega} \in M_{\mathcal{F}_T}$. Then we have

1. $\tilde{\Gamma} \odot \tilde{\Lambda} \neq \tilde{\Lambda} \odot \tilde{\Gamma}$ and $\tilde{\Gamma} \otimes \tilde{\Lambda} \neq \tilde{\Lambda} \otimes \tilde{\Gamma}$
2. $(\tilde{\Gamma} \otimes \tilde{\Lambda}) \otimes \tilde{\Omega} = \tilde{\Gamma} \otimes (\tilde{\Lambda} \otimes \tilde{\Omega})$
3. $(\tilde{\Gamma} \otimes \tilde{\Lambda}) \odot \tilde{\Omega} = \tilde{\Gamma} \odot (\tilde{\Lambda} \otimes \tilde{\Omega})$
4. $(\tilde{\Gamma} \oplus \tilde{\Lambda}) \odot \tilde{\Omega} = (\tilde{\Gamma} \odot \tilde{\Omega}) \oplus (\tilde{\Lambda} \odot \tilde{\Omega})$
5. $(\tilde{\Gamma} \oplus \tilde{\Lambda}) \otimes \tilde{\Omega} = (\tilde{\Gamma} \otimes \tilde{\Omega}) \oplus (\tilde{\Lambda} \otimes \tilde{\Omega})$
6. $(\tilde{\Gamma} \odot \tilde{\Lambda})^T = \tilde{\Lambda}^T \odot \tilde{\Gamma}^T$ and $(\tilde{\Gamma} \otimes \tilde{\Lambda})^T = \tilde{\Lambda}^T \otimes \tilde{\Gamma}^T$.

Proposition 3. For $\tilde{A}, \tilde{B} \in M_{\mathbb{R}^*}$ and $\tilde{\Gamma}, \tilde{\Lambda} \in M_{\mathcal{F}_T}$ the following properties hold:

1. $\tilde{A}\tilde{\Gamma} \neq \tilde{\Gamma}\tilde{A}$
2. $(\tilde{A}\tilde{\Gamma}) \odot \tilde{\Lambda} = \tilde{A}(\tilde{\Gamma} \odot \tilde{\Lambda})$ and $(\tilde{A}\tilde{\Gamma}) \otimes \tilde{\Lambda} = \tilde{A}(\tilde{\Gamma} \otimes \tilde{\Lambda})$
3. $\tilde{\Gamma} \odot (\tilde{\Lambda}\tilde{A}) = (\tilde{\Gamma} \odot \tilde{\Lambda})\tilde{A}$ and $\tilde{\Gamma} \otimes (\tilde{\Lambda}\tilde{A}) = (\tilde{\Gamma} \otimes \tilde{\Lambda})\tilde{A}$
4. $(\tilde{A}\tilde{\Gamma})^T = \tilde{\Gamma}^T \tilde{A}^T$ and $(\tilde{\Gamma}\tilde{A})^T = \tilde{A}^T \tilde{\Gamma}^T$
5. $(\tilde{A} + \tilde{B})\tilde{\Gamma} = \tilde{A}\tilde{\Gamma} \oplus \tilde{B}\tilde{\Gamma}$
6. $(\tilde{A}\tilde{B})\tilde{\Gamma} = \tilde{A}(\tilde{B}\tilde{\Gamma})$.

Proposition 4. For $\tilde{A} \in M_{\mathcal{R}}$, $X \in \mathcal{F}_{\mathcal{T}}$ and $\tilde{\Gamma}, \tilde{\Lambda} \in M_{\mathcal{F}_{\mathcal{T}}}$ we have

1. $X \odot \tilde{\Gamma} = \tilde{\Gamma} \odot X$ and $X \otimes \tilde{\Gamma} = \tilde{\Gamma} \otimes X$
2. $(X \otimes \tilde{\Gamma}) \otimes \tilde{\Lambda} = X \otimes (\tilde{\Gamma} \otimes \tilde{\Lambda}) = \tilde{\Gamma} \otimes (X \otimes \tilde{\Lambda})$
3. $(X \otimes \tilde{\Gamma}) \odot \tilde{\Lambda} = X \odot (\tilde{\Gamma} \otimes \tilde{\Lambda}) = \tilde{\Gamma} \odot (X \otimes \tilde{\Lambda})$
4. $(\tilde{A}\tilde{\Gamma}) \odot X = \tilde{A}(\tilde{\Gamma} \odot X) = \tilde{\Gamma} \odot (\tilde{A}X) = (\tilde{A}X) \odot \tilde{\Gamma}$
5. $(\tilde{A}\tilde{\Gamma}) \otimes X = \tilde{A}(\tilde{\Gamma} \otimes X) = \tilde{\Gamma} \otimes (\tilde{A}X) = (\tilde{A}X) \otimes \tilde{\Gamma}$
6. $\tilde{\Gamma} \odot (X\tilde{A}) = (\tilde{\Gamma} \odot X)\tilde{A} = (\tilde{\Gamma}\tilde{A}) \odot X = X \odot (\tilde{\Gamma}\tilde{A})$
7. $\tilde{\Gamma} \otimes (X\tilde{A}) = (\tilde{\Gamma} \otimes X)\tilde{A} = (\tilde{\Gamma}\tilde{A}) \otimes X = X \otimes (\tilde{\Gamma}\tilde{A})$.

3. Fuzzy Least Squares Estimation

Consider a classical simple linear model

$$Y = X\beta + \epsilon \tag{13}$$

where $Y = (Y_1, \dots, Y_n)^T$ is a vector of observed responses, X is a design matrix with size $n \times (p + 1)$ of explanatory variables x_{ij} , $\beta = (\beta_0, \beta_1, \dots, \beta_p)^T$ denotes a p -dimensional vector of unknown parameters, and $\epsilon = (\epsilon_1, \dots, \epsilon_n)^T$ is a vector of errors. We usually assume that $E\epsilon = \mathbf{0}$ and $Var\epsilon = \sigma^2 I_n$, where $\sigma^2 < \infty$. It is also usual to take $x_{i0} \equiv 1, i = 1, \dots, n$.

The most common estimator in the simple regression model is the least squares estimator (LSE) given by

$$\hat{\beta}_n = (X^T X)^{-1} X^T Y, \tag{14}$$

where the design matrix X is supposed to have the full rank. By the Gauss–Markov theorem (14) is the best linear unbiased estimator (BLUE) of the parameters, where “best” means giving the lowest variance of the estimate, as compared to other unbiased, linear estimators. On the other hand, estimator (14) is strong consistent under certain conditions for the design matrix, i.e., $(X_n^T X_n)^{-1} \rightarrow 0$.

In this paper, we consider a fuzzy generalization of (13), i.e., the following linear regression model with fuzzy inputs and fuzzy outputs

$$Y_i = \beta_0 \oplus \beta_1 X_{i1} \oplus \dots \oplus \beta_p X_{ip} \oplus \Phi_i, \quad i = 1, \dots, n, \tag{15}$$

where $X_{ij} = (l_{x_{ij}}, x_{ij}, r_{x_{ij}})$ and $Y_i = (l_{y_i}, y_i, r_{y_i})$, while β_1, \dots, β_p denote unknown crisp regression parameters to be estimated. Moreover, let $\Phi_i, i = 1, \dots, n$, denote fuzzy error terms which express both randomness and fuzziness allowing negative spreads [8,16,38], i.e., $\Phi_i = (\theta_i^l, \epsilon_i, \theta_i^r)$, where $\theta_i^l, \epsilon_i, \theta_i^r$ are crisp random variables which satisfy the following assumptions:

Assumption 1.

