

Statistical Analysis of Fuzzy Linear Regression Model Based on Centroid Method

Aiwu Zhang

School of Mathematical Science, Yancheng Teachers University, Yancheng, China

Email: zaw_017@163.com

Received 9 September 2015; accepted 15 April 2016; published 18 April 2016

Copyright © 2016 by author and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper transforms fuzzy number into clear number using the centroid method, thus we can research the traditional linear regression model which is transformed from the fuzzy linear regression model. The model's input and output are fuzzy numbers, and the regression coefficients are clear numbers. This paper considers the parameter estimation and impact analysis based on data deletion. Through the study of example and comparison with other models, it can be concluded that the model in this paper is applied easily and better.

Keywords

Centroid Method, Fuzzy Linear Regression Model, Parameter Estimation, Data Deletion Model, Cook Distance

1. Introduction

Regression analysis is an important and comprehensive approach to analyze relationship between dependent variable and other one or more independent variables; it has a very wide range of applications in engineering sciences, social sciences, economic and financial fields. Traditional regression analysis methods often require that both independent variable and dependent variable are clear data. However, practical problem is often not clear data but fuzzy data. For example, the amount of observations described in language, such as something large, something heavy, or approximately equal to 3, etc. Because of some ambiguous indicators, analyzing these issues only by traditional regression can not get satisfactory and completely results.

By means of Zadeh's [1] fuzzy set theory, researchers established a different fuzzy regression model and solved its solutions. After Tanaka *et al.* [2], Diamond [3] estimated regression coefficients using least squares method (FLS), which is similar to traditional LS estimate. Savic and Pedrycz [4] established a two-step model of fuzzy regression analysis by combining FLS with linear programming. Recently, Chang [5] compared fuzzy re-

gression methods, summed up three methods of fuzzy regression: minimum fuzzy criteria, least squares fitting criteria and interval regression analysis method. The main difference between fuzzy regression and conventional regression is that the residual in fuzzy regression is a fuzzy variable, but a random variable in traditional regression. The fuzzy regression model discussed here can be divided into several cases, such as the regression coefficient being expressed in fuzzy numbers; or part of variables being ambiguous; or input and output variables being ambiguous [6]-[10].

This paper starts from the fuzzy input and output variables, transforms them into clear data by centroid method [11], and then, the problem of fuzzy linear regression analysis (regression coefficient is clear number) can be transformed into traditional linear regression analysis. Thus, the problems of fuzzy linear regression analysis can be addressed by the estimation and statistical diagnosis method of traditional linear regression model.

2. Fuzzy Regression Model and Parameter Estimation

Assume that $(\tilde{x}_{ij}, \tilde{y}_i) (i = 1, 2, \dots, n, j = 1, 2, \dots, p)$ is a observational data set of fuzzy input and fuzzy output, and $\tilde{x}_{ij}, \tilde{y}_i \in \tilde{R}$ (\tilde{R} is the fuzzy number set of the real number set R). Then fuzzy linear regression model can be expressed as

$$\tilde{y}_i = \beta_0 + \beta_1 \tilde{x}_{i1} + \beta_2 \tilde{x}_{i2} + \dots + \beta_p \tilde{x}_{ip} + \tilde{\varepsilon}_i \tag{1}$$

where $\beta_0, \beta_1, \dots, \beta_p \in R, \tilde{\varepsilon}_i \in \tilde{R}$, $\tilde{\varepsilon}_i$ is the i -th observation error.

Assume

$$\begin{aligned} \tilde{Y} &= (\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_n)^T, \quad \beta = (\beta_0, \beta_1, \dots, \beta_p)^T, \\ \tilde{E} &= (\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \dots, \tilde{\varepsilon}_n)^T, \quad \tilde{X} = \begin{pmatrix} 1 & \tilde{x}_{11} & \dots & \tilde{x}_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \tilde{x}_{n1} & \dots & \tilde{x}_{np} \end{pmatrix} \end{aligned}$$

Then the fuzzy linear regression model (1) can be expressed in the form of matrix as

$$\tilde{Y} = \tilde{X} \beta + \tilde{E}, \tag{2}$$

where the membership functions of $\tilde{x}_{ij}, \tilde{y}_i, \tilde{\varepsilon}_i$ are $\mu_{\tilde{x}_{ij}}, \mu_{\tilde{y}_i}, \mu_{\tilde{\varepsilon}_i}, i = 1, 2, \dots, n, j = 1, 2, \dots, p$.

