Regression Analysis Model Based on Normal Fuzzy Numbers

Cui-Ling Gu, Wei Wang and Han-Yu Wei

Abstract Fuzzy regression analysis plays an important role in analyzing the correlation between the dependent and explanatory variables in the fuzzy system. This paper put forward the FLS (Fuzzy Least Squares) method for parameter estimating of the fuzzy linear regression model with input, output variables and regression coefficients that are normal fuzzy numbers. Our improved method proves the statistical properties, i.e., linearity and unbiasedness of the fuzzy least square estimators. Residuals, residual sum of squares and coefficient of determination are given to illustrate the fitting degree of the regression model. Finally, the method is validated in both rationality and validity by solving a practical parameter estimation problem.

Keywords Normal fuzzy numbers · Fuzzy regression analysis · Fuzzy least squares · Coefficient of determination

1 Introduction

The term regression was introduced by Francis Galton. Now, regression analysis is a fundamental analytic tool in many research fields. The method gives a crisp relationship between the dependent and explanatory variables with an estimated variance of measurement errors. Fuzzy regression [1] techniques provide a useful means to model the functional relationships between the dependent variable and independent variables in a fuzzy environment. After the introduction of fuzzy linear

C.-L. Gu · H.-Y. Wei
Mathematics and Statistical Institute, Zhoukou Normal University,
Zhoukou 466001, Henan, China
e-mail: 20121024@zknu.edu.cn

H.-Y. Wei
e-mail: weihanyu8207@163.com

W. Wang (✉)
School of Network Engineering, Zhoukou Normal University,
Zhoukou 466001, Henan, China
e-mail: wangwei@zknu.edu.cn

T.-H. Fan et al. (eds.), *Quantitative Logic and Soft Computing 2016*,
Advances in Intelligent Systems and Computing 510,
DOI 10.1007/978-3-319-46206-6_46

regression by Tanaka et al. [2], there has been a great deal of literatures on this topic [3–13]. Diamond [3] defined the distance between two fuzzy numbers and the estimated fuzzy regression parameters by minimizing the sum of the squares of the deviation. Chang [4] summarized three kinds of fuzzy regression methods from existing regression models: minimum fuzzy rule, the rule of least squares fitting and interval regression analysis method. For the purpose of integration of fuzziness and randomness, mixed regression model is put forward in [5]. Chang proposed the triangular fuzzy regression parameters least squares estimation by using the weighted fuzzy arithmetic and least-square fitting criterion. Sakawa and Yano [6] studied the fuzzy linear regression relation between the dependent variable and the fuzzy explanatory variable based on three given linear programming methods. In order to estimate the parameters of fuzzy linear regression model with input, output variables and regression coefficients are LR typed fuzzy numbers, Zhang [7] first represented the observed fuzzy data by using intervals, and then used the left, right point and the midpoint data sets of intervals to derive the corresponding regression coefficients of conventional linear regression models. Zhang [8] discussed the least squares estimation and the error estimate of the fuzzy regression analysis when the coefficient is described by trapezoidal fuzzy numbers depicting the fuzzy concept by using the gaussian membership function corresponding to human mind. To our knowledge, few researches are conducted on fuzzy regression analysis based on normal fuzzy numbers. Therefore, in this paper, we first calculate the least squares estimator of the fuzzy linear regression model, and then discuss statistical properties of the fuzzy least squares (FLS) estimator. Then, we give residuals, residual sum of squares and coefficient of determination and illustrate the fitting degree of the regression model. Last, we also verify the rationality and validity of the parameter estimation method by a numerical example (Fig. 1).

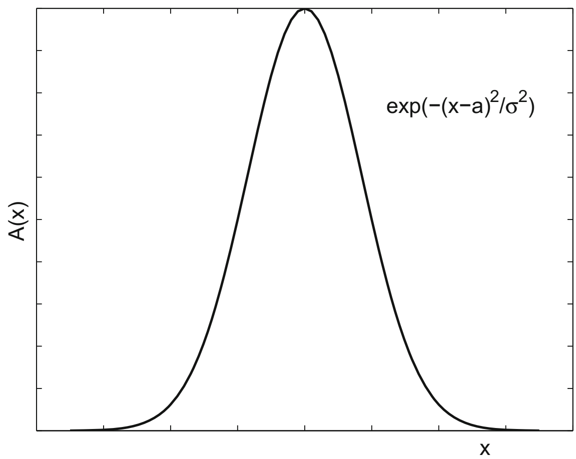

Fig. 1 The schematic diagram of fuzzy normal numbers

2 Preliminaries

Definition 1 ([14]) If fuzzy number $\tilde{A}$ has the following membership function

$$\tilde{A}(x) = exp\left\{-\frac{(x-a)^2}{\sigma^2}\right\}, x, a \in R, \sigma > 0$$

where R is a set of real numbers, then $\tilde{A}$ is called a normal fuzzy number determined by a and σ^2, and thus denoted by $\tilde{A} = (a, \sigma^2)$.

Let $\tilde{A} = (a, \sigma_a^2)$ and $\tilde{B} = (b, \sigma_b^2)$, then three operations of the normal fuzzy numbers are defined as follows: (1) $\tilde{A} + \tilde{B} = (a + b, \sigma_a^2 + \sigma_b^2)$; (2) $\lambda\tilde{A} = (\lambda a, \lambda\sigma_a^2)$; (3)$\frac{1}{\tilde{A}} = (\frac{1}{a}, \frac{1}{\sigma_a^2})$, where $a \neq 0$.

Definition 2 ([15]) The expectation of fuzzy number $\tilde{A}$ is

$$E(\tilde{A}) \triangleq \frac{\int_{-\infty}^{+\infty} x\tilde{A}(x)dx}{\int_{-\infty}^{+\infty} \tilde{A}(x)dx} \tag{1}$$

where $\int_{-\infty}^{+\infty} \tilde{A}(x)dx > 0$. The average of $\tilde{A}$ is denoted by the expectation $E(\tilde{A})$ of fuzzy number $\tilde{A}$. In particular, when $\tilde{A} = (a, \sigma_a^2)$, $E(\tilde{A}) = a$.

Definition 3 ([15]) The variance of fuzzy number $\tilde{A}$ is

$$D(\tilde{A}) \triangleq \frac{\int_{-\infty}^{+\infty} \tilde{A}(x)(x - E(\tilde{A}))^2 dx}{\int_{-\infty}^{+\infty} \tilde{A}(x)dx} \tag{2}$$

where $\int_{-\infty}^{+\infty} \tilde{A}(x)dx > 0$. The spread of $\tilde{A}$ is denoted by the variance $D(\tilde{A})$ of fuzzy number $\tilde{A}$. In particular, when $\tilde{A} = (a, \sigma_a^2)$, $D(\tilde{A}) = \frac{\sigma^2}{2}$.

Definition 4 ([15]) Multiplication between fuzzy numbers $\tilde{A}$ and $\tilde{B}$ is defined as:

$$\tilde{A} \otimes \tilde{B} \triangleq \int_{-\infty}^{+\infty} \tilde{A}(x)dx \int_{-\infty}^{+\infty} \tilde{B}(y)dy \tag{3}$$

when $\tilde{A} = \tilde{B}$, and $\tilde{A} \otimes \tilde{B} = \tilde{A} \otimes \tilde{A} = [\int_{\infty}^{\infty} \tilde{A}(x)dx]^2$, $\tilde{A} \otimes \tilde{A} = ||\tilde{A}||^2$ is called the module of $\tilde{A}$.