- (A1) $\epsilon_1, \dots, \epsilon_n$ are i.i.d. r.v.’s such that $E(\epsilon_i) = 0$ and $Var(\epsilon_i) = \sigma_{\epsilon}^2 < \infty$
- (A2) $\theta_1^r, \dots, \theta_n^r$ are i.i.d. r.v.’s such that $E(\theta_i^r) = 0$ and $Var(\theta_i^r) = \sigma_r^2 < \infty$
- (A3) θ_1^l, θ_n^l are i.i.d. r.v.’s such that $E(\theta_i^l) = 0$ and $Var(\theta_i^l) = \sigma_l^2 < \infty$
- (A4) ϵ_i, θ_i^r and θ_i^l are mutually uncorrelated.

It can be shown (see [33]) that defining the following design matrix

$$\tilde{X} = \begin{bmatrix} (1, 1, 1) & (l_{x_{11}}, x_{11}, r_{x_{11}}) & \dots & (l_{x_{1p}}, x_{1p}, r_{x_{1p}}) \\ \vdots & \vdots & \ddots & \vdots \\ (1, 1, 1) & (l_{x_{n1}}, x_{n1}, r_{x_{n1}}) & \dots & (l_{x_{np}}, x_{np}, r_{x_{np}}) \end{bmatrix},$$

and a vector $\tilde{\mathbf{y}} = [(l_{y_i}, y_i, r_{y_i})]_{n \times 1} = [(l_{y_1}, y_1, r_{y_1}), \dots, (l_{y_n}, y_n, r_{y_n})]^T$ and assuming that $\det(\tilde{X}^T \odot \tilde{X}) \neq 0$, we obtain the following least squares estimator

$$\hat{\boldsymbol{\beta}} = (\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \odot \tilde{\mathbf{y}}. \tag{16}$$

In the next section, we explore basic properties (16) of the show that it is BLUE. Then, in Section 4 we examine the asymptotic behavior of (16).

4. BLUE in Fuzzy Regression Model

The following theorems provide formulas to find the expectation and variance of our estimator.

Lemma 1. Let $\tilde{\Gamma}_{m \times n} \in M_{\mathcal{F}_T}$, $\tilde{\mathbf{y}}_{n \times 1} = [Y_j] = [(l_{y_j}, y_j, r_{y_j})] \in M_{\mathcal{F}_T}$, where $j = 1, \dots, n$. Then

$$E(\tilde{\Gamma} \odot \tilde{\mathbf{y}}) = \tilde{\Gamma} \odot E(\tilde{\mathbf{y}}). \tag{17}$$

Proof. Let $\tilde{\Gamma} = [Z_{ij}] = [(l_{z_{ij}}, z_{ij}, r_{z_{ij}})]$. Then

$$\begin{aligned} E(\tilde{\Gamma} \odot \tilde{\mathbf{y}}) &= E\left(\left[\sum_{j=1}^n Z_{ij} \odot Y_j\right]\right) \\ &= E\left(\left[\sum_{j=1}^n (l_{z_{ij}} l_{y_j} + z_{ij} y_j + r_{z_{ij}} r_{y_j})\right]\right) \\ &= \left[\sum_{j=1}^n l_{z_{ij}} E(l_{y_j}) + z_{ij} E(y_j) + r_{z_{ij}} E(r_{y_j})\right] \\ &= \left[\sum_{j=1}^n Z_{ij} \odot E(Y_j)\right] = \tilde{\Gamma} \odot [E(Y_j)], \end{aligned}$$

where $E(Y_j) = (E(l_{y_j}), E(y_j), E(r_{y_j}))$ for $j = 1, \dots, n$. \square

Let $Var(l_{y_j}) = \sigma_{l_{y_j}}^2$, $Var(y_j) = \sigma_{y_j}^2$ and $Var(r_{y_j}) = \sigma_{r_{y_j}}^2$. The matrix $\tilde{\Sigma}_{n \times n} \in M_{\mathcal{F}_T}$ is the diagonal matrix with the diagonals $\tilde{\Sigma}_{jj} = (\sigma_{l_{y_j}}^2, \sigma_{y_j}^2, \sigma_{r_{y_j}}^2) = (\sigma_l^2, \sigma_e^2, \sigma_r^2) = \sigma_{\Phi}^2$, while $\tilde{\Sigma}_{jl} = (0, 0, 0)$ for $j, l = 1, \dots, n$ and $j \neq l$.

Lemma 2. Let $\tilde{\Gamma}_{m \times n} \in M_{\mathcal{F}_T}$, $\tilde{\mathbf{y}}_{n \times 1} = [Y_j] = [(l_{y_j}, y_j, r_{y_j})] \in M_{\mathcal{F}_T}$, where $j = 1, \dots, n$. Then

$$Var(\tilde{\Gamma} \odot \tilde{\mathbf{y}}) = (\tilde{\Gamma} \otimes \tilde{\Sigma}) \odot \tilde{\Gamma}^T = \sigma_{\Phi}^2 \odot (\tilde{\Gamma} \otimes \tilde{\Gamma}^T). \tag{18}$$

Proof. Let $\tilde{\Gamma} = [Z_{ij}] = [(l_{z_{ij}}, z_{ij}, r_{z_{ij}})]$ and $\tilde{\mathbf{y}} = [Y_j] = [(l_{y_j}, y_j, r_{y_j})]$, where $i = 1, \dots, m, j = 1, \dots, n$. Since $\tilde{\Gamma} \odot \tilde{\mathbf{y}} = [\sum_{j=1}^n Z_{ij} \odot Y_j]$, we have

$$Var(\tilde{\Gamma} \odot \tilde{\mathbf{y}}) = \left[Cov\left(\sum_{j=1}^n Z_{ij} \odot Y_j, \sum_{j=1}^n Z_{kj} \odot Y_j\right) \right],$$

where $i, k = 1, \dots, m$. The diagonal elements of $Var(\tilde{\Gamma} \odot \tilde{\mathbf{y}})$, i.e., for $i = k$, are $Var\left(\sum_{j=1}^n Z_{ij} \odot Y_j\right)$, where $\sum_{j=1}^n Z_{ij} \odot Y_j = \sum_{j=1}^n (l_{z_{ij}}, z_{ij}, r_{z_{ij}}) \odot (l_{y_j}, y_j, r_{y_j}) = \sum_{j=1}^n (l_{z_{ij}} l_{y_j} + z_{ij} y_j + r_{z_{ij}} r_{y_j})$. Hence, by (A4), we obtain

$$\begin{aligned}
 \text{Var} \left(\sum_{j=1}^n Z_{ij} \odot Y_j \right) &= \text{Var} \left(\sum_{j=1}^n l_{z_{ij}} l_{y_j} + z_{ij} y_j + r_{z_{ij}} r_{y_j} \right) \\
 &= \sum_{j=1}^n (l_{z_{ij}}^2 \sigma_{l_{y_j}}^2 + z_{ij}^2 \sigma_{y_j}^2 + r_{z_{ij}}^2 \sigma_{r_{y_j}}^2) \\
 &= \sum_{j=1}^n \{ (l_{z_{ij}}, z_{ij}, r_{z_{ij}}) \otimes (l_{z_{ij}}, z_{ij}, r_{z_{ij}}) \} \odot (\sigma_l^2, \sigma_e^2, \sigma_r^2) \\
 &= \sum_{j=1}^n (Z_{ij} \otimes Z_{ij}) \odot \sigma_{\Phi}^2.
 \end{aligned}$$

Note that σ_{Φ}^2 is a crisp vector, not a triangular fuzzy number.

