A Comparison of Several Nonparametric Fuzzy Regressions with Trapezoidal Data

T. Razzaghnia [1] ,S.Danesh [2] , A. Maleki [3*]

Abstract–In this paper, three methods of nonparametric fuzzy regression with crisp input and asymmetric trapezoidal fuzzy output, are compared. It analyzes the three nonparametric techniques in statistics, namely local linear smoothing (L-L-S), K- nearest neighbor Smoothing (K-NN) and kernel smoothing (K-S) with trapezoidal fuzzy data to obtain the best smoothing parameters. In addition, it makes an analysis on three real-world datasets and calculates the goodness of fit to illustrate the application of the proposed method.

Keywords:Nonparametric Fuzzy Regression; Trapezoidal Fuzzy Numbers; Local Linear Smoothing (L-L-S); K-Nearest Neighbor Smoothing (K-NN); Kernel Smoothing (K-S)

1. Introduction

In 1982 Tanaka et al.[1] introduced fuzzy regression analysis. After that time, several fuzzy regression approaches have been proposed, including the mathematical programming based methods [1], least squares based methods [2],[3] and other methods [4],[5]. In many real-world problems, it may be unrealistic to predetermine a fuzzy parametric regression relationship especially for a large dataset with a complicated underlying variation trend. Along this line of consideration, some other approaches have been developed to handle the fuzzy regression problems without predefining a specific form of the underlying regression relationship. For instance, Ishibushi and Tanaka [6],[7] have suggested several fuzzy nonparametric regression methods by using the traditional back propagation networks. Also, statistical nonparametric smoothing techniques have achieved significant development in recent years [8],[9],[10]. These smoothing techniques are especially useful to handle the nonparametric regression problems and therefore there may be other promising tools for developing fuzzy nonparametric regression. In this aspect, Cheng and Lee [4] have extended the k-nearest neighbor (K-NN) and kernel smoothing (K-S) methods for the context of fuzzy nonparametric regression. In Wang et al. [11], the local linear smoothing method, the special case of the local polynomial smoothing technique, is fuzzified to handle the fuzzy nonparametric regression with crisp input and LR fuzzy output based on the distance measure proposed by Diamond [12]. Farnoosh et al. [5] used ridge estimation in nonparametric regression with triangular fuzzy data.

In this paper, we propose to analyze the three nonparametric regression techniques in statistical regression, namely local linear smoothing (L-L-S), the K- nearest neighbor smoothing (K-NN) and the kernel smoothing techniques (K-S) with trapezoidal fuzzy data.

This article is organized as follows: In section 2, we have some preliminaries about fuzzy nonparametric regression and trapezoidal fuzzy data. In section 3, smoothing methods for trapezoidal fuzzy numbers are proposed and in section 4, two numerical examples are solved.

2. Preliminaries

A fuzzy number $\widetilde{A}$ is a convex normalized fuzzy subset of the real line **R** with an upper semi-continuous membership functionof bounded support [12].

Definition 2.1. An asymmetric trapezoidal fuzzy number $\widetilde{A}$, denoted by $\widetilde{A}=(a^{(1)},a^{(2)},a^{(3)},a^{(4)})$ is defined as:

1 Department of Statistics, North Tehran Branch, Islamic Azad University, Tehran, Iran

2 Young Researchers and Elite Club, East Tehran Branch, Islamic Azad university, Tehran, Iran.

3* Corresponding Author: Department of Statistics, West tehran Branch, Islamic Azad University, Tehran, Iran. E-mail: maleki.ali@wtiau.ac.ir

Received: 2021.12.07; Accepted: 2022.02.17

$$\tilde{A}(x)=\begin{cases} L(\frac{a^{(2)}-x}{a^{(2)}-a^{(1)}}) & x<a^{(2)} \\ 1 & a^{(2)}\le x\le a^{(3)} \\ R(\frac{x-a^{(3)}}{a^{(4)}-a^{(3)}}) & x>a^{(3)} \end{cases}$$

where $a^{(1)},a^{(2)}a^{(3)},a^{(4)}$ are four parameters of the asymmetric trapezoidal fuzzy number.