Let $\tilde{A}$ and $\tilde{B}$ denote the fuzzy numbers $\tilde{A} = (a, \sigma_a^2)$, and $\tilde{B} = (b, \sigma_b^2)$ respectively, then

$$\tilde{A} \otimes \tilde{B} \triangleq \int_{\infty}^{\infty} \tilde{A}(x)dx \int_{\infty}^{\infty} \tilde{B}(y)dy = \int_{\infty}^{\infty} e^{-\frac{(x-a)^2}{\sigma_a^2}} dx \int_{\infty}^{\infty} e^{-\frac{(y-b)^2}{\sigma_b^2}} dy = \pi\sigma_a\sigma_b,$$

Specifically, when $\tilde{A} = \tilde{B}, \tilde{A} \otimes \tilde{A} = ||\tilde{A}||^2$.

Definition 5 ([16]) Let $\tilde{A} = (a, \sigma_a^2), \tilde{B} = (b, \sigma_b^2)$, then the distance between $\tilde{A}$ and $\tilde{B}$ is defined as:

$$d^2(\tilde{A}, \tilde{B}) = (a - b)^2 + \frac{1}{2}(\sigma_a^2 - \sigma_b^2)^2 \tag{4}$$

3 The Least Squares Estimator of Fuzzy Linear Regression Model

The classical linear regression model is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \tag{5}$$

where Y is explained as variable and $X_1, X_2, \ldots, X_k$ are explanatory variables, β_0, $\beta_1, \ldots, \beta_k$ are regression coefficients. Let$\{(X_i, Y_i) : i = 1, 2, \ldots, n\}$ be a set of sample observations, ordinary least squares estimation is frequently based on the fact that the overall error between the estimated $\hat{Y}_i$ and the observations Y_i should be as small as possible. That is, the corresponding Q residual between the estimated $\hat{Y}_i$ and the observations Y_i should be as small as possible. Symbolically,

$$Q = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n}(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \cdots + \hat{\beta}_k X_{ki}))^2 \tag{6}$$

According to the principle of differential and integral calculus, Q will be the minimum value when the first order partial derivative of Q about $\beta_0, \beta_1, \ldots, \beta_k$ is equal to zero.

However, in many cases, the fuzzy relations in formula (5) must be considered. In general, there are the following three conditions [9]:
(a) $\tilde{Y}_i = \beta_0 + \beta_1 \tilde{X}_{1i} + \beta_2 \tilde{X}_{2i} + \cdots + \beta_k \tilde{X}_{ki}, \beta_0, \beta_1, \ldots, \beta_k \in R, \tilde{X}_1, \ldots, \tilde{X}_k, \tilde{Y}_i \in \widetilde{F}(R), i = 1, 2, \ldots, n$;
(b) $\tilde{Y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 X_{1i} + \tilde{\beta}_2 X_{2i} + \cdots + \tilde{\beta}_k X_{ki}, \tilde{\beta}_0, \tilde{\beta}_1, \ldots, \tilde{\beta}_k, \tilde{Y}_i \in \widetilde{F}(R), X_1, \ldots, X_k \in R, i = 1, 2, \ldots, n$;
(c) $\tilde{Y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{X}_{1i} + \tilde{\beta}_2 \tilde{X}_{2i} + \cdots \tilde{\beta}_k \tilde{X}_{ki}, \tilde{\beta}_0, \tilde{\beta}_1, \ldots, \tilde{\beta}_k, \tilde{X}_1, \ldots, \tilde{X}_k, \tilde{Y}_i \in \widetilde{F}(R), i = 1, 2, \ldots, n$.

In fact, (b) is the most common conditions. For (b), we focus on the fuzzy linear regression model in which dependent variables are the form of real numbers and explanatory variables and the regression coefficients are the form of normal fuzzy numbers.

Theorem 1 *Assume the fuzzy multiple linear regression model is as follows:*

$$\tilde{Y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 X_{1i} + \tilde{\beta}_2 X_{2i} + \cdots + \tilde{\beta}_k X_{ki}$$

then

$$\begin{aligned}\tilde{Y}_i = (a_i, \sigma_i^2) &= (a_{\tilde{\beta}_0}, \sigma^2_{\tilde{\beta}_0}) + (a_{\tilde{\beta}_1}, \sigma^2_{\tilde{\beta}_1}) X_{1i} + (a_{\tilde{\beta}_2}, \sigma^2_{\tilde{\beta}_2}) X_{2i} + \cdots + (a_{\tilde{\beta}_k}, \sigma^2_{\tilde{\beta}_k}) X_{ki} \\ &= (a_{\tilde{\beta}_0} + a_{\tilde{\beta}_1} X_{1i} + \cdots + a_{\tilde{\beta}_k} X_{ki}) + (\sigma^2_{\tilde{\beta}_0} + \sigma^2_{\tilde{\beta}_1} X_{1i}^2 + \cdots + \sigma^2_{\tilde{\beta}_k} X_{ki}^2)\end{aligned}$$

Let $\boldsymbol{a} = \boldsymbol{X}\boldsymbol{\psi}$, $\boldsymbol{b} = \boldsymbol{X_1}\boldsymbol{\zeta}$, *where*

$$\boldsymbol{\psi} = [a_{\tilde{\beta}_0}, a_{\tilde{\beta}_1}, \ldots, a_{\tilde{\beta}_k}]^T, \boldsymbol{\zeta} = [\sigma^2_{\tilde{\beta}_0}, \sigma^2_{\tilde{\beta}_1}, \ldots, \sigma^2_{\tilde{\beta}_k}]^T, \boldsymbol{b} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2]^T$$

$$\boldsymbol{a} = [a_1, a_2, \ldots, a_n]^T, \boldsymbol{A} = [\hat{a}_{\tilde{\beta}_0}, \hat{a}_{\tilde{\beta}_1}, \ldots, \hat{a}_{\tilde{\beta}_k}]^T, \boldsymbol{\sigma} = [\hat{\sigma}^2_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_1}, \ldots, \hat{\sigma}^2_{\tilde{\beta}_k}]^T$$

$$\boldsymbol{X} = \begin{pmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{pmatrix}$$

$$\boldsymbol{X_1} = \begin{pmatrix} 1 & X_{11}^2 & X_{21}^2 & \cdots & X_{k1}^2 \\ 1 & X_{12}^2 & X_{22}^2 & \cdots & X_{k2}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n}^2 & X_{2n}^2 & \cdots & X_{kn}^2 \end{pmatrix}$$

where $i = 1, 2, \ldots, n$; $\tilde{\beta}_0, \tilde{\beta}_1, \ldots, \tilde{\beta}_k, \tilde{Y} \in \tilde{F}(R)$; $X_1, X_2, \ldots, X_k \in R$. *Then, the FLS of* $\tilde{\beta}_0, \tilde{\beta}_1, \ldots, \tilde{\beta}_k$ *are defined as:*

$$\begin{cases} \boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{a} \\ \boldsymbol{\sigma} = (\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'\boldsymbol{b} \end{cases}$$