If $i \neq k$, by (A1)–(A3), the off-diagonal elements of $\text{Var}(\tilde{\Gamma} \odot \tilde{\mathbf{y}})$ are

$$\begin{aligned}
 \text{Cov} \left(\sum_{j=1}^n Z_{ij} \odot Y_j, \sum_{j=1}^n Z_{kj} \odot Y_j \right) &= \text{Cov} \left(\sum_{j=1}^n l_{z_{ij}} l_{y_j} + z_{ij} y_j + r_{z_{ij}} r_{y_j}, \sum_{j=1}^n l_{z_{kj}} l_{y_j} + z_{kj} y_j + r_{z_{kj}} r_{y_j} \right) \\
 &= \sum_{j=1}^n (l_{z_{ij}} l_{z_{kj}} \sigma_{l_{y_j}}^2 + z_{ij} z_{kj} \sigma_{y_j}^2 + r_{z_{ij}} r_{z_{kj}} \sigma_{r_{y_j}}^2) \\
 &= \sum_{j=1}^n (Z_{ij} \otimes Z_{kj}) \odot \sigma_{\Phi}^2.
 \end{aligned}$$

On the other hand

$$\begin{aligned}
 (\tilde{\Gamma} \otimes \tilde{\Sigma}) \odot \tilde{\Gamma}^T &= [Z_{ij} \otimes \sigma_{\Phi}^2] \odot \tilde{\Gamma}^T \\
 &= \left[\sum_{j=1}^n (Z_{ij} \otimes \sigma_{\Phi}^2) \odot Z_{kj} \right] = \left[\sum_{j=1}^n (Z_{ij} \otimes Z_{kj}) \odot \sigma_{\Phi}^2 \right],
 \end{aligned}$$

where $i, k = 1, \dots, m, j = 1, \dots, n$. Thus, we obtain $\text{Var}(\tilde{\Gamma} \odot \tilde{\mathbf{y}}) = (\tilde{\Gamma} \otimes \tilde{\Sigma}) \odot \tilde{\Gamma}^T$.

Keeping also in mind that $\tilde{\Sigma} = \sigma_{\Phi}^2 \tilde{I}$, where \tilde{I} is the identity in $M_{\mathbb{R}}$, we obtain

$$\begin{aligned}
 (\tilde{\Gamma} \otimes \tilde{\Sigma}) \odot \tilde{\Gamma}^T &= (\tilde{\Gamma} \otimes \sigma_{\Phi}^2 \tilde{I}) \odot \tilde{\Gamma}^T = (\sigma_{\Phi}^2 \otimes (\tilde{\Gamma} \tilde{I})) \odot \tilde{\Gamma}^T \\
 &= (\sigma_{\Phi}^2 \otimes \tilde{\Gamma}) \odot \tilde{\Gamma}^T = \sigma_{\Phi}^2 \odot (\tilde{\Gamma} \otimes \tilde{\Gamma}^T).
 \end{aligned}$$

Thus the proof is completed. \square

Corollary 1. Let $\hat{\beta}$ be the least squares estimator of β . Then

$$\text{Var}(\hat{\beta}) = \sigma_{\Phi}^2 \odot \left((\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \otimes \tilde{X}) (\tilde{X}^T \odot \tilde{X})^{-1} \right).$$

Proof. By (16) we have $\hat{\beta} = (\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \odot \tilde{\mathbf{y}} = ((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T) \odot \tilde{\mathbf{y}}$, where $(\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \in M_{\mathcal{F}_T}$. Hence, by Lemma 2, we get

$$\begin{aligned}
 \text{Var}(\hat{\beta}) &= \sigma_{\Phi}^2 \odot \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \right) \otimes \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \right)^T \\
 &= \sigma_{\Phi}^2 \odot \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \right) \otimes \left(\tilde{X} (\tilde{X}^T \odot \tilde{X})^{-1} \right) \\
 &= \sigma_{\Phi}^2 \odot \left((\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \otimes \tilde{X}) (\tilde{X}^T \odot \tilde{X})^{-1} \right).
 \end{aligned}$$

which completes the proof. \square

Theorem 1. Let $\hat{\beta}$ be the least squares estimator given by (16). Then $\hat{\beta}$ is the unbiased estimator of β .

Proof. Firstly, we want to show that $E(\tilde{\mathbf{y}}) = \tilde{X}\beta$. Since

$$\begin{aligned}
 l_{y_i} &= \beta_0 + \beta_1 l_{x_{i1}} + \cdots + \beta_{ip} l_{x_{ip}} + \theta_i^l = \sum_{j=0}^p \beta_j l_{x_{ij}} + \theta_i^l, \\
 y_i &= \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{ip} x_{ip} + \epsilon_i = \sum_{j=0}^p \beta_j x_{ij} + \epsilon_i, \\
 r_{y_i} &= \beta_0 + \beta_1 \zeta_{i1}^r + \cdots + \beta_{ip} \zeta_{ip}^r + \theta_i^r = \sum_{j=0}^p \beta_j r_{x_{ij}} + \theta_i^r,
 \end{aligned}$$

thus, by Assumptions A we obtain

$$\begin{aligned}
 E(\tilde{\mathbf{y}}) &= [E(Y_j)] = [(E(l_{y_j}), E(y_j), E(r_{y_j}))] \\
 &= [(\sum_{j=0}^p \beta_j l_{x_{ij}}, \sum_{j=0}^p \beta_j x_{ij}, \sum_{j=0}^p \beta_j r_{x_{ij}})] = \tilde{X}\beta.
 \end{aligned}$$

Now, by Lemma 1, we conclude that

$$\begin{aligned}
 E(\hat{\beta}) &= E\left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \odot \tilde{\mathbf{y}}\right) = E\left(\left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T\right) \odot \tilde{\mathbf{y}}\right) \\
 &= \left(\left(\tilde{X}^T \odot \tilde{X}\right)^{-1} \tilde{X}^T\right) \odot E(\tilde{\mathbf{y}}) = \left(\left(\tilde{X}^T \odot \tilde{X}\right)^{-1} \tilde{X}^T\right) \odot \tilde{X}\beta \\
 &= (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \odot \tilde{X})\beta = \beta,
 \end{aligned}$$

which means that $\hat{\beta}$ is the unbiased estimator of β . \square

Now we are able to state the main theorem in this section.

Theorem 2. The least squares estimator $\hat{\beta}$, given by (16), is the best linear unbiased estimator of β .

Proof. It is clear that $\hat{\beta} = (\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \odot \tilde{\mathbf{y}}$ is the linear estimator with the fuzzy operation \odot , i.e., fuzzy-type linear estimator. By Theorem we also know that $\hat{\beta}$ is unbiased. Therefore, it is enough to prove that $\hat{\beta}$ has the minimum variance among the linear unbiased estimators of β .

Let $\hat{\beta}^*$ be arbitrary linear unbiased estimator for β . Then

$$\hat{\beta}^* = ((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T) \odot \tilde{\mathbf{y}} + \tilde{\Lambda} \odot \tilde{\mathbf{y}} = \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \oplus \tilde{\Lambda}\right) \odot \tilde{\mathbf{y}}$$

for some $\tilde{\Lambda} \in M_{\mathcal{F}_T}$. By Corollary 1 we have

$$\begin{aligned} \text{Var}(\hat{\beta}^*) &= \text{Var} \left\{ \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \oplus \tilde{\Lambda} \right) \odot \tilde{\mathbf{y}} \right\} \\ &= \sigma_{\Phi}^2 \odot \left\{ \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \oplus \tilde{\Lambda} \right) \otimes \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \oplus \tilde{\Lambda} \right)^T \right\} \\ &= \sigma_{\Phi}^2 \odot \left\{ \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \oplus \tilde{\Lambda} \right) \otimes \left(\tilde{X}(\tilde{X}^T \odot \tilde{X})^{-1} \oplus \tilde{\Lambda}^T \right) \right\} \\ &= \sigma_{\Phi}^2 \odot \left\{ (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \otimes \tilde{X}) (\tilde{X}^T \odot \tilde{X})^{-1} \right. \\ &\quad \left. + \sigma_{\Phi}^2 \odot \left\{ (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \otimes \tilde{\Lambda}^T) \right. \right. \\ &\quad \left. \left. + (\tilde{\Lambda} \otimes \tilde{X}) (\tilde{X}^T \odot \tilde{X})^{-1} + (\tilde{\Lambda} \otimes \tilde{\Lambda}^T) \right\} \right\}. \end{aligned} \tag{19}$$