Definition 2.2.Suppose that $\tilde{A}=(a^{(1)},a^{(2)},a^{(3)},a^{(4)})$ and $\tilde{B}=(b^{(1)},b^{(2)},b^{(3)},b^{(4)})$ are two trapezoidal fuzzy numbers. Diamond distance between $\tilde{A}$ and $\tilde{B}$ can be expressed as:

$$d^2(\tilde{A},\tilde{B})=(a^{(1)}-b^{(1)})^2+(a^{(2)}-b^{(2)})^2$$
$$+(a^{(3)}-b^{(3)})^2+(a^{(4)}-b^{(4)})^2$$

This distance measures the closeness between two trapezoidal fuzzy membership functions when $d^2(\tilde{A},\tilde{B})=0$. It means that the membership functions of $\tilde{A}$ and $\tilde{B}$ are equal.

Let $F=\left\{\tilde{Y}:\tilde{Y}=(y^{(1)},y^{(2)},y^{(3)},y^{(4)})\right\}$ be a set of all trapezoidal fuzzy numbers. The following univariate fuzzy nonparametric regression model is considered by $Y=F(x)\{+\}\varepsilon$. In this model, X is a crisp independent variable (input) and Y is a symmetric trapezoidal fuzzy dependent variable (output). ε is an error term, and $\{+\}$ is an operator whose definition depends on the fuzzy ranking method used.

In this paper, for the nonparametric regression techniques, K-N-N and K-S are based on the concept of local averaging. In other words, the estimated value of the regression surface at point k_0 is the weighted average of the responses of the observations in the neighborhood of k_0.

Definition 2.3.Let K_i, $i=1,2,\ldots,n$ where the index is in ascending order, then the smoothing function based on local averaging can be represented as:

$$S(K=K_i)=\underset{i-k\le j\le i+k}{AVE}(Y_j)$$
$$=\underset{i-k\le j\le i+k}{AVE}(y_j^{(1)},y_j^{(2)},y_j^{(3)},y_j^{(4)})$$

where AVE denotes the mean, median or any weighted average.

3. Smoothing methods for trapezoidal fuzzy numbers

The basic idea of smoothing is that if a function f is fairly smooth, then the observations made at and near x should contain information about value of x. Thus, it should be possible to use local averaging of the data x to construct an estimator for $F(x)$ which is called the smoother. There are several smoothing techniques. We proposed K-nearest neighbor smoothing (K-NN), kernel-smoothing (K-S) and local linear smoothing (L-L-S) methods for trapezoidal variable in this section.

In the following discussion, asymmetric trapezoidal fuzzy numbers are applied as asymmetric trapezoidal membership functions for deriving nonparametric regression model based on the smoothing parameters.

These models are considered univariate fuzzy nonparametric regression model as:

$$\tilde{Y}=F(x)\{+\}\varepsilon$$
$$=\left(Y^{(1)}(x),Y^{(2)}(x),Y^{(3)}(x),Y^{(4)}(x)\right)\{+\}\varepsilon \quad (1)$$

where Y is a trapezoidal fuzzy dependent variable as output. x is a crisp independent variable as input , $x\in\mathbb{R}$, and x domain is assumed to be D. $F(x)$ is a mapping $D\to F$. The definition of the three smoothing methods for trapezoidal fuzzy variables is as follows:

3.1. Local linear smoothing method (L-L-S)

In the following discussion, Razzaghnia et al. [13] proposed the first linear regression analysis with trapezoidal coefficients. Asymmetric trapezoidal fuzzy numbers are applied as asymmetric trapezoidal membership functions for deriving bivariate regression model. A univariate regression model can be expressed as:

$$\hat{\tilde{Y}}_i=\tilde{A}_0+\tilde{A}_1X_i=\left(a_0^{(1)},a_0^{(2)},a_0^{(3)},a_0^{(4)}\right)$$
$$+\left(a_1^{(1)},a_1^{(2)},a_1^{(3)},a_1^{(4)}\right)\mathrm{X_i} \quad (2)$$

This model can be rewritten as

$$\hat{\tilde{Y}}_i=\begin{pmatrix} a_0^{(1)}+a_1^{(1)}X_i,a_0^{(2)}+a_1^{(2)}\mathrm{X_i},a_0^{(3)} \\ +a_1^{(3)}\mathrm{X_i},a_0^{(4)}+a_1^{(4)}\mathrm{X_i} \end{pmatrix}$$

where $i=1,\ldots,n$ and n is the sample size.

and $\tilde{Y}_i=\left(Y_i^{(1)},Y_i^{(2)},Y_i^{(3)},Y_i^{(4)}\right)$ is an observed value for $i=1,\ldots,n$. So $\hat{\tilde{Y}}_{i.L}$ and $\hat{\tilde{Y}}_{i.R}$ are the left bound and right bound of the predicted $\hat{\tilde{Y}}_i$ at membership h level. Also $\tilde{Y}_{i,L}$ and $\tilde{Y}_{i,R}$ are left bound and right bounds of observed $\tilde{Y}_i$ at membership h level.