Proof Assuming that $\{(X_i, \tilde{Y}_i), i = 1, 2, \ldots, n\}$ are the set of known samples, and $\tilde{Y}_i = (a_i, \sigma_i^2)$, the sum Q of the squares of the dispersion between the estimated $\hat{\tilde{Y}}_i$ and the observations $\tilde{Y}_i$ should be minimized. That is,

$$\begin{aligned}\tilde{Q} &= \sum_{i=1}^{n}(\tilde{Y}_i - \hat{\tilde{Y}}_i)^2 = \sum_{i=1}^{n}\{(a_i, \sigma_i^2) - [(\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0}) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_{1i} + \cdots \\ &\quad + (\hat{a}_{\tilde{\beta}_k}, \hat{\sigma}^2_{\tilde{\beta}_k})X_{ki}]\} \\ &= \sum_{i=1}^{n}[(a_i, \sigma_i^2) - (\hat{a}_{\tilde{\beta}_0} + \hat{a}_{\tilde{\beta}_1}X_{1i} + \cdots + \hat{a}_{\tilde{\beta}_k}X_{ki}, \\ &\quad \hat{\sigma}^2_{\tilde{\beta}_0} + \hat{\sigma}^2_{\tilde{\beta}_1}X_{1i}^2 + \cdots + \hat{\sigma}^2_{\tilde{\beta}_1}X_{ki}^2)] \\ &= \sum_{i=1}^{n}[(a_i - \hat{a}_{\tilde{\beta}_0} - \hat{a}_{\tilde{\beta}_1}X_{1i} - \cdots - \hat{a}_{\tilde{\beta}_k}X_{ki})^2 \\ &\quad + \frac{1}{2}(\sigma_i^2 - \hat{\sigma}^2_{\tilde{\beta}_0} - \hat{\sigma}^2_{\tilde{\beta}_1}X_{1i}^2 - \cdots - \hat{\sigma}^2_{\tilde{\beta}_k}X_{ki}^2)^2]\end{aligned}$$

should be minimized. $\tilde{Q}$ will be the minimum value when the first order partial derivatives of Q about $\hat{\tilde{\beta}}_0, \hat{\tilde{\beta}}_1, \ldots, \hat{\tilde{\beta}}_k$ are equal to zero. In this case, fuzzy ordinary least squares estimator can be calculated.

$$\begin{cases}\dfrac{\partial\tilde{Q}}{\partial\hat{a}_{\tilde{\beta}_0}} = -2\sum_{i=1}^{n}(a_i - \hat{a}_{\tilde{\beta}_0} - \hat{a}_{\tilde{\beta}_1}X_{1i} - \cdots - \hat{a}_{\tilde{\beta}_k}X_{ki}) = 0 \\ \dfrac{\partial\tilde{Q}}{\partial\hat{a}_{\tilde{\beta}_1}} = -2\sum_{i=1}^{n}(a_i - \hat{a}_{\tilde{\beta}_0} - \hat{a}_{\tilde{\beta}_1}X_{1i} - \cdots - \hat{a}_{\tilde{\beta}_k}X_{ki})X_{1i} = 0 \\ \cdots\cdots \\ \dfrac{\partial\tilde{Q}}{\partial\hat{a}_{\tilde{\beta}_k}} = -2\sum_{i=1}^{n}(a_i - \hat{a}_{\tilde{\beta}_0} - \hat{a}_{\tilde{\beta}_1}X_{1i} - \cdots - \hat{a}_{\tilde{\beta}_k}X_{ki})X_{ki} = 0 \\ \dfrac{\partial\tilde{Q}}{\partial\hat{\sigma}^2_{\tilde{\beta}_0}} = -\sum_{i=1}^{n}(\sigma_i^2 - \hat{\sigma}^2_{\tilde{\beta}_0} - \hat{\sigma}^2_{\tilde{\beta}_1}X_{1i}^2 - \cdots - \hat{\sigma}^2_{\tilde{\beta}_k}X_{ki}^2) = 0 \\ \dfrac{\partial\tilde{Q}}{\partial\hat{\sigma}^2_{\tilde{\beta}_1}} = -\sum_{i=1}^{n}(\sigma_i^2 - \hat{\sigma}^2_{\tilde{\beta}_0} - \hat{\sigma}^2_{\tilde{\beta}_1}X_{1i}^2 - \cdots - \hat{\sigma}^2_{\tilde{\beta}_k}X_{ki}^2)X_{1i}^2 = 0 \\ \cdots\cdots \\ \dfrac{\partial\tilde{Q}}{\partial\hat{\sigma}^2_{\tilde{\beta}_k}} = -\sum_{i=1}^{n}(\sigma_i^2 - \hat{\sigma}^2_{\tilde{\beta}_0} - \hat{\sigma}^2_{\tilde{\beta}_1}X_{1i}^2 - \cdots - \hat{\sigma}^2_{\tilde{\beta}_k}X_{ki}^2)X_{ki}^2 = 0\end{cases}$$

Then, the above equations can be simplified as

$$\begin{cases} \sum_{i=1}^{n} a_i = n\hat{a}_{\tilde{\beta}_0} + \hat{a}_{\tilde{\beta}_1} \sum_{i=1}^{n} X_{1i} + \cdots + \hat{a}_{\tilde{\beta}_k} \sum_{i=1}^{n} X_{ki} \\ \sum_{i=1}^{n} a_i X_{1i} = \hat{a}_{\tilde{\beta}_0} \sum_{i=1}^{n} X_{1i} + \hat{a}_{\tilde{\beta}_1} \sum_{i=1}^{n} X_{1i}^2 + \cdots + \hat{a}_{\tilde{\beta}_k} \sum_{i=1}^{n} X_{ki} X_{1i} \\ \cdots\cdots \\ \sum_{i=1}^{n} a_i X_{ki} = \hat{a}_{\tilde{\beta}_0} \sum_{i=1}^{n} X_{ki} + \hat{a}_{\tilde{\beta}_1} \sum_{i=1}^{n} X_{ki} X_{ki} + \cdots + \hat{a}_{\tilde{\beta}_k} \sum_{i=1}^{n} X_{ki}^2 \\ -\sum_{i=1}^{n} \sigma_i^2 = n\hat{\sigma}^2_{\tilde{\beta}_0} + \hat{\sigma}^2_{\tilde{\beta}_1} \sum_{i=1}^{n} X_{1i}^2 + \cdots + \hat{\sigma}^2_{\tilde{\beta}_k} \sum_{i=1}^{n} X_{ki}^2 \\ -\sum_{i=1}^{n} \sigma_i^2 X_{1i}^2 = \hat{\sigma}^2_{\tilde{\beta}_0} \sum_{i=1}^{n} X_{1i}^2 + \hat{\sigma}^2_{\tilde{\beta}_1} \sum_{i=1}^{n} X_{1i}^4 + \cdots + \hat{\sigma}^2_{\tilde{\beta}_k} \sum_{i=1}^{n} X_{ki}^2 X_{1i}^2 \\ \cdots\cdots \\ -\sum_{i=1}^{n} \sigma_i^2 X_{ki}^2 = \hat{\sigma}^2_{\tilde{\beta}_0} \sum_{i=1}^{n} X_{ki}^2 + \hat{\sigma}^2_{\tilde{\beta}_1} \sum_{i=1}^{n} X_{ki}^2 X_{1i}^2 + \cdots + \hat{\sigma}^2_{\tilde{\beta}_k} \sum_{i=1}^{n} X_{ki}^4 \end{cases}$$