Since

$$\begin{aligned} E(\hat{\beta}^*) &= \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \oplus \tilde{\Lambda} \right) \odot E(\tilde{\mathbf{y}}) \\ &= \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \oplus \tilde{\Lambda} \right) \odot (\tilde{X}\beta) \\ &= (\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \odot (\tilde{X}\beta) + \tilde{\Lambda} \odot (\tilde{X}\beta) \\ &= (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \odot \tilde{X})\beta + (\tilde{\Lambda} \odot \tilde{X})\beta \\ &= \beta + (\tilde{\Lambda} \odot \tilde{X})\beta, \end{aligned}$$

hence, by the unbiasedness of $\hat{\beta}^*$, it means $\tilde{\Lambda} \odot \tilde{X} = \tilde{O}$, where \tilde{O} is the zero matrix in M_R . Let $\tilde{\Lambda} = [Z_{ij}] = [(l_{z_{ij}}, z_{ij}, r_{z_{ij}})]$, where $i = 1, \dots, p + 1$ and $j = 1, \dots, n$. Then

$$\tilde{\Lambda} \odot \tilde{X} = \left[\sum_{k=1}^n Z_{ik} \odot X_{kl} \right] = \left[\sum_{k=1}^n (l_{z_{ik}} l_{x_{kl}} + z_{ik} x_{kl} + r_{z_{ik}} r_{x_{kl}}) \right] = \tilde{O},$$

where $i, l = 1, \dots, p + 1$. So $\sum_{k=1}^n (l_{z_{ik}} l_{x_{kl}} + z_{ik} x_{kl} + r_{z_{ik}} r_{x_{kl}}) = 0$ for all $i, l = 1, \dots, p + 1$. Since we assumed $\tilde{\Lambda}, \tilde{X} \in M_{\mathcal{F}_T} = M_{\mathcal{F}_T(\mathcal{R}^+)}$, hence all components are non-negative, i.e., $l_{z_{ik}} l_{x_{kl}} + z_{ik} x_{kl} + r_{z_{ik}} r_{x_{kl}} = 0$ for all $k = 1, \dots, n$. Thus we conclude that $l_{z_{ik}} l_{x_{kl}} + z_{ik} x_{kl} + r_{z_{ik}} r_{x_{kl}} = 0$ for all $k = 1, \dots, n$. Moreover, since $l_{z_{ik}} l_{x_{kl}} = 0, z_{ik} x_{kl} = 0$ and $r_{z_{ik}} r_{x_{kl}} = 0$, hence $l_{z_{ik}} l_{x_{kl}} = 0, z_{ik} x_{kl} = 0$ and $r_{z_{ik}} r_{x_{kl}} = 0$, so $Z_{ik} \odot X_{kl} = (l_{z_{ik}} l_{x_{kl}}, z_{ik} x_{kl}, r_{z_{ik}} r_{x_{kl}}) = (0, 0, 0)$ for all $k = 1, \dots, n$. Thus

$$\tilde{\Lambda} \otimes \tilde{X} = \left[\sum_{k=1}^n Z_{ik} \otimes X_{kl} \right] = \left[\sum_{k=1}^n (l_{z_{ik}} l_{x_{kl}}, z_{ik} x_{kl}, r_{z_{ik}} r_{x_{kl}}) \right] = [(0, 0, 0)] = \tilde{\Theta},$$

where $\tilde{\Theta}$ is the zero triangular fuzzy matrix in $M_{\mathcal{F}_T}$ which has all elements $(0, 0, 0) \in \mathcal{F}_T$. Of course, $\tilde{X}^T \otimes \tilde{\Lambda}^T = (\tilde{\Lambda} \otimes \tilde{X})^T$. Therefore, Equation (19) reduces to

$$\begin{aligned} \text{Var}(\hat{\beta}^*) &= \sigma_{\Phi}^2 \odot \left((\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \otimes \tilde{X}) (\tilde{X}^T \odot \tilde{X})^{-1} + (\tilde{\Lambda} \otimes \tilde{\Lambda}^T) \right) \\ &= \text{Var}(\hat{\beta}) + \sigma_{\Phi}^2 \odot (\tilde{\Lambda} \otimes \tilde{\Lambda}^T). \end{aligned}$$

Let $\hat{\beta}^* = [\hat{\beta}_0^*, \hat{\beta}_1^*, \dots, \hat{\beta}_p^*]^T$. Then $\text{Var}(\hat{\beta}_k^*)$ for $k = 0, \dots, p$, appear on the diagonal of $\text{Var}(\hat{\beta}^*)$. The diagonal elements of $\tilde{\Lambda} \otimes \tilde{\Lambda}^T$ are $\sum_{k=1}^n Z_{ik} \otimes Z_{ik} = (\sum_{k=1}^n l_{z_{ik}}^2, \sum_{k=1}^n z_{ik}^2, \sum_{k=1}^n r_{z_{ik}}^2)$, where $i = 1, \dots, n$. Thus the diagonal elements of $\sigma_{\Phi}^2 \odot (\tilde{\Lambda} \otimes \tilde{\Lambda}^T)$ are $\sum_{k=1}^n \sigma_l^2 l_{z_{ik}}^2 + \sum_{k=1}^n z_{ik}^2 \sigma_{\epsilon}^2 + \sum_{k=1}^n r_{z_{ik}}^2 \sigma_r^2$, where $i = 1, \dots, n$. Hence,

$$\text{Var}(\hat{\beta}_k^*) = \text{Var}(\hat{\beta}_k) + \sum_{k=1}^n \sigma_l^2 l_{z_{ik}}^2 + \sum_{k=1}^n z_{ik}^2 \sigma_{\epsilon}^2 + \sum_{k=1}^n r_{z_{ik}}^2 \sigma_r^2.$$

For any linear unbiased estimator $\hat{\beta}^*$ we get $Var(\hat{\beta}_k^*)Var(\hat{\beta}_k)$ for all $k = 0, \dots, p$. Thus, $Var(\hat{\beta}_k)$ has the minimum variance among the linear unbiased estimators for all $k = 0, \dots, p$. This proves the theorem. \square

Example 1. With the modified data of [7], in Table 1, we evaluate the covariance matrix of some fuzzy-type linear estimators to compare the variances of some estimators. The data involves student grades and family income. The data were fuzzified as triangular fuzzy numbers. We compare four fuzzy-type linear estimators.

Table 1. Numerical data for an example.

| ID | $Y_i = (l_{y_i}, y_i, r_{y_i})$ | $X_i = (l_{x_i}, x_i, r_{x_i})$ |
|----|---------------------------------|---------------------------------|
| 1 | (3.40, 4.00, 4.80) | (16.80, 21.00, 23.10) |
| 2 | (2.70, 3.00, 3.30) | (12.75, 15.00, 17.25) |
| 3 | (3.15, 3.50, 3.85) | (13.50, 15.00, 17.25) |
| 4 | (1.60, 2.00, 2.40) | (7.65, 9.00, 10.35) |
| 5 | (2.70, 3.00, 3.45) | (10.80, 12.00, 13.20) |
| 6 | (2.97, 3.50, 4.20) | (14.40, 18.00, 19.80) |
| 7 | (2.25, 2.50, 2.88) | (5.40, 5.50, 7.20) |
| 8 | (2.00, 2.50, 3.00) | (10.20, 12.00, 14.40) |