For convenience of discussion, we assume that all observations are triangular fuzzy numbers

$$\tilde{x}_{ij} = (x_{ijl}, x_{ijm}, x_{iju}), \tilde{y}_i = (y_{il}, y_{im}, y_{iu}), \tilde{\varepsilon}_i = (\varepsilon_{il}, 0, \varepsilon_{iu})$$

where the $x_{ijl}, x_{ijm}, x_{iju}, y_{il}, y_{im}, y_{iu}, \varepsilon_{il}, \varepsilon_{iu}$ are all real numbers. And their membership functions are

$$\begin{aligned} \mu_{\tilde{x}_{ij}}(x) &= \begin{cases} \frac{x - x_{ijl}}{x_{ijm} - x_{ijl}}, & x_{ijl} \leq x \leq x_{ijm}, \\ \frac{x_{iju} - x}{x_{iju} - x_{ijm}}, & x_{ijm} \leq x \leq x_{iju}, \\ 0, & \text{otherwise,} \end{cases} \\ \mu_{\tilde{y}_i}(y) &= \begin{cases} \frac{y - y_{il}}{y_{im} - y_{il}}, & y_{il} \leq y \leq y_{im}, \\ \frac{y_{iu} - y}{y_{iu} - y_{im}}, & y_{im} \leq y \leq y_{iu}, \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

In order to obtain the estimator of regression coefficient β , a natural idea is to try to turn fuzzy observation data $\tilde{x}_{ij}, \tilde{y}_i$ into crisp data, and then use traditional least squares method to calculate the estimator of β [12]. There are many ways to transform fuzzy data into crisp data, one of the most common methods is the centroid method [7].

The fuzzy data $\tilde{x}_{ij}, \tilde{y}_i$ is transformed into crisp data x_{ijc}, y_{ic} (x_{ijc}, y_{ic} are usually called the centroid of $\tilde{x}_{ij}, \tilde{y}_i$) with the formula

$$\begin{aligned} x_{ijc} &= \frac{\int_{-\infty}^{+\infty} x \mu_{\tilde{x}_{ij}}(x) dx}{\int_{-\infty}^{+\infty} \mu_{\tilde{x}_{ij}}(x) dx} = \frac{1}{3}(x_{ijl} + x_{ijm} + x_{iju}), \\ y_{ic} &= \frac{\int_{-\infty}^{+\infty} y \mu_{\tilde{y}_i}(y) dy}{\int_{-\infty}^{+\infty} \mu_{\tilde{y}_i}(y) dy} = \frac{1}{3}(y_{il} + y_{im} + y_{iu}). \end{aligned} \tag{3}$$

Obviously, when the observation data is a symmetric triangular fuzzy data, the centroid of fuzzy observation data is the symmetric center of the symmetric triangular fuzzy data.

Lemma 1 [13] The traditional linear regression model

$$Y = X\beta + \varepsilon$$

where $Y = (y_1, y_2, \dots, y_n)^T$, $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$, $\varepsilon \sim N(0, \sigma^2 I)$,

$$X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix}, \beta = (\beta_0, \beta_1, \dots, \beta_p)^T.$$

Then the least squares estimator of β is

$$\hat{\beta} = (X^T X)^{-1} X^T Y.$$

According to Lemma 1, it is easy to get the estimator for the parameter in model (1) or (2).

Theorem 2.1 Assume that the fuzzy linear regression model is

$$\tilde{Y} = \tilde{X}\beta + \tilde{E}$$

where $(\tilde{x}_{ij}, \tilde{y}_i) (i = 1, 2, \dots, n, j = 1, 2, \dots, p)$ are observed triangular fuzzy data and $\tilde{x}_{ij} = (x_{ijl}, x_{ijm}, x_{iju})$, $\tilde{y}_i = (y_{il}, y_{im}, y_{iu})$. Then the estimator for β is

$$\hat{\beta} = (X_c^T X_c)^{-1} X_c^T Y_c \tag{4}$$

where $X_c = \begin{pmatrix} 1 & x_{11c} & \dots & x_{1pc} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1c} & \dots & x_{npc} \end{pmatrix}$, $Y_c = (y_{1c}, y_{2c}, \dots, y_{nc})^T$, x_{ijc}, y_{ic} can be calculated by (3).