The matrix expression of the normal equations is as follows

$$\begin{cases} (\boldsymbol{X}'\boldsymbol{X})\boldsymbol{A} = \boldsymbol{X}'\boldsymbol{a} \\ (\boldsymbol{X_1}'\boldsymbol{X_1})\boldsymbol{\sigma} = \boldsymbol{X_1}'\boldsymbol{b} \end{cases}$$

And least squares estimator of parameters are as follows

$$\begin{cases} \boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{a} \\ \boldsymbol{\sigma} = (\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'\boldsymbol{b} \end{cases}$$

□

Corollary 1 *Assume that the fuzzy simple linear regression model is as follows*

$$\tilde{Y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 X_i, i = 1, 2, \ldots, n, \tilde{\beta}_0, \tilde{\beta}_1, \tilde{Y}_i \in \tilde{F}(R), X_i \in R, \tilde{Y}_i = (a_i, \sigma_i^2),$$

that is

$$\tilde{Y}_i = (a_i, \sigma_i^2) = (a_{\tilde{\beta}_0}, \sigma^2_{\tilde{\beta}_0}) + (a_{\tilde{\beta}_1}, \sigma^2_{\tilde{\beta}_1})X_i, i = 1, 2, \ldots, n, \tilde{\beta}_0, \tilde{\beta}_1 \tilde{Y}_i \in \tilde{F}(R), X_i \in R$$

where $\hat{\tilde{\beta}}_0 = (\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0})$ *and* $\hat{\tilde{\beta}}_1 = (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})$ *are respectively the FLS of* $\tilde{\beta}_0$ *and* $\tilde{\beta}_1$. *then*

$$\begin{cases} \hat{a}_{\tilde{\beta}_0} = \frac{\sum_{i=1}^{n} a_i \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} a_i X_i}{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2}, & \hat{\sigma}^2_{\tilde{\beta}_0} = \frac{\sum_{i=1}^{n} \sigma_i^2 \sum_{i=1}^{n} X_i^4 - \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} X_i^2 \sigma_i^2}{-n\sum_{i=1}^{n} X_i^4 + (\sum_{i=1}^{n} X_i^2)^2} \\ \hat{a}_{\tilde{\beta}_1} = \frac{n\sum_{i=1}^{n} a_i X_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} a_i}{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2}, & \hat{\sigma}^2_{\tilde{\beta}_1} = \frac{-n\sum_{i=1}^{n} X_i^2 \sigma_i^2 - \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} \sigma_i^2}{-n\sum_{i=1}^{n} X_i^4 + (\sum_{i=1}^{n} X_i^2)^2} \end{cases}$$

Proof Let $X_i, \tilde{Y}_i, i = 1, 2, \ldots, n$ be a set of sample observations and $\tilde{Y}_i = (a_i, \sigma_i^2)$, according to Theorem 1, is revised as follows:

$$\begin{aligned}
\tilde{Q} &= \sum_{i=1}^{n}(\tilde{Y}_i - \hat{\tilde{Y}}_i)^2 \\
&= \sum_{i=1}^{n}\{(a_i, \sigma_i^2) - [(\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0}) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i]\}^2 \\
&= \sum_{i=1}^{n}[(a_i, \sigma_i^2) - (\hat{a}_{\tilde{\beta}_0} + \hat{a}_{\tilde{\beta}_1}X_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \hat{\sigma}^2_{\tilde{\beta}_1}X_i^2)]^2 \\
&= \sum_{i=1}^{n}\left[(a - \hat{a}_{\tilde{\beta}_0} - \hat{a}_{\tilde{\beta}_1}X_i)^2 + \frac{1}{2}(\sigma_i^2 - \hat{\sigma}^2_{\tilde{\beta}_0} + \hat{\sigma}^2_{\tilde{\beta}_1}X_i^2)^2\right]
\end{aligned}$$

Obviously, $\tilde{Q}$ will be minimized when the first order partial derivatives of $\tilde{Q}$ about $\hat{\tilde{\beta}}_0$, $\hat{\tilde{\beta}}_1$ and are equal to zero. That is, we can solve the question by making the first order partial derivatives of $\tilde{Q}$ about $\hat{a}_{\tilde{\beta}_0}, \hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_1}$ respectively equal to zero.

$$\begin{cases}
\frac{\partial \tilde{Q}}{\partial \hat{a}_{\tilde{\beta}_0}} = -2\sum_{i=1}^{n}(a_i - \hat{a}_{\tilde{\beta}_0} - \hat{a}_{\tilde{\beta}_1}X_i) = 0 \\
\frac{\partial \tilde{Q}}{\partial \hat{a}_{\tilde{\beta}_1}} = -2\sum_{i=1}^{n}(a_i - \hat{a}_{\tilde{\beta}_0} - \hat{a}_{\tilde{\beta}_1}X_i)X_i = 0 \\
\frac{\partial \tilde{Q}}{\partial \hat{\sigma}^2_{\tilde{\beta}_0}} = -\sum_{i=1}^{n}(\sigma_i^2 - \hat{\sigma}^2_{\tilde{\beta}_0} + \hat{\sigma}^2_{\tilde{\beta}_1}X_i^2) = 0 \\
\frac{\partial \tilde{Q}}{\partial \hat{\sigma}^2_{\tilde{\beta}_1}} = -\sum_{i=1}^{n}(\sigma_i^2 - \hat{\sigma}^2_{\tilde{\beta}_0} + \hat{\sigma}^2_{\tilde{\beta}_1}X_i^2)X_i^2 = 0
\end{cases}$$

The above equations may be written as

$$\begin{cases}
n\hat{a}_{\tilde{\beta}_0} + \hat{a}_{\tilde{\beta}1}\sum_{i=1}^{n}X_i = \sum_{i=1}^{n}a_i \\
\hat{a}_{\tilde{\beta}_0}\sum_{i=1}^{n}X_i + \hat{a}_{\tilde{\beta}_1}\sum_{i=1}^{n}X_i^2 = \sum_{i=1}^{n}a_iX_i \\
n\hat{\sigma}^2_{\tilde{\beta}_0} - \hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}X_i^2 = -\sum_{i=1}^{n}\sigma_i^2 \\
\hat{\sigma}^2_{\tilde{\beta}_0}\sum_{i=1}^{n}X_i^2 - \hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}X_i^4 = -\sum_{i=1}^{n}\sigma_i^2X_i^2
\end{cases}$$

Then, in terms of Cramer's rule, we can obtain the linear fuzzy least squares estimator of the simple linear regression model by solving the above equations. □

4 The Statistical Properties of Fuzzy Least Squares Estimator

Theorem 2 *Fuzzy least squares estimator*

$$\begin{cases} \boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{a} \\ \boldsymbol{\sigma} = (\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'\boldsymbol{b} \end{cases}$$

is a linear estimator.