$\hat{\beta}$ is the our estimator and $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ are the several modified four fuzzy-type linear estimators. We define the estimators as followings:

$$\hat{\beta} = (\tilde{X}^t \diamond \tilde{X})^{-1} \tilde{X}^t \diamond \tilde{y},$$

$$\hat{\beta}_1 = (\tilde{X}^t \diamond \tilde{X})^{-1} \tilde{X}^t \diamond \tilde{y} \oplus (\tilde{A} \diamond \tilde{y}),$$

$$\hat{\beta}_2 = (\tilde{X}^t \diamond \tilde{X})^{-1} \tilde{X}^t \diamond \tilde{y} \oplus (\tilde{B} \diamond \tilde{y}),$$

and

$$\hat{\beta}_3 = (\tilde{X}^t \diamond \tilde{X})^{-1} \tilde{X}^t \diamond \tilde{y} \oplus (\tilde{C} \diamond \tilde{y}),$$

where $\tilde{A} = [A_{ij}]$, $\tilde{B} = [B_{ij}]$ and $\tilde{C} = [C_{ij}]$ and $A_{ij} = (a_{ij}^l, a_{ij}^m, a_{ij}^r)$, $B_{ij} = (b_{ij}^l, b_{ij}^m, b_{ij}^r)$ and $C_{ij} = (c_{ij}^l, c_{ij}^m, c_{ij}^r)$ ($i = 1, 2, 3$ and $j = 1, \dots, 10$). a_{ij}^m is the mode of A_{ij} , a_{ij}^l, a_{ij}^r are the left and right spreads of A_{ij} , respectively. We define all $a_{ij}^l = a_{ij}^m = a_{ij}^r = 0$ except $a_{19}^l = -0.1$, <https://www.overleaf.com/project/5d494ddde255ad39c9816390> $a_{25}^m = 0.1$ and $a_{32}^m = -0.1$. All $b_{ij}^l = b_{ij}^m = b_{ij}^r = 0$ except $b_{12}^m = 0.01, b_{25}^r = 0.01$ and $b_{39}^l = 0.01$, and $c_{ij}^l = c_{ij}^m = c_{ij}^r = 0$ for all i, j , except $c_{14}^l = 0.1, c_{27}^m = -0.1$ and $c_{3,10}^r = 0.1$. We get the estimates: $\hat{\beta} = [11.9986 \ 4.0970 \ 2.9970]^t$, $\hat{\beta}_1 = [11.5416 \ 4.8970 \ 2.4650]^t$, $\hat{\beta}_2 = [12.5306 \ 5.3540 \ 2.9970]^t$, and $\hat{\beta}_3 = [18.1686 \ 9.1170 \ 8.8870]^t$.

By applying Theorems 1 and 2, we obtain the covariance matrices of the estimators as follows:

$$\Sigma = \begin{bmatrix} 1.2366 & -0.0831 & -0.1110 \\ -0.0831 & 0.0090 & 0.0035 \\ -0.1110 & 0.0035 & 0.0155 \end{bmatrix},$$

$$\Sigma_1 = \begin{bmatrix} 1.2367 & -0.0831 & -0.1110 \\ -0.0831 & 0.0091 & 0.0035 \\ -0.1110 & 0.0035 & 0.0156 \end{bmatrix},$$

$$\Sigma_2 = \begin{bmatrix} 1.2367 & -0.0831 & -0.1110 \\ -0.0831 & 0.0092 & 0.0035 \\ -0.1110 & 0.0035 & 0.0155 \end{bmatrix},$$

and

$$\Sigma_3 = \begin{bmatrix} 1.2466 & -0.0831 & -0.1110 \\ -0.0831 & 0.0190 & 0.0035 \\ -0.1110 & 0.0035 & 0.0255 \end{bmatrix}.$$

$Var(\hat{\beta}_0) = 1.2366$ is smaller than 1.2367 and 1.2466. $Var(\hat{\beta}_1) = 0.0090$ is the smallest among 0.0090, 0.0091, 0.0092 and 0.0190. Finally, $Var(\hat{\beta}_2) = 0.0155$ is the smaller than 0.0156, 0.0255. Hence we conclude that the estimators $\hat{\beta}$ have the minimum variances among $\hat{\beta}, \hat{\beta}_1, \hat{\beta}_2$ and $\hat{\beta}_3$.

5. Asymptotic Normality

We start this section by citing some theorems concerning the Central Limit Theorem (CLT) and Strong Law of Large Numbers (SLLN) for martingales which will be useful in the proof of the main result.

Theorem 3 (SLLN for martingales). *Let $S_n = \sum_{i=1}^n X_i, n \geq 1$, be a martingale such that $E|X_k|^p < \infty$ for $k \geq 1$ and $1 \leq p \leq 2$. Suppose that $\{b_n\}$ is a sequence of positive constants increasing to ∞ as $n \rightarrow \infty$, and $\sum_{i=1}^n E[X_i^2]/b_i^2 < \infty$. Then $S_n/b_n \xrightarrow{a.s.} 0$, where the notation $\xrightarrow{a.s.}$ means the almost surely converges.*

Theorem 4 (Hajék–Sidák CLT). *Let $\{X_n\}$ be a sequence of i.i.d. random variables (r.v.'s) with the mean μ and a finite variance σ^2 . Let $\{c_n\}$ be a sequence of real vectors $c_n = (c_{n1}, \dots, c_{nm})^T$. If*

$$\left(\max_{1 \leq i \leq n} c_{ni}^2 \right) \left(\sum_{i=1}^n c_{ni}^2 \right)^{-1} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

then

$$Z_n = \frac{\sum_{i=1}^n c_{ni}(X_i - \mu)}{\sigma^2 \sum_{i=1}^n c_{ni}^2} \xrightarrow{L} N(0, 1),$$

where the notation \xrightarrow{L} stands for the convergence in law.

The proof of Theorem 4 can be found in [39].

Theorem 5 (Courant–Fisher minimax theorem). *For any $n \times n$ real symmetric matrix A its eigenvalues $\lambda_1 \dots \lambda_n$ satisfy*

$$\lambda_k = \min_{\dim(C)=k} \max_{\|x\|=1, x \in C} \langle Ax, x \rangle,$$

where C is a subspace of \mathbb{R}^n .

To obtain some asymptotic properties of our least squares estimator in the generalized fuzzy regression model, the following additional assumption is required besides Assumption A given in Section 3:

Assumption 2.

- (B1) $\max_{1 \leq i \leq n} (\tilde{x}_i^T (X^T \odot \tilde{X})^{-1}) \odot \tilde{x}_i \rightarrow 0$ as $n \rightarrow \infty$, where \tilde{x}_i^T denotes the i -th row in a fuzzy matrix \tilde{X} .
- (B2) $n(\tilde{X}^T \odot \tilde{X})^{-1}(\tilde{X}^T \otimes \tilde{X})(\tilde{X}^T \odot \tilde{X})^{-1} \rightarrow \tilde{\Gamma}$ as $n \rightarrow \infty$ for some $\tilde{\Gamma} \in M_{\mathcal{F}_T}$.

Now we are able to formulate one of the main results of this contribution. Let $\hat{\beta}_n$ be the estimator for β of size n .

Theorem 6. If model (15) satisfies Assumptions A and B, then the least squares estimator $\hat{\beta}_n$ is asymptotically normal, i.e.,

$$\sqrt{n} \left(\hat{\beta}_n - \beta \right) \xrightarrow{L} N_{p+1} \left(\mathbf{0}, \sigma_{\Phi}^2 \odot \tilde{\Pi} \right), \tag{20}$$

where $\sigma_{\Phi}^2 = (\sigma_l^2, \sigma_e^2, \sigma_r^2)$.

Proof. By (16) one can find that

$$\begin{aligned} \hat{\beta}_n &= (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \odot \tilde{y}) = \left((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \right) \odot \tilde{y} \\ &= (\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \odot (\tilde{X}\beta + \Phi) = \beta + (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \odot \Phi) \\ &= \left(\beta + (\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \right) \odot \Phi, \end{aligned}$$

so, consequently, $\hat{\beta}_n - \beta = ((\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T) \odot \Phi$.