When fuzzy data is reduced to crisp data, the least squares estimation of β is a conventional least squares estimation.

Specifically, when the fuzzy linear regression model is $\tilde{y}_i = \beta_0 + \beta_1 \tilde{x}_i + \tilde{\varepsilon}_i$, we have

Corrary Let $(\tilde{x}_i, \tilde{y}_i) (i = 1, 2, \dots, n)$ is a set of triangular fuzzy data of the fuzzy linear regression model $\tilde{y}_i = \beta_0 + \beta_1 \tilde{x}_i + \tilde{\varepsilon}_i$, it has

$$\hat{\beta}_0 = \bar{y}_c - \hat{\beta}_1 \bar{x}_c, \hat{\beta}_1 = \frac{\sum_{i=1}^n x_{ic} y_{ic} - n \bar{x}_c \bar{y}_c}{\sum_{i=1}^n x_{ic}^2 - n \bar{x}_c^2}. \tag{5}$$

where $\tilde{x}_i = (x_{il}, x_{im}, x_{iu})$, $\tilde{y}_i = (y_{il}, y_{im}, y_{iu})$, $x_{ic} = \frac{1}{3}(x_{il} + x_{im} + x_{iu})$, $y_{ic} = \frac{1}{3}(y_{il} + y_{im} + y_{iu})$,

where $\bar{x}_c = \frac{1}{n} \sum_{i=1}^n x_{ic}$, $\bar{y}_c = \frac{1}{n} \sum_{i=1}^n y_{ic}$.

If the fuzzy observation data $(\tilde{x}_i, \tilde{y}_i)$ is not triangular fuzzy data, the centroid method can also apply.

3. The Evaluation Performance of Fuzzy Linear Regression Model

In order to evaluate the performance of fuzzy regression model, Kim and Bishu [11] introduced an absolute difference of the observed fuzzy dependent variable and estimated one as

$$E_i = \int_{S_{\hat{y}_i} \cup S_{\tilde{y}_i}} \left| \mu_{\hat{y}_i}(y) - \mu_{\tilde{y}_i}(y) \right| dy, \tag{6}$$

where $S_{\hat{y}_i}$ and $S_{\tilde{y}_i}$ are the support of $\mu_{\hat{y}_i}$ and $\mu_{\tilde{y}_i}$, respectively.

Essentially, E_i is estimated error term, the smaller value of E_i , the better fit of fuzzy linear regression model. Nasrabadi and Nasrabadi [12] showed the general calculation steps. Kao and Chyu [13] showed with the fuzzy linear regression model $\hat{y}_i = \beta_0 + \beta_1 \tilde{x}_{ij} + \tilde{\varepsilon}$, $\tilde{\varepsilon} = (-l, 0, r)$, the formula of E_i is

$$E_i = 0.5(y_{iu} - y_{il}) + 0.5[r + l + \hat{\beta}_1(x_{iju} - x_{ijl})] - 2(0.5)[y_{iu} - (\hat{\beta}_0 + \hat{\beta}_1 x_{ijl} - l)] \alpha_i,$$

When putting $\hat{\beta}_0, \hat{\beta}_1$ in (5) into the above formula, it has

$$E_i = 0.5(y_{iu} - y_{il}) + 0.5 \left[r + l + \frac{\sum_{i=1}^n x_{ic} y_{ic} - n \bar{x}_c \bar{y}_c}{\sum_{i=1}^n x_{ic}^2 - n \bar{x}_c^2} (x_{iju} - x_{ijl}) \right] - 2(0.5) \left[y_{iu} - \left(\bar{y}_c - \frac{\sum_{i=1}^n x_{ic} y_{ic} - n \bar{x}_c \bar{y}_c}{\sum_{i=1}^n x_{ic}^2 - n \bar{x}_c^2} \bar{x}_c + \frac{\sum_{i=1}^n x_{ic} y_{ic} - n \bar{x}_c \bar{y}_c}{\sum_{i=1}^n x_{ic}^2 - n \bar{x}_c^2} x_{ijl} - l \right) \right] \alpha_i, \tag{7}$$

where $\alpha_i = \frac{y_{iu} - \hat{y}_{il}}{y_{iu} - y_{im} + \hat{y}_{im} - \hat{y}_{il}}$. The value of l, r can be determined by using the following method