Proof Since

$$\begin{cases} \boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{a} = \boldsymbol{Ca} \\ \boldsymbol{\sigma} = (\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'\boldsymbol{b} = \boldsymbol{Db} \end{cases}$$

where $\boldsymbol{C} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}', \boldsymbol{D} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$, the parameter estimator is a linear combination of explanatory variables. □

In order to know statistic properties of the parameter estimator in simple fuzzy regression mode1, let $x_i = X_i - \bar{X}$, where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. When $\bar{\tilde{Y}} = (\bar{a}, \bar{\sigma}^2,)y_i = (\check{a}, \check{\sigma}^2) = \tilde{Y}_i - \bar{\tilde{Y}} = (a_i, \sigma_i^2) - (\bar{a}, \bar{\sigma}^2) = (a_i - \bar{a}, \sigma_i^2 - \bar{\sigma}^2)$, where $\bar{a} = E(\frac{1}{n}\sum_{i=1}^{n}\tilde{Y}_i)$ $= \frac{1}{n}\sum_{i=1}^{n} E(\tilde{Y}_i) = \frac{1}{n}\sum_{i=1}^{n} a_i, \bar{\sigma^2} = Var(\frac{1}{n}\sum_{i=1}^{n}\tilde{Y}_i) = \frac{1}{n^2}\sum_{i=1}^{n} Var(\tilde{Y}_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2$, then

$$\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \sum_{i=1}^{n} X_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right)^2$$

that is

$$n\sum_{i=1}^{n} x_i^2 = n\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2$$

so

$$\begin{cases} \hat{a}_{\tilde{\beta}_0} = \bar{a} - \hat{a}_{\tilde{\beta}_1}\bar{X} \\ \hat{a}_{\tilde{\beta}_1} = \frac{\sum_{i=1}^{n}\check{a}_1 x_i}{\sum_{i=1}^{n} x_i^2} \end{cases}$$

Corollary 2 *Expectations* $\hat{a}_{\tilde{\beta}_0}$ *and* $\hat{a}_{\tilde{\beta}_1}$ *of fuzzy least squares estimator* $\hat{\tilde{\beta}}_0 = (\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0})$ *and* $\hat{\tilde{\beta}}_1 = (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})$ *are linear estimators.*

Proof

$$\hat{a}_{\tilde{\beta}_1} = \frac{\sum_{i=1}^{n} \check{a}_i x_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} (a_i - \bar{a}) x_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} a_i x_i}{\sum_{i=1}^{n} x_i^2} - \frac{\sum_{i=1}^{n} \bar{a} x_i}{\sum_{i=1}^{n} x_i^2}$$
$$= \frac{\sum_{i=1}^{n} a_i x_i}{\sum_{i=1}^{n} x_i^2} - \frac{\bar{a} \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} = \sum_{i=1}^{n} k_i a_i$$

where $k_i = \frac{x_i}{\sum_{i=1}^{n} x_i^2}$, $\sum_{i=1}^{n} x_i = 0$;

$$\hat{a}_{\tilde{\beta}_0} = \bar{a} - \hat{a}_{\tilde{\beta}_1}\bar{X} = \frac{1}{n}\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} k_i a_i \bar{X} = \sum_{i=1}^{n} \left(\frac{1}{n} - \bar{X}k_i\right) a_i = \sum_{i=1}^{n} w_i a_i$$

where $w_i = \frac{1}{n} - \bar{X}k_i$. □

Theorem 3 *Fuzzy least squares estimator*

$$\begin{cases} \boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{a} \\ \boldsymbol{\sigma} = (\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'\boldsymbol{b} \end{cases}$$

are unbiased estimators.

Proof

$$E(\boldsymbol{A}) = E[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{a})] = E[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\psi}] = E(\boldsymbol{\psi}) = \boldsymbol{\psi}$$
$$E(\boldsymbol{\sigma}) = E[(\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'\boldsymbol{b}] = E[(\boldsymbol{X_1}'\boldsymbol{X_1})^{-1}\boldsymbol{X_1}'(\boldsymbol{X_1}\boldsymbol{\zeta})] = \boldsymbol{\zeta}$$

So fuzzy least squares estimators are unbiased. □

Corollary 3 *Expectations $a_{\tilde{\beta}_0}$ and $a_{\tilde{\beta}_1}$ of fuzzy least squares estimator $\hat{\tilde{\beta}}_0 = (\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0})$ and $\hat{\tilde{\beta}}_1 = (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})$ are unbiased estimators of the parameters $\tilde{\beta}_0$, $\tilde{\beta}_1$.*

Proof

$$\hat{a}_{\tilde{\beta}_1} = \sum_{i=1}^{n} k_i a_i = \sum_{i=1}^{n} k_i (a_{\tilde{\beta}_0} + a_{\tilde{\beta}_1} X_i) = a_{\tilde{\beta}_0} \sum_{i=1}^{n} k_i + a_{\tilde{\beta}_1} \sum_{i=1}^{n} k_i X_i$$

where $k_i = \frac{x_i}{\sum_{i=1}^{n} x_i^2}$, $\sum_{i=1}^{n} k_i = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} = 0$

$$\sum_{i=1}^{n} k_i X_i = \frac{\sum_{i=1}^{n} x_i X_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i (x_i + \bar{X})}{\sum_{i=1}^{n} x_i^2} = \frac{\bar{X} \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} + \frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2} = 1$$

so $E(\hat{a}_{\tilde{\beta}_1}) = a_{\tilde{\beta}_1}$;

$$\hat{a}_{\tilde{\beta}_0} = \sum_{i=1}^{n} w_i a_i = \sum_{i=1}^{n} [w_i (a_{\tilde{\beta}_0} + a_{\tilde{\beta}_1} X_i)] = a_{\tilde{\beta}_0} \sum_{i=1}^{n} w_i + a_{\tilde{\beta}_1} \sum_{i=1}^{n} w_i X_i$$

where, $w_i = \frac{1}{n} - \bar{X} k_i$,$E\left(\sum_{i=1}^{n} w_i\right) = E\left(\sum_{i=1}^{n} \left(\frac{1}{n} - \bar{X} k_i\right)\right) = 1$,

$$\sum_{i=1}^{n} w_i X_i = \sum_{i=1}^{n} \left(\frac{1}{n} - \bar{X} k_i\right) X_i = \frac{1}{n} \sum_{i=1}^{n} X_i - \bar{X} \sum_{i=1}^{n} k_i X_i = \bar{X} - \bar{X} = 0$$

so $E(\hat{a}_{\tilde{\beta}_0}) = a_{\tilde{\beta}_0}$. □

5 Assessment on Fuzzy Multiple Linear Regression Model

Regression analysis is a useful statistical method for analyzing quantitative relationships between two or more variables. It is important for the regression analysis to assess the performance of fitting regression model. That is to say, after estimating parameter of fuzzy liner regression model, how far is it from the parameter estimation to the true value? In fuzzy regression analysis, the simplest method evaluating the fuzzy regression model is to take the residual and the Coefficient of Determination as metrics. According to Classical Regression Mode [17], we can calculate the residual and the Coefficient of Determination about Fuzzy Regression Model by using fuzzy calculation rule which listed previously.

Theorem 4 give the module formula of residual $|\check{e}_i|$ and require that it is as small as possible. The fuzzy total sum of squares(FTSS) and the fuzzy explained sum of squares(FESS) are given in Theorem 5, and we obtain fuzzy coefficient of determination $\tilde{R}^2$ in Theorem 6, $\tilde{R}$ is bigger, and more better.