Let $\lambda_n \in \mathbb{R}^{p+1} (\lambda_n \neq \mathbf{0})$ be an arbitrary but fixed vector. Moreover, let $Z_n = \lambda_n^T (\hat{\beta}_n - \beta) = \tilde{C}_n^T \odot \Phi \in \mathbb{R}$, where $\tilde{C}_n = \tilde{X} (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n \in M_{\mathcal{F}_T}$. If we denote $\tilde{C}_n^T = [C_{n1}, \dots, C_{nm}]$, where $C_{n1}, \dots, C_{nm} \in \mathcal{F}_T$, then (see [33])

$$\begin{aligned} \sum_{i=1}^n C_{ni} \odot C_{ni} &= \tilde{C}_n^T \odot \tilde{C}_n \\ &= \left(\lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} \tilde{X}^T \right) \odot \left(\tilde{X} (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n \right) \\ &= \lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \odot \tilde{X}) (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n \\ &= \lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n. \end{aligned}$$

We claim that \tilde{C}_n satisfies the regularity condition of Theorem 3. Then we can obtain the asymptotic distribution of Z_n . Let \tilde{x}_i^T be the i -th row of \tilde{X} . Then we get $C_{ni} = \mathbf{x}_i^T (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n$. Since $C_{ni} \in \mathcal{F}_T$ we have $C_{ni}^T = C_{ni}$. Hence

$$\begin{aligned} C_{ni} \odot C_{ni} &= C_{ni}^T \odot C_{ni} \\ &= (\lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} \mathbf{x}_i) \odot (\mathbf{x}_i^T (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n) \\ &= \lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} (\mathbf{x}_i \odot \mathbf{x}_i^T) (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n. \end{aligned}$$

Therefore, by Theorem 5 (see [33])

$$\sup_{\lambda_n} \frac{\lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{x}_i \odot \tilde{x}_i^T) (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n}{\lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n} \tag{21}$$

becomes

$$\text{ch}_{\max} [(\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{x}_i \odot \tilde{x}_i^T)] = \left(\tilde{x}_i^T (\tilde{X}^T \odot \tilde{X})^{-1} \right) \odot \tilde{x}_i,$$

where $\text{ch}_{\max}(Q)$ stands for the largest characteristic value of matrix Q . Thus

$$\begin{aligned} &\sup_{\lambda_n} \max_i \frac{C_{ni} \odot C_{ni}}{\sum_{i=1}^n C_{ni} \odot C_{ni}} \\ &= \max_i \sup_{\lambda_n} \frac{\lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{x}_i \odot \tilde{x}_i^T) (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n}{\lambda_n^T (\tilde{X}^T \odot \tilde{X})^{-1} \lambda_n} \\ &= \max_i \left(\tilde{x}_i^T (\tilde{X}^T \odot \tilde{X})^{-1} \right) \odot \tilde{x}_i, \end{aligned}$$

which, by Assumption (B1), converges to 0 as $n \rightarrow \infty$. It means that

$$\max_{1 \leq i \leq n} \frac{C_{ni} \odot C_{ni}}{\sum_{i=1}^n C_{ni} \odot C_{ni}} \rightarrow 0$$

as $n \rightarrow \infty$. So, consequently,

$$Z_n = \lambda_n^T (\hat{\beta}_n - \beta) \xrightarrow{L} N(0, 1)$$

by the Hajék–Sidák Central Limit Theorem 4.

On the other hand, one may notice that

$$\text{Var} \left(\sqrt{n}(\hat{\beta}_n - \beta) \right) = n\sigma_{\Phi}^2 \odot \left((\tilde{X}^T \odot \tilde{X})^{-1} (\tilde{X}^T \otimes \tilde{X}) (\tilde{X}^T \odot \tilde{X})^{-1} \right) \rightarrow \sigma_{\Phi}^2 \odot \tilde{\Pi},$$

as $n \rightarrow \infty$ (by Assumption B2). Thus

$$\sqrt{n} \left(\hat{\beta}_n - \beta \right) \xrightarrow{L} N_{p+1} \left(\mathbf{0}, \sigma_{\Phi}^2 \odot \tilde{\Pi} \right),$$

which completes the proof. \square

6. Strong Consistency and Confidence Region

The weak consistency is a direct consequence of the asymptotic normality.

Theorem 7. For the model (1), under Assumptions A and B, the $\hat{\beta}$ is weakly consistent estimator for β , that is

$$\hat{\beta}_n \xrightarrow{P} \beta,$$

where the notation \xrightarrow{P} means converges in probability.

Proof. From the fact that $\sqrt{n} \left(\hat{\beta}_n - \beta \right)$ converges a random variable in law, which is non-degenerated, it follows that $\sqrt{n} \left(\hat{\beta}_n - \beta \right) = O_P(\mathbf{1})$, where $O_P(\mathbf{1})$ stands for bounded in probability. This implies that each $\hat{\beta}_i$ is weakly consistent, and then $\hat{\beta}_n - \beta \xrightarrow{P} \mathbf{0}$. \square

One of main results of this section is the following theorem, the strong consistency of fuzzy least squares estimators. Moreover, this theorem explains that for strong consistency property, asymptotic normality is not needed, and hence, some of the following theorems may be relaxed.

Theorem 8. For the model (15), suppose that Assumptions A and B are fulfilled. Furthermore, assume that $s_j^2 = \sum_{i=1}^n (l_{x_{ij}} + x_{ij} + r_{x_{ij}})^2 \rightarrow \infty$ ($j = 0, 1, \dots, p$) as $n \rightarrow \infty$, and the sequence of matrices $(\tilde{X}^t \diamond \tilde{X})^{-1} \text{diag}(s_0^2, \dots, s_p^2)$ is bounded. Then $\hat{\beta}_n$ defined on (16) is strongly consistent for β , that is

$$\hat{\beta}_n \xrightarrow{a.s.} \beta, \tag{22}$$

where the notation $\xrightarrow{a.s.}$ means converges in almost surely.

Proof. From the fact that $\hat{\beta}_n - \beta = ((\tilde{X}^t \diamond \tilde{X})^{-1} \tilde{X}^t) \diamond \Phi$ and under the assumption $s_j^2 \rightarrow \infty$ for all j , it can be proved by Theorem 3 that

$$\text{diag}(1/s_0^2, \dots, 1/s_p^2) \tilde{X}^t \diamond \Phi \rightarrow \mathbf{0}$$

almost surely. The result of the theorem is now an obvious consequence of the assumption. The proof is completed. \square

Next, we provide an approximate confidence region for β based on the large-sample normality of the FLSE. The asymptotic normality of $\sqrt{n}(\hat{\beta}_n - \beta)$, derived in Theorem 6, under the regularity conditions, suggests the use of the pivotal quantity of the form

$$Q_n(\hat{\beta}_n) = n(\hat{\beta}_n - \beta)^t (\sigma_\Phi^2 \diamond \tilde{\Gamma}_n)^{-1} (\hat{\beta}_n - \beta),$$

where $\tilde{\Gamma}_n = n(\tilde{X}^t \diamond \tilde{X})^{-1} (\tilde{X}^t \otimes \tilde{X}) (\tilde{X}^t \diamond \tilde{X})^{-1}$.

The following theorem gives a large sample distribution of $Q_n(\hat{\beta}_n)$.

Corollary 2. Under the conditions of Theorem 6, $Q_n(\hat{\beta}_n)$ has asymptotically a chi-squared distribution with $p + 1$ degrees of freedom.

Proof. Corollary 2 follows immediately from Theorem 6. \square

Corollary 3. We can define a confidence region $C_{1-\alpha}(\beta)$ as the set of $\hat{\beta}_n$ such that

$$(\hat{\beta}_n - \beta)^t (\sigma_\Phi^2 \diamond \tilde{\Gamma}_n)^{-1} (\hat{\beta}_n - \beta) \leq \delta, \tag{23}$$

where $\delta = \frac{1}{n} \chi_{1-\alpha}^2(p + 1)$ and $\chi_{1-\alpha}^2(p + 1)$ is $(1 - \alpha)$ th the quantile of the chi-squared distribution. Then, for large n , $C_{1-\alpha}(\beta)$ provides a $100(1 - \alpha)$ percent confidence region for β .