$$\begin{aligned} & \min \sum_{i=1}^n E_i, \\ & \text{s.t. } E_i = \int_{S_{\hat{y}_i} \cup S_{\tilde{y}_i}} \left| \mu_{\hat{y}_i}(y) - \mu_{\tilde{y}_i}(y) \right| dy, \\ & \hat{y}_i = \beta_0 + \beta_1 \tilde{x}_{ij} + \tilde{\varepsilon}, \\ & \tilde{\varepsilon} = (-l, 0, r), \\ & (\hat{y}_{im} - \hat{y}_{il}) \geq l_{\min}, (\hat{y}_{iu} - \hat{y}_{im}) \geq r_{\min}. \end{aligned}$$

4. Parameter Estimation and Impact Analysis of the Data Deleted Fuzzy Linear Regression Model

4.1. The Fuzzy Linear Regression Model Based on Data Deletion

For the fuzzy linear regression model (1) or (2), in order to evaluate the role and impact of the i -th data point $(\tilde{x}_{ij}, \tilde{y}_i)$ in the regression analysis, we can compare the inference results of before and after deleting the i -th data point $(\tilde{x}_{ij}, \tilde{y}_i)$. And we can test this point whether it is an outlier point or not. The fuzzy linear regression model with the i -th point deleted is called a case deletion fuzzy linear regression model (FCDM), and its component form and matrix form are respectively

$$\tilde{y}_k = \beta_0 + \beta_1 \tilde{x}_{k1} + \beta_2 \tilde{x}_{k2} + \dots + \beta_p \tilde{x}_{kp} + \tilde{\varepsilon}_k, \quad k = 1, 2, \dots, n, \text{ but } k \neq i \tag{8}$$

$$\tilde{Y}(i) = \tilde{X}(i) \beta(i) + \tilde{\varepsilon}(i) \tag{9}$$

where $\tilde{Y}(i), \tilde{X}(i), \tilde{\varepsilon}(i)$ is the vector or matrix after deleting the i -th data of $\tilde{Y}, \tilde{X}, \tilde{\varepsilon}$ respectively and $\hat{\beta}(i)$ denotes the least squares estimator of $\beta(i)$ in model (7).

4.2. The Parameter Estimate of Case Deletion Fuzzy Linear Regression Model

According to Lemma 1 and Theorem 2.1, we can obtain the least squares estimator $\hat{\beta}(i)$.

Theorem 4.1 For the case deletion fuzzy linear regression model (6) or (7), the least squares estimator of $\beta(i)$ is

$$\hat{\beta}(i) = [X_c^T(i) X_c(i)]^{-1} X_c^T(i) Y_c(i) \quad \text{and}$$

$$\hat{\beta}(i) = \hat{\beta} - \frac{(X_c^T X_c)^{-1} x_{ic} \hat{e}_i}{1 - p_{ii}} \quad (10)$$

where $\hat{e}_i = y_{ic} - x_{ic}^T \hat{\beta}$, and x_{ic} is the vector which is composed of the i -th row's element of matrix X_c , p_{ii} is the main diagonal element of matrix $P = X_c (X_c^T X_c)^{-1} X_c^T$. The proof of formula (10) can be obtained in Wei *et al.* [13].

Theorem 4.1 gives a calculation formula of the regression coefficient after the i -th data point deleted and also shows the relationship of the regression coefficient before and after the i -th data point deleted. It is the basis that we evaluate whether this point is a outlier point or a strong impact point. If the i -th data point is a normal point, $\hat{\beta}(i)$ and $\hat{\beta}$ should be little difference. If they have a large difference, it shows that the existence of the i -th data point seriously affect the estimation of β , and this data point may be a outlier point or a strong impact point.

4.3. The Impact Analysis of the Case Deletion Fuzzy Linear Regression Model

At present, the existing method on fuzzy regression model did not consider the actual data which often contains outlier point or strong impact point. However, because of the gross error, rounding error, and other factor's interference, it's difficult to avoid that actual data mixed with a certain proportion of outlier points or strong impact points. Once mixed with outliers, these methods will face serious challenges, and even lead to wrong conclusions. Research about the impact of the data on the model is an important part of statistical diagnosis, and one of the most straightforward way is to delete data [6]. As we transform fuzzy data into clear data, the problem of the fuzzy linear regression analysis is transformed into a traditional linear regression analysis problem. Therefore, the discussion of impact analysis based on data deleted fuzzy regression model can be transformed into a traditional data-deleted linear regression model.