Theorem 4 *The residual produced by the fuzzy multiple linear regression model based on normal fuzzy numbers is defined as*

$$|\check{e}_i| = \sqrt{\pi}\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2} + \sqrt{\pi}\sigma_{\tilde{\beta}_1}|X_{1i}| + \sqrt{\pi}\sigma_{\tilde{\beta}_2}|X_{2i}| + \cdots + \sqrt{\pi}\sigma_{\tilde{\beta}_k}|X_{ki}|$$

Proof

$$\begin{aligned}
|\check{e}_i| &= |\hat{\tilde{Y}}_i - \tilde{Y}_i| \\
&= |(\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0}) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_{1i} + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_2})X_{2i} + \cdots + (\hat{a}_{\tilde{\beta}_k}, \hat{\sigma}^2_{\tilde{\beta}_k})X_{ki} - (a_i, \sigma_i^2)| \\
&= |(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_{1i} + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_2})X_{2i} + \cdots + (\hat{a}_{\tilde{\beta}_k}, \hat{\sigma}^2_{\tilde{\beta}_k})X_{ki}| \\
&= |(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2| + |(\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})||X_{1i}| + \cdots + |(\hat{a}_{\tilde{\beta}_k}, \hat{\sigma}^2_{\tilde{\beta}_k})||X_{ki}| \\
&= \sqrt{\pi}\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2} + \sqrt{\pi}\sigma_{\tilde{\beta}_1}|X_{1i}| + \sqrt{\pi}\sigma_{\tilde{\beta}_2}|X_{2i}| + \cdots + \sqrt{\pi}\sigma_{\tilde{\beta}_k}|X_{ki}| \qquad \square
\end{aligned}$$

Corollary 4 *The residual produced by the fuzzy simple linear regression model based on normal fuzzy numbers is expressed as*

$$\check{e}_i = \sqrt{\pi}\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2} + \sqrt{\pi}\hat{\sigma}_{\tilde{\beta}_1}|X_i|$$

Proof

$$\begin{aligned}
\check{e}_i &= |\hat{\tilde{Y}}_i - \tilde{Y}_i| = |(\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0}) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i - (a_i, \sigma_i^2)| \\
&= |(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i| \\
&= |(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)| + |(\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})||X_i| \\
&= \sqrt{\pi}\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2} + \sqrt{\pi}\hat{\sigma}_{\tilde{\beta}_1}|X_i| \qquad \square
\end{aligned}$$

Theorem 5 *The residual sum of squares produced by the fuzzy multiple linear regression model based on normal fuzzy numbers is defined as*

$$\begin{aligned}
FTSS = \pi \sum_{i=1}^{n}(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + \pi \sum_{i=1}^{n}\sum_{j=1}^{k}\hat{\sigma}^2_{\tilde{\beta}_j}X_{ji}^2 \\
+ 2\pi \sum_{i=1}^{n}\sum_{j=1}^{k}\hat{\sigma}_{\tilde{\beta}_j}X_{ji}\sqrt{(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)} + \pi \sum_{i=1}^{n}\sum_{j\neq r}^{k}\hat{\sigma}_{\tilde{\beta}_j}\hat{\sigma}_{\tilde{\beta}_r}X_{ri}X_{ji}
\end{aligned}$$

The explained sum of squares produced by the fuzzy multiple linear regression model based on normal fuzzy numbers is defined as

$$FESS = n\pi(\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + \pi \sum_{i=1}^{n} \sum_{j=1}^{k} \hat{\sigma}^2_{\tilde{\beta}_j} X^2_{ji} + 2\pi\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2} \sum_{i=1}^{n} \sum_{j=1}^{k} \hat{\sigma}_{\tilde{\beta}_j} X_{ji} + \pi \sum_{i=1}^{n} \sum_{j \neq r}^{k} \hat{\sigma}_{\tilde{\beta}_j} \hat{\sigma}_{\tilde{\beta}_r} X_{ri} X_{ji}$$

Proof

$$\begin{aligned}
FTSS &= \textstyle\sum_{i=1}^{n} (\hat{\tilde{Y}}_i - \tilde{Y}_i)^2 \\
&= \textstyle\sum_{i=1}^{n} [(\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0}) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1}) X_{1i} + \cdots + (\hat{a}_{\tilde{\beta}_k}, \hat{\sigma}^2_{\tilde{\beta}_k}) X_{ki} - (a_i, \sigma_i^2)]^2 \\
&= \textstyle\sum_{i=1}^{n} [(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1}) X_{1i} + \cdots + (\hat{a}_{\tilde{\beta}_k}, \hat{\sigma}^2_{\tilde{\beta}_k}) X_{ki}]^2 \\
&= \textstyle\sum_{i=1}^{n} \Big[(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)^2 + \sum_{j=1}^{k} (\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_1})^2 X^2_{ji} \\
&\quad + 2 \textstyle\sum_{j=1}^{k} (\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)(\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j}) X_{ji} \\
&\quad + \textstyle\sum_{r \neq j}^{k} (\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j})(\hat{a}_{\tilde{\beta}_r}, \hat{\sigma}^2_{\tilde{\beta}_r}) X_{ji} X_{ri} \Big] \\
&= \textstyle\sum_{i=1}^{n} \Big[(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)^2 + \sum_{i=1}^{n} \sum_{j=1}^{k} (\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_1})^2 X^2_{ji} \\
&\quad + 2 \textstyle\sum_{i=1}^{n} \sum_{j=1}^{k} (\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j})(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) X_{ji} \\
&\quad + \textstyle\sum_{i=1}^{n} \sum_{r \neq j}^{k} (\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j})(\hat{a}_{\tilde{\beta}_r}, \hat{\sigma}^2_{\tilde{\beta}_r}) X_{ji} X_{ri} \Big] \\
&= \pi \textstyle\sum_{i=1}^{n} (\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + \pi \sum_{i=1}^{n} \sum_{j=1}^{k} \hat{\sigma}^2_{\tilde{\beta}_j} X^2_{ji} \\
&\quad + 2 \textstyle\sum_{i=1}^{n} \sum_{j=1}^{k} [(\hat{a}_{\tilde{\beta}_0} X_{ji}, \hat{\sigma}^2_{\tilde{\beta}_0} X^2_{ji})(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)] \\
&\quad + \textstyle\sum_{i=1}^{n} \sum_{r \neq j}^{k} (\hat{a}_{\tilde{\beta}_j} X_{ji}, \hat{\sigma}^2_{\tilde{\beta}_j} X^2_{ji})(\hat{a}_{\tilde{\beta}_r} X_{ri}, \hat{\sigma}^2_{\tilde{\beta}_r} X^2_{ri}) \\
&= \pi \textstyle\sum_{i=1}^{n} (\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + \pi \sum_{i=1}^{n} \sum_{j=1}^{k} \hat{\sigma}^2_{\tilde{\beta}_j} X^2_{ji} \\
&\quad + 2\pi \textstyle\sum_{i=1}^{n} \sum_{j=1}^{k} \hat{\sigma}_{\tilde{\beta}_j} X_{ji} \sqrt{(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)} + \pi \sum_{i=1}^{n} \sum_{j \neq r}^{k} \hat{\sigma}_{\tilde{\beta}_j} \hat{\sigma}_{\tilde{\beta}_r} X_{ri} X_{ji}
\end{aligned}$$