Proof. It is directly obtained from Corollary 2. \square

Remark 1. Note that it is well known that under certain regularity conditions, the sequence of the crisp LSE's $\check{\beta}_n$ has asymptotically a normal distribution in the sense that

$$\sqrt{n}(\check{\beta}_n - \beta) \xrightarrow{L} N_{p+1}(\mathbf{0}, \sigma_\epsilon^{*2} V^{-1}), \tag{24}$$

where σ_ϵ^{*2} is the variance of errors ϵ_i in model and V is given by $V_n = \frac{1}{n}(\mathbf{X}'\mathbf{X}) \rightarrow V$ as $n \rightarrow \infty$. Thus, a $100(1 - \alpha)$ percent approximate confidence region based on the LSE, denoted by $C_{1-\alpha}^*(\beta)$, is the set of $\check{\beta}_n$ such that

$$(\check{\beta}_n - \beta)^t V (\check{\beta}_n - \beta) \leq \delta^*, \tag{25}$$

where $\delta^* = \frac{\sigma_\epsilon^{*2}}{n} \chi_{1-\alpha}^2(p)$ and $\sigma_\epsilon^{*2} = \text{Var}[\epsilon]$. Then, for n large $C_{1-\alpha}^*(\beta)$ provides a $100(1 - \alpha)$ percent confidence region for β .

Now we compare the sequences of the FLSE $\{\hat{\beta}_n\}$ and the classical crisp LSE $\{\check{\beta}_n\}$. A numerical measure of the asymptotic relative efficiency (ARE) of $\{\hat{\beta}_n\}$ with respect to $\{\check{\beta}_n\}$ on the inverse ratio of their generalized limiting variances, and denote it by $e_{(F,C)}$, which implies a strictly smaller asymptotic confidence region, is given by [40]

$$e_{(F,C)} = \left(\frac{|\sigma_\epsilon^{*2} V^{-1}|}{|\sigma_\Phi^2 \diamond \tilde{\Gamma}|} \right)^{\frac{1}{p+1}}. \tag{26}$$

If the ARE of the FLSE's $\{\hat{\beta}_n\}$ with respect to the crisp LSE $\{\check{\beta}_n\}$ is greater than 1, we can then say that $\{\hat{\beta}_n\}$ is more efficiency than $\{\check{\beta}_n\}$.

Example 2. The dataset from in Table 1 is used asan example of ARE. If we put $\sigma_{\theta^l}^2 = \sigma_{\theta^r}^2 = \sigma_{\epsilon}^2 = 1$ and $\sigma_{\epsilon^{*2}} = 1$, then we obtain

$$e_{(F,C)} = \left(\frac{|\sigma_{\epsilon^{*2}}^2 V^{-1}|}{|\sigma_{\Phi}^2 \diamond \tilde{\Gamma}|} \right)^{\frac{1}{p+1}} = \left(\frac{0.0494}{0.0058} \right)^{\frac{1}{2}} = 2.9101.$$

In this case, it can be concluded that our estimator $\{\hat{\beta}_n\}$ is more efficient than $\{\check{\beta}_n\}$. Let us regard the data as crisp numbers, i.e., $Y_i = (y_i, y_i, y_i)$ and $X_i = (x_i, x_i, x_i)$. We take $\sigma_{\theta^l}^2 = \sigma_{\theta^r}^2 = 0$, $\sigma_{\epsilon}^2 = 1$ and $\sigma_{\epsilon^{*2}} = 1$. Then we have

$$e_{(F,C)} = \left(\frac{|\sigma_{\epsilon^{*2}}^2 V^{-1}|}{|\sigma_{\Phi}^2 \diamond \tilde{\Gamma}|} \right)^{\frac{1}{p+1}} = \left(\frac{0.0494}{0.0555} \right)^{\frac{1}{2}} = 0.9435.$$

We can verify that the efficiency is approximately 1, as we wished.

Example 3 (Monte Carlo Simulation). We performed Monte Carlo simulations to examine the performance of the proposed estimator with fuzzy observations discussed in this paper. The asymptotic behavior and the accuracies for some finite sample datasets are investigated. Two independent variables of parameters are chosen for this simulation: $(\beta_0, \beta_1) = (1.50, 0.20)$. For the two independent variables, the spreads and the modes are selected from the normal distributions $N(15, 4^2)$, $N(2, 0.5^2)$, respectively. In addition, the measurement errors of the modes and spreads are chosen to be Gaussian white noise with mean zero and variance 0.25: $\Phi_i = (\theta_i^l, \epsilon_i, \theta_i^r)$, $\epsilon_i, \theta_i^l, \theta_i^r \sim \text{NID}(0, 0.5^2)$ “(i = 1, . . . , n). In this paper, sample sizes for n = 10, 50, 100 have been used for the small, moderate and large samples. Here 1000 different datasets were generated for a particular sample size n. For each data set we estimated the parameters β_0, β_1 by the proposed estimators, and provided the average estimates and average mean squared error (MMSE) over 1000 simulations. These results are shown in Table 2. In addition, the percentiles, minimum, 1st quartile, median, 3rd quartile, 95 percent point, and the maximum of the 1000 estimation errors are given in Table 3 and 4.

Table 2. Average estimates and average mean squared errors.

| | n = 10 | (MMSE) | n = 50 | (MMSE) | n = 100 | (MMSE) |
|-----------------|--------|----------|--------|----------|---------|----------|
| $\hat{\beta}_0$ | 1.50 | (0.0124) | 1.50 | (0.0124) | 1.50 | (0.0124) |
| $\hat{\beta}_1$ | 0.20 | (0.0124) | 0.20 | (0.0124) | 0.20 | (0.0124) |

Table 3. Characteristics of estimate errors for $\hat{\beta}_0$.

| n | Min. | 1st Quar. | Median | 3rd Quar. | 95% | Max. |
|-----|------|-----------|--------|-----------|------|--------|
| 10 | 1.50 | 0.0124 | 1.50 | 0.0124 | 1.50 | 0.0124 |
| 50 | 0.20 | 0.0124 | 0.20 | 0.0124 | 0.20 | 0.0124 |
| 100 | 0.20 | 0.0124 | 0.20 | 0.0124 | 0.20 | 0.0124 |

Table 4. Characteristics of estimate errors for $\hat{\beta}_1$.

| n | Min. | 1st Quar. | Median | 3rd Quar. | 95% | Max. |
|-----|------|-----------|--------|-----------|------|--------|
| 10 | 1.50 | 0.0124 | 1.50 | 0.0124 | 1.50 | 0.0124 |
| 50 | 0.20 | 0.0124 | 0.20 | 0.0124 | 0.20 | 0.0124 |
| 100 | 0.20 | 0.0124 | 0.20 | 0.0124 | 0.20 | 0.0124 |

Thus, the accuracy of estimators was assessed for several values of sample sizes, our simulation results indicate that our estimation procedures have smaller mean bias and smaller mean squared error in the estimates of parameters as the sample size increases.

7. Conclusions

In this paper consider a multiple fuzzy-input–fuzzy-output regression model with fuzzy error terms, described by a suitable matrix, called the triangular fuzzy matrix. We proposed a simple formula for the fuzzy least squares estimator and examine its fundamental properties. It appears that the suggested estimator is BLUE so it is in some sense optimal. Moreover, due to its asymptotical normality under quite general assumptions, we open a new perspective for constructing statistical tests and confidence intervals useful both in the model validation and in forecasting. These very topics will be considered in further research. Another open problem to be undertaken is to discover the analogous results in more general fuzzy regression models based on trapezoidal fuzzy numbers, LR-fuzzy numbers and so on.