Although we can get $\hat{\beta} - \hat{\beta}(i)$ by formula (8), it's a vector which is difficult to compare. In practice, Cook's distance is often used to measure $\hat{\beta} - \hat{\beta}(i)$. Cook's distance is one of the most important diagnostic statistics, and was originally proposed based on the statistical significance of parameter confidence region by Cook in 1977 [14].

Cook's distance is defined as

$$D_i = D_i(X^T X, p\hat{\sigma}^2) = \frac{(\hat{\beta} - \hat{\beta}(i))^T X^T X (\hat{\beta} - \hat{\beta}(i))}{p\hat{\sigma}^2}$$

here $\hat{\sigma}^2 = \frac{1}{n-p} Y^T (I - P) Y$.

A big D_i shows the estimate $\hat{\beta}(i)$ is far away from the true parameter $\hat{\beta}$, and $\hat{\beta}(i)$ and $\hat{\beta}$ have large difference.

The following theorem is a simple formula for calculating Cook's distance.

Theorem 4.2 [13] For the fuzzy linear regression model (2) or (7), Cook's distance can be expressed as

$$D_i = \frac{\|\hat{Y}_c - \hat{Y}_c(i)\|^2}{p\hat{\sigma}^2} = \frac{p_{ii}}{1 - p_{ii}} \frac{r_i^2}{p}, r_i = \frac{\hat{e}_i}{\hat{\sigma} \sqrt{1 - p_{ii}}}$$

where $\hat{Y}_c = X_c \hat{\beta}$. $\hat{Y}_c(i) = X_c \hat{\beta}(i)$ is the fitted values with before and after deleting the i -th data point.

During the specific data analysis, we first calculate Cook's distance point by point written as D_1, D_2, \dots, D_n ,

and then find one or more particularly large D_i through a list or figure (maybe not particularly large). The data point, with a big effect on parameter estimate, may be the outlier or the strong impact point.

5. Analysis of Practical Example

The following study shows an application of the centroid method and compares the proposed method in this paper with Diamond and Sakawa and Yano’s method.

The data in **Table 1** are triangular fuzzy numbers, and we establish the fuzzy linear regression model $\tilde{Y} = \beta_0 + \beta_1 \tilde{X}$, $\beta_0, \beta_1 \in R$, discuss the model’s error term and the outlier data point.

By Theorem 2.1 (centroid), we can get the fuzzy linear regression equation

$$\hat{Y} = 3.5724 + 0.5139\tilde{X} + (-0.24, 0, 0.24).$$

Using Diamond’s method [3], we can obtain the fuzzy linear regression equation

$$\hat{Y} = (3.2636, 3.5632, 3.8628) + 0.5206\tilde{X}.$$

Also when using Sakawa-Yano’s method [11], we can obtain:

$$\hat{Y} = (3.0307, 3.2010, 3.3713) + (0.4982, 0.5788, 0.6594)\tilde{X}.$$

By formula (7), we calculate the model’s error term E_i and it’s sum using centroid method, Diamond’s method and Sakawa-Yano’s method, and the results are listed in **Table 1**. From **Table 1**, we can find that the sum of the model’s error term using centroid method is less than using Diamond’s method and Sakawa-Yano’s method. Thus, the result of fuzzy linear regression model using centroid method is better than using Diamond’s method and Sakawa-Yano’s method.

Figure 1 obtained by Matlab programming.

Table 2 and **Figure 1** show Cook’s distance under centroid method and their scatter plot, respectively. Because of being with the Cook’s distance. These results indicate that the data point No. 7 is an outlier or strong impact point.