$$
\begin{aligned}
FESS &= \textstyle\sum_{i=1}^{n}(\hat{\tilde{Y}}_i - \bar{\tilde{Y}})^2 \\
&= \textstyle\sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0}) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_{1i} + \cdots + (\hat{a}_{\tilde{\beta}_k}, \hat{\sigma}^2_{\tilde{\beta}_k})X_{ki} - (\bar{a}, \bar{\sigma}^2)]^2 \\
&= \textstyle\sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_{1i} + \cdots + (\hat{a}_{\tilde{\beta}_k}, \hat{\sigma}^2_{\tilde{\beta}_k})X_{ki}]^2 \\
&= \textstyle\sum_{i=1}^{n}\Big[(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2)^2 + \sum_{j=1}^{k}(\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j})^2 X_{ji}^2 \\
&\quad +2(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2)\textstyle\sum_{j=1}^{k}(\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j})X_{ji} \\
&\quad + \textstyle\sum_{r\neq j}^{k}(\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j})(\hat{a}_{\tilde{\beta}_r}, \hat{\sigma}^2_{\tilde{\beta}_r})X_{ji}X_{ri}\Big] \\
&= n(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2)^2 + \textstyle\sum_{i=1}^{n}\sum_{j=1}^{k}(\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})^2 X_{ji}^2 \\
&\quad +2(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2)\textstyle\sum_{i=1}^{n}\sum_{j=1}^{k}(\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j})X_{ji} \\
&\quad + \textstyle\sum_{i=1}^{n}\sum_{r\neq j}^{k}(\hat{a}_{\tilde{\beta}_j}, \hat{\sigma}^2_{\tilde{\beta}_j})(\hat{a}_{\tilde{\beta}_r}, \hat{\sigma}^2_{\tilde{\beta}_r})X_{ji}X_{ri} \\
&= n\pi(\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + \pi\textstyle\sum_{j=1}^{k}\sum_{i=1}^{n}\hat{\sigma}^2_{\tilde{\beta}_j}X_{ji}^2 \\
&\quad +2\sqrt{\pi}\sqrt{\hat{\sigma}_{\tilde{\beta}_0} + \bar{\sigma}^2}\textstyle\sum_{i=1}^{n}\sum_{j=1}^{k}\sqrt{\pi}\hat{\sigma}_{\tilde{\beta}_j}X_{ji} \\
&\quad + \textstyle\sum_{i=1}^{n}\sum_{r\neq j}^{k}(\hat{a}_{\tilde{\beta}_j}X_{ji}, \hat{\sigma}^2_{\tilde{\beta}_j}X_{ji}^2)(\hat{a}_{\tilde{\beta}_r}X_{ri}, \hat{\sigma}^2_{\tilde{\beta}_r}X_{ri}^2) \\
&= n\pi(\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + \pi\textstyle\sum_{i=1}^{n}\sum_{j=1}^{k}\hat{\sigma}^2_{\tilde{\beta}_j}X_{ji}^2 \\
&\quad +2\pi\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2}\textstyle\sum_{i=1}^{n}\sum_{j=1}^{k}\hat{\sigma}_{\tilde{\beta}_j}X_{ji} + \pi\sum_{i=1}^{n}\sum_{j\neq r}^{k}\hat{\sigma}_{\tilde{\beta}_j}\hat{\sigma}_{\tilde{\beta}_r}X_{ri}X_{ji}
\end{aligned}
$$

□

Corollary 5 *The residual produced by the fuzzy simple linear regression model based on normal fuzzy numbers is expressed as*

$$
FTSS = \pi\sum_{i=1}^{n}(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + 2\pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}\sqrt{(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)X_i^2} + \pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}X_i^2
$$

The explained sum of squares produced by fuzzy simple linear regression model based on normal fuzzy numbers is defined as

$$
FESS = n\pi(\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + 2\sqrt{\pi}\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2}\sum_{i=1}^{n}X_i + \pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}X_i^2
$$

Proof

$$
\begin{aligned}
FTSS &= \sum_{i=1}^{n}(\hat{\tilde{Y}}_i - \tilde{Y}_i)^2 \\
&= \sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0}) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i - (a_i, \sigma_i^2)]^2 \\
&= \sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i]^2 \\
&= \sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)^2 + 2(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)(\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i \\
&\quad + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})^2 X_i^2] \\
&= \sum_{i=1}^{n}(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + 2(\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})\sum_{i=1}^{n}(\hat{a}_{\tilde{\beta}_0} - a_i, \hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)X_i \\
&\quad + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})^2 \sum_{i=1}^{n} X_i^2 \\
&= \pi\sum_{i=1}^{n}(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + 2\sqrt{\pi}\hat{\sigma}_{\tilde{\beta}_1}\sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0} - a_i)X_i, (\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)X_i^2] \\
&\quad + \pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n} X_i^2 \\
&= \pi\sum_{i=1}^{n}(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2) + 2\pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}\sqrt{(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma_i^2)X_i^2} + \pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n} X_i^2
\end{aligned}
$$

$$
\begin{aligned}
FESS &= \sum_{i=1}^{n}(\hat{\tilde{Y}}_i - \bar{\tilde{Y}}_i)^2 \\
&= \sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0}, \hat{\sigma}^2_{\tilde{\beta}_0}) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i - (\bar{a}, \bar{\sigma}^2)]^2 \\
&= \sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i]^2 \\
&= \sum_{i=1}^{n}[(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2)^2 + 2(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} \\
&\quad + \bar{\sigma}^2)(\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})X_i + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})^2 X_i^2] \\
&= n(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2)^2 \\
&\quad + 2(\hat{a}_{\tilde{\beta}_0} - \bar{a}, \hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2)(\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})\sum_{i=1}^{n} X_i + (\hat{a}_{\tilde{\beta}_1}, \hat{\sigma}^2_{\tilde{\beta}_1})^2 \sum_{i=1}^{n} X_i^2 \\
&= n\pi(\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + 2\sqrt{\pi}\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2}\sum_{i=1}^{n} X_i + \pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n} X_i^2
\end{aligned}
$$

□

The greater the regression sum of squares, the smaller the sum of squared residuals, and the better the fitting between regression line and the sample points.

Theorem 6 *The coefficient of determination of the fuzzy multiple linear regression model based on normal fuzzy numbers is defined as*

$$\tilde{R}^2 = \frac{FESS}{FTSS} = \frac{n\pi(\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + \pi\sum_{i=1}^{n}\sum_{j=1}^{k}\hat{\sigma}^2_{\tilde{\beta}_j}X^2_{ji} + 2\pi\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2}\sum_{i=1}^{n}\sum_{j=1}^{k}\hat{\sigma}_{\tilde{\beta}_j}X_{ji} + \pi\sum_{i=1}^{n}\sum_{j\neq r}^{k}\hat{\sigma}_{\tilde{\beta}_j}\hat{\sigma}_{\tilde{\beta}_r}X_{ri}X_{ji}}{\pi\sum_{i=1}^{n}(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma^2_i) + \pi\sum_{i=1}^{n}\sum_{j=1}^{k}\hat{\sigma}^2_{\tilde{\beta}_j}X^2_{ji} + 2\pi\sum_{i=1}^{n}\sum_{j=1}^{k}\hat{\sigma}_{\tilde{\beta}_j}X_{ji}\sqrt{(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma^2_i)} + \pi\sum_{i=1}^{n}\sum_{j\neq r}^{k}\hat{\sigma}_{\tilde{\beta}_j}\hat{\sigma}_{\tilde{\beta}_r}X_{ri}X_{ji}}$$