Author Contributions: J.H.Y. developed the conceptualization, proving the theorems, data analysis and drafted the manuscript. Moreover, investigation, reviewing, and editing were done by P.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2020R1A2C1A01011131).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [[CrossRef](#)]
2. Tanaka, H.; Uejima, S.; Asai, K. Linear regression analysis with fuzzy model. *IEEE Trans. Syst. Man Cybernet* **1982**, *12*, 903–907.
3. Diamond, P. Fuzzy Least Squares. *Inform. Sci.* **1988**, *46*, 141–157. [[CrossRef](#)]
4. Bargiela, A.; Pedrycz, W.; Nakashima, T. Multiple regression with fuzzy data. *Fuzzy Sets Syst.* **2007**, *158*, 2169–2188. [[CrossRef](#)]
5. Celminš, A. Least squares model fitting to fuzzy vector data. *Fuzzy Sets Syst.* **1987**, *22*, 245–269. [[CrossRef](#)]
6. Choi, S.H.; Yoon, J.H. General fuzzy regression using least squares method. *Int. J. Syst. Sci.* **2010**, *4*, 477–485. [[CrossRef](#)]
7. Choi, S.H.; Jung, H.Y.; Lee, W.J.; Yoon, J.H. Fuzzy regression model with monotonic response function. *Commun. Korean Math. Soc.* **2018**, *3*, 973–983.
8. Diamond, P.; Körner, R.K. Extended fuzzy linear models and least squares estimates. *Fuzzy Sets Syst.* **1997**, *33*, 15–32. [[CrossRef](#)]
9. D’Urso, P.; Gastaldi, T. A least-squares approach to fuzzy linear regression analysis. *Comput. Stat. Data Anal.* **2000**, *34*, 427–440. [[CrossRef](#)]
10. D’Urso, P. Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data. *Comput. Stat. Data Anal.* **2003**, *42*, 47–72. [[CrossRef](#)]
11. González-Rodríguez, G.; Blanco, A.; Colubi, A.; Lubiano, M.A. Estimation of a simple linear regression model for fuzzy random variables. *Fuzzy Sets Syst.* **2009**, *160*, 357–370. [[CrossRef](#)]
12. Jung, H.; Yoon, J.H.; Choi, S.H. Fuzzy linear regression using rank transform method. *Fuzzy Sets Syst.* **2015**, *274*, 97–108. [[CrossRef](#)]
13. Grzegorzewski, P.; Mrowka, E. Linear Regression Analysis for Fuzzy Data. In Proceedings of the 10th IFSA World Congress-IFSA, Istanbul, Turkey, 29 June–2 July 2003; pp. 228–231.
14. Grzegorzewski, P.; Mrowka, E. Regression Analysis with Fuzzy Data. In *Soft Computing-Tools, Techniques and Applications*; Grzegorzewski P., Krawczak M., Eds.; Zadro Exit: Warszawa, Poland, 2004; pp. 65–76.
15. Hong, D.H.; Hwang, C. Extended fuzzy regression models using regularization method. *Inform. Sci.* **2004**, *164*, 31–46. [[CrossRef](#)]
16. Kim, H.K.; Yoon, J.H.; Li, Y. Asymptotic properties of least squares estimation with fuzzy observations. *Inform. Sci.* **2008**, *178*, 439–451. [[CrossRef](#)]
17. Kim, I. K.; Lee, W.-J.; Yoon, J.H.; Choi, S.H. Fuzzy Regression Model Using Trapezoidal Fuzzy Numbers for Re-auction Data. *Int. J. Fuzzy Log. Intell. Syst.* **2016**, *16*, 72–80. [[CrossRef](#)]
18. Lee, W.-J.; Jung, H.; Yoon, J.H.; Choi, S.H. The statistical inferences of fuzzy regression based on bootstrap techniques. *Soft Comput.* **2015**, *19*, 883–890. [[CrossRef](#)]

19. Ming, M.; Friedman, M.; Kandel, A. General fuzzy least squares. *Fuzzy Sets Syst.* **1997**, *88*, 107–118. [[CrossRef](#)]
20. Nasrabadi, M.M.; Nasrabadi, E. A mathematical-programming approach to fuzzy linear regression analysis. *Appl. Math. Comput.* **2004**, *155*, 873–881. [[CrossRef](#)]
21. Sakawa, M.; Yano, H. Multiobjective fuzzy linear regression analysis for fuzzy input-output data. *Fuzzy Sets Syst.* **1992**, *47*, 173–181. [[CrossRef](#)]
22. Yang, M.; Lin, T. Fuzzy least-squares linear regression analysis for fuzzy input-output data. *Fuzzy Sets Syst.* **2002**, *126*, 389–399. [[CrossRef](#)]
23. Yoon, J.H.; Choi, S.H. Componentwise fuzzy linear regression using least squares estimation. *J. Mult.-Valued Log. Soft Comp.* **2009**, *55*, 137–153.
24. Yoon, J.H.; Choi, S.H. Fuzzy rank linear regression model. In *AI 2009: Advances in Artificial Intelligence*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5866, pp. 617–626.
25. Diamond, P. Fuzzy Kriging. *Fuzzy Sets Syst.* **1989**, *46*, 315–332. [[CrossRef](#)]
26. Körner, R.; Näther, W. Linear regression with random fuzzy variables: Extended classical estimates, best linear estimates, least squares estimates. *Inform. Sci.* **1998**, *109*, 95–118. [[CrossRef](#)]
27. Näther, W. Linear statistical inference for random fuzzy data. *Statistics* **1997**, *29*, 221–240. [[CrossRef](#)]
28. Näther, W. On random fuzzy variables of second order and their application to linear statistical inference with fuzzy data. *Metrika* **2000**, *51*, 201–221. [[CrossRef](#)]
29. Näther, W. Regression with fuzzy random data *Comput. Stat. Data Anal.* **2006**, *51*, 235–252. [[CrossRef](#)]
30. Parchami, A.; Mashinchi, M. Fuzzy estimation for process capability indices. *Inform. Sci.* **2006**, *177*, 1452–1462. [[CrossRef](#)]
31. Chang, P.-T.; Lee, E.S. Fuzzy least absolute deviations regression and the conflicting trends in fuzzy parameters. *Comput. Math. Appl.* **1994**, *28*, 89–101. [[CrossRef](#)]
32. Yoon, J.H.; Choi, S.H. Separate fuzzy regression with crisp input and fuzzy output *J. Korean Data Inform. Sci.* **2007**, *18*, 301–314. (In Korean)
33. Yoon, J.H.; Choi, S.H. Fuzzy least squares estimation with new fuzzy operations. *Adv. Intell. Syst. Comput.* **2013**, *190*, 193–202.
34. Yoon J.H.; Choi S.H.; Grzegorzewski, P. On asymptotic properties of the multiple fuzzy least squares estimator. In *Soft Methods for Data Science*; Ferraro, M.B., Giordani, P., Vantaggi, B., Gagolewski, M., Ángeles Gil, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 525–532.
35. Yoon, J.H.; Jung, H.Y.; Lee, W.J.; Choi, S.H. Optimal properties of a fuzzy least estimator based on new operations. In *Proceedings of the SCIS&ISIS 2014*, Kitakyushu, Japan, 3–6 December 2014; pp. 36–41.
36. Yoon, J.H. Fuzzy regression analysis for generalized fuzzy data based on L_2 -distance. *Int. J. Fuzzy Syst.* **2020**, submitted.
37. Pedrycz, W. Why triangular membership functions? *Fuzzy Sets Syst.* **1994**, *64*, 21–30 [[CrossRef](#)]
38. Chang, P.-T.; Lee, E.S. Fuzzy linear regression with spreads unrestricted in sign. *Comput. Math. Appl.* **1994**, *28*, 61–70. [[CrossRef](#)]
39. Drygas, H. Weak and strong consistency of the least square estimates in regression models. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **1976**, *34*, 119–127. [[CrossRef](#)]
40. Serfling, R.J. *Approximation Theorems of Mathematical Statistics*; John Wiley & Sons, Inc.: New York, NY, USA; London, UK; Sydney, Australia; Toronto, ON, Canada, 1980.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).