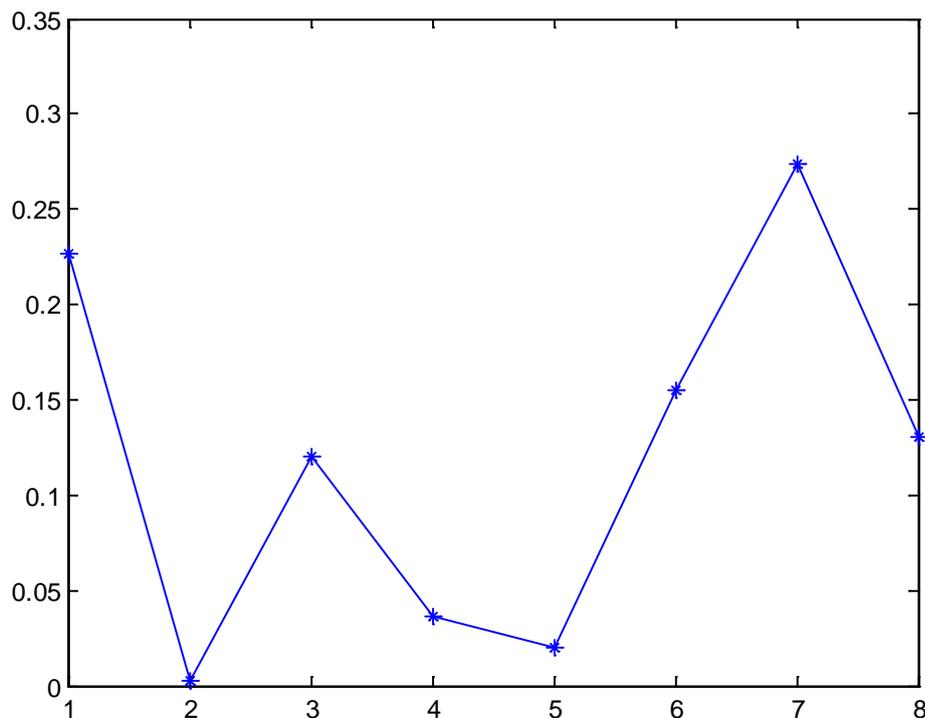


Figure 1. The scatter plot of Cook’s distance under centroid method.

Table 1. Data from Sakawa and Yano [15].

No.	$\tilde{x}_i = (x_{il}, x_{im}, x_{iu})$	$\tilde{y}_i = (y_{il}, y_{im}, y_{iu})$	error term E_i		
			Centroid method	Diamond	Sakawa-Yano
1	(1.5, 2.0, 2.5)	(3.5, 4.0, 4.5)	0.848	0.861	0.633
2	(3.0, 3.5, 4.0)	(5.0, 5.5, 6.0)	0.208	0.227	0.453
3	(4.5, 5.5, 6.5)	(6.5, 7.5, 8.5)	1.489	1.520	1.613
4	(6.5, 7.0, 7.5)	(6.0, 6.5, 7.0)	0.910	0.945	1.165
5	(8.0, 8.5, 9.0)	(8.0, 8.5, 9.0)	0.760	0.785	0.770
6	(9.5, 10.5, 11.5)	(7.0, 8.0, 9.0)	1.449	1.477	1.977
7	(10.5, 11.0, 11.5)	(10.0, 10.5, 11.0)	1.000	1.060	1.368
8	(12.0, 12.5, 13.0)	(9.0, 9.5, 10.0)	0.806	0.834	1.452
The sum of error t error term			7.470	7.709	9.431

Table 2. Cook's distance under centroid method.

No.	$\tilde{x}_i = (x_{il}, x_{im}, x_{iu})$	$\tilde{y}_i = (y_{il}, y_{im}, y_{iu})$	D_i
1	(1.5, 2.0, 2.5)	(3.5, 4.0, 4.5)	0.2267
2	(3.0, 3.5, 4.0)	(5.0, 5.5, 6.0)	0.0030
3	(4.5, 5.5, 6.5)	(6.5, 7.5, 8.5)	0.1199
4	(6.5, 7.0, 7.5)	(6.0, 6.5, 7.0)	0.0362
5	(8.0, 8.5, 9.0)	(8.0, 8.5, 9.0)	0.0202
6	(9.5, 10.5, 11.5)	(7.0, 8.0, 9.0)	0.1554
7	(10.5, 11.0, 11.5)	(10.0, 10.5, 11.0)	0.2735
8	(12.0, 12.5, 13.0)	(9.0, 9.5, 10.0)	0.1306

6. Conclusion

By transforming fuzzy data into clear data, the fuzzy linear regression model is transformed into traditional linear regression model. We study the parameter estimation and impact analysis of the case-deletion fuzzy linear regression model. By comparing with other methods through a practical example, we can conclude that the proposed method in this paper can be used easily and have a good fitting performance.