Proof It is easy to prove Theorem 6 by using Theorem 5. □

Corollary 6 *The coefficient of the determination of fuzzy simple linear regression model based on normal fuzzy numbers is expressed as*

$$\tilde{R}^2 = \frac{FESS}{FTSS} = \frac{n\pi(\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2) + 2\sqrt{\pi}\sqrt{\hat{\sigma}^2_{\tilde{\beta}_0} + \bar{\sigma}^2}\sum_{i=1}^{n}X_i + \pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}X^2_i}{\pi\sum_{i=1}^{n}(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma^2_i) + 2\pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}\sqrt{(\hat{\sigma}^2_{\tilde{\beta}_0} + \sigma^2_i)X^2_i} + \pi\hat{\sigma}^2_{\tilde{\beta}_1}\sum_{i=1}^{n}X^2_i}$$

6 Numerical Example

Assume that the fuzzy linear regression model is as follows:

$$\tilde{Y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 X_{1i} + \tilde{\beta}_2 X_{2i}$$

where, $\tilde{Y}$ is the dependent variable, X_1 and X_2 the explanatory variables, and $(X_{1i}, X_{2i}, \tilde{Y}_i)$, $i = 1, 2, \ldots, n$, $X_1, X_2 \in R$, $\tilde{Y} \in \tilde{F}(R)$. Now, our goal is to solve the fuzzy regression and evaluate the model with the observed data shown in Table 1.

Then, the fuzzy regression mode can be obtained by our proposed method.

$$\tilde{Y}_i = (20.5371, 0.0542^2) + (41.5827, 0.0087^2)X_{1i} + (14.5884, 0.0035^2)X_{2i}$$

Residual series of the regression model are shown in Table 2. According to the formulas in Theorem 5, we can calculate the evaluation indexes of the fuzzy model i.e., $FTSS = 7.2531$, $FESS = 6.9836$, and $\tilde{R}^2 = 0.9628$. Clearly, the uncertainty of the practical problem is better considered by the fuzzy linear regression analysis. Using fuzzy numbers to represent the observation data makes it more effective to

Table 1 The observed data

Order	X_1	X_2	$\tilde{Y}$	Order	X_1	X_2	$\tilde{Y}$
1	0.16	0.86	$(40.0, 0.31^2)$	7	0.28	1.15	$(48.6, 0.18^2)$
2	0.18	0.89	$(41.0, 0.22^2)$	8	0.29	1.18	$(49.4, 0.19^2)$
3	0.23	0.94	$(42.0, 0.25^2)$	9	0.32	1.25	$(50.8, 0.23^2)$
4	0.24	0.96	$(43.0, 0.16^2)$	10	0.35	1.29	$(54.3, 0.24^2)$
5	0.22	0.98	$(46.5, 0.17^2)$	11	0.39	1.33	$(57.0, 0.25^2)$
6	0.26	0.99	$(47.2, 0.20^2)$	12	0.45	1.37	$(59.2, 0.21^2)$

Table 2 Residual series of the regression model

Order	Fuzzy residual	Order	Fuzzy residual	Order	Fuzzy residual
1	0.1833	5	0.0659	9	0.1160
2	0.0993	6	0.0862	10	0.1207
3	0.1253	7	0.0741	11	0.1302
4	0.0602	8	0.0810	12	0.0988

resolve the problem. In this example, the residual sequence of the regression model and the coefficient of determination help to understand how well the regression model can fit the sample points. The coefficient of determination 96.28 % implies that the change 96.28 % of the explained variable can be explained by the change of explanatory variables.

7 Conclusions

The paper proposes an improved FLS method for parameter estimating of the fuzzy linear regression model when the explanatory variables are precise and the explained variables and regression parameters are normal fuzzy numbers. Specifically, the paper figures out the fuzzy least squares estimation of multivariate linear regression analysis and gets some statistical properties, i.e., linearity and unbiasedness, of the fuzzy least square estimators. Finally, it illustrates the feasibility and effectiveness of the proposed method by the numerical example.

Acknowledgments This work is supported in part by the Science and Technology Department of Henan Province (Grant No. 152300410230) and the Key Scientific Research Projects of Henan Province (Grant No. 17A110040).

References

1. Zhang, A.W.: Statistical analysis of fuzzy linear regression model based on centroid method. J. Fuzzy Syst. Math. **5**, 172–177 (2012)
2. Tanaka, H., Uejima, S., Asai, K.: Linear regression analysis with fuzzy model. J. Syst. Man Cyber. **12**, 903–907 (1982)
3. Diamond, P.: Fuzzy least squares. J. Inf. Sci. **46**, 141–157 (1988)
4. Chang, Y.H., Ayyub, B.M.: Fuzzy regression methods–a comparative assessment. J. Fuzzy Sets Syst. **119**, 187–203 (2011)
5. Chang, Y.H.: Hybrid fuzzy least-squares regression analysis and its reliability measures. J. Fuzzy Sets Syst. **119**, 225–246 (2011)
6. Sakawa, M., Yano, H.: Multiobjective fuzzy liear regression analysis for fuzzy input-out data. J. Fuzzy Set Syst. **157**, 137–181 (1992)
7. Zhang, A.W.: Parameter estimation of the fuzzy regressive model with lr typed fuzzy coefficients. J. Fuzzy Syst. Math. **27**, 140–147 (2013)
8. Zhang, A.W.: A least-squares approach to fuzzy regression analysis with trapezoidal fuzzy number. J. Math. Pract. Theory **42**, 235–244 (2012)
9. Parvathi, R., Malathi, C., Akram, M.: Intuitionistic fuzzy linear regression analysis. J. Fuzzy Optim. Decis. Mak. **2**, 215–229 (2013)
10. Xu, R.N., Li, C.L.: Multidimensional least squares fitting with a fuzzy model. J. Fuzzy Sets Syst. **119**, 215–223 (2001)
11. Cope, R., D'Urso, P., Giordani, P., et al.: Least squares estimation of a linear regression model with LR fuzzy response. J. Comput. Stat. Data Anal. **51**, 267–286 (2006)
12. Bisserier, A., Boukezzoula, B., Galichet, S.: Linear Fuzzy Regression Using Trapezoidal Fuzzy Intervals. I PMU, Malaga (2008)
13. Liang, Y., Wei, L.L.: Fuzzy linear least-squares regression with LR-type fuzzy coefficients. J. Fuzzy Syst. Math. **3**, 112–117 (2007)
14. Peng, Z.Y., Sun, Y.Y.: The Fuzzy Mathematics and Its Application. Wuhan University Press, Wuhan (2007)
15. Li, A.G., Zhang, Z.H., Meng, Y., et al.: Fuzzy Mathematics and Its Application. Metallurgical Industry Press, Beijing (2005)
16. Xu, R.N.: Research on time series prediction problem with the normal fuzzy number. J. J. Guangzhou Univ. **12**, 82–89 (1988)
17. Li, Z.N.: Econmetrics. Higher Education Press, Beijing (2012)