Acknowledgements

This research is supported by National Natural Science Foundation Grant No. 11171065, the National Statistical Scientific Foundation Grant No. 2014LY059.

References

- [1] Zadeh, L.A. (1975) The Concept of Linguistic Variable and Its Application to Approximate Reasoning. *Information Sciences*, 8, 99-244, 301-357.
- [2] Thanaka, H., Uejina, S. and Asai, K. (1982) Linear Regression Analysis with Fuzzy Model. *IEEE Trans Systems Man Cybernetics*, 12, 903-907.
- [3] Diamond, P. (1988) Fuzzy Least Squares. *Information Science*, 46, 141-157. [http://dx.doi.org/10.1016/0020-0255\(88\)90047-3](http://dx.doi.org/10.1016/0020-0255(88)90047-3)
- [4] Savic, D.A. and Pedrycz, W. (1988) Evaluation of Fuzzy Linear Regression Models. *Fuzzy Sets and Systems*, 46, 141-157.
- [5] Chang, Y.-H.O. and Ayyub, B.M. (2001) Fuzzy Regression Methods—A Comparative Assessment. *Fuzzy Sets and*

- Systems*, **119**, 225-246. [http://dx.doi.org/10.1016/S0165-0114\(99\)00092-5](http://dx.doi.org/10.1016/S0165-0114(99)00092-5)
- [6] Kim, B. and Bishu, R.R. (1998) Evaluation of Fuzzy Linear Regression Model by Comparison Membership Function. *Fuzzy Set and Systems*, **100**, 343-352. [http://dx.doi.org/10.1016/S0165-0114\(97\)00100-0](http://dx.doi.org/10.1016/S0165-0114(97)00100-0)
- [7] Nasrabadi, M.M. and Nasrabadi, E. (2004) A Mathematical-Programming Approach to Fuzzy Linear Regression Analysis. *Applied Mathematics and Computation*, **155**, 873-881. <http://dx.doi.org/10.1016/j.amc.2003.07.031>
- [8] Kao, C. and Chyu, C.L. (2002) A Fuzzy Linear Regression Model with Better Explanatory Power. *Fuzzy Sets and Systems*, **126**, 401-409. [http://dx.doi.org/10.1016/S0165-0114\(01\)00069-0](http://dx.doi.org/10.1016/S0165-0114(01)00069-0)
- [9] Yeh, C.-T. (2011) A Formula for Fuzzy Linear Regression Analysis. 2011 *IEEE International Conference on Fuzzy Systems*, Taipei, 27-30 June 2011, 2845-2850.
- [10] Azadeh, A., Neshat, N. and Rafiee, K. (2015) An Adaptive Neural Network-Fuzzy Linear Regression Approach for Improved Car Ownership Estimation and Forecasting in Complex and Uncertain Environments: The Case of Iran. *Transportation Planning and Technology*, **35**, 221. <http://dx.doi.org/10.1080/03081060.2011.651887>
- [11] Yager, R.R. (1980) On a General Class of Fuzzy Connectives. *Fuzzy Set and Systems*, **4**, 235-242. [http://dx.doi.org/10.1016/0165-0114\(80\)90013-5](http://dx.doi.org/10.1016/0165-0114(80)90013-5)
- [12] Chen, S.J. and Hwang, C.L. (1992) *Fuzzy Multiple Attribute Decision Making*. Springer, NY. <http://dx.doi.org/10.1007/978-3-642-46768-4>
- [13] Wei, B.C., Lin, J.G. and Xie, F.C. (2009) *Diagnostic Statistics*. Higher Education Press, Beijing, 19-44.
- [14] Cook, R.D. (1977) Detection of Influential Observations in Linear Regression. *Technometrics*, **19**, 15-18. <http://dx.doi.org/10.1080/00401706.1977.10489493>
- [15] Sakawa, M. and Yano, H. (1992) Multiobjective Fuzzy Linear Regression Analysis for Fuzzy Input-Output Data. *Fuzzy Set and Systems*, **47**, 173-181. [http://dx.doi.org/10.1016/0165-0114\(92\)90175-4](http://dx.doi.org/10.1016/0165-0114(92)90175-4)