Thereupon,

$$\hat{\tilde{Y}}_{i,L} = ha_0^{(2)} + ha_1^{(2)}X_i + (1-h)a_0^{(1)} + (1-h)a_1^{(1)}X_i$$

$$\hat{\tilde{Y}}_{i,R} = ha_0^{(3)} + ha_1^{(3)}X_i + (1-h)a_0^{(4)} + (1-h)a_1^{(4)}X_i$$

$$\tilde{Y}_{i,L} = hY_i^{(2)} + (1-h)Y_i^{(1)}$$

$$\tilde{Y}_{i,R} = hY_i^{(3)} + (1-h)Y_i^{(4)}.$$

Let $\left(X_i, \tilde{Y}_i\right)$ be a sample of the observed crisp inputs and trapezoidal fuzzy outputs with underlying fuzzy regression function of model (2).

$F(x)$ is estimated at any $x \in D$ based on $\left(x_i, \tilde{Y}_i\right)$ for $i = 1,\ldots,n$. When the local linear smoothing technique is used, we shall estimate $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ for each $x \in D$ by using the distance proposed by Diamond [12] as a measure of the fit (Definition 2.2).

This distance is used to fit the fuzzy nonparametric model (1).

Let $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ have continuous derivatives in the domain $x \in D$. Then for a given $x_0 \in D$ and Taylors expansion, $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ can be locally approximated in neighborhood of x_0 , respectively by the following linear functions:

$$Y^{(1)}(x) \simeq \hat{Y}^{(1)}(x) = Y^{(1)}(x_0) + Y'^{(1)}(x_0)(x - x_0) \quad (3)$$

$$Y^{(2)}(x) \simeq \hat{Y}^{(2)}(x) = Y^{(2)}(x_0) + Y'^{(2)}(x_0)(x - x_0) \quad (4)$$

$$Y^{(3)}(x) \simeq \hat{Y}^{(3)}(x) = Y^{(3)}(x_0) + Y'^{(3)}(x_0)(x - x_0) \quad (5)$$

$$Y^{(4)}(x) \simeq \hat{Y}^{(4)}(x) = Y^{(4)}(x_0) + Y'^{(4)}(x_0)(x - x_0) \quad (6)$$

where $Y'^{(1)}(x_0), Y'^{(2)}(x_0), Y'^{(3)}(x_0)$ and $Y'^{(4)}(x_0)$ are respectively, the derivatives of $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ based on Diamond distance (Definition 2.2) and the local linear smoothing method is estimated at x_0 ,

$$F(x_0) = \left(Y^{(1)}(x_0), Y^{(2)}(x_0), Y^{(3)}(x_0), Y^{(4)}(x_0)\right)$$

by minimizing

$$\sum_{i=1}^{n} d^2\left(\tilde{Y}_i, \hat{\tilde{Y}}_i\right) = \sum_{i=1}^{n} d^2\left(\left(Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}\right), \left(\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}\right)\right) K_h(|x_i - x_0|) \quad (7)$$

With respect to $Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}$ and $\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}$ for the given kernel k(.) and smoothing parameter h, where $K_h\left(|x_i - x_0|\right) = k\left(\frac{\frac{|x_i - x_0|}{h}}{h}\right)$ for $i = 1,\ldots,n$ are a sequence of weights at x_0 .Two commonly used kernel functions are parabolic shape functions:

$$k_1(x) = \begin{cases} 0.75(1-x^2) & \text{if } |x| \le 1 \\ 0 & \text{otherwise} \end{cases}$$

and Gaussian function:

$$k_2(x) = (2\pi)^{-1/2} \exp\left(\frac{-x^2}{2}\right)$$

Also, by substituting (3), (4), (5) and (6) at (7), the following can be obtained

$$\sum_{i=1}^{n} d^2\left(\tilde{Y}_i, \hat{\tilde{Y}}_i\right) = \sum_{i=1}^{n} d^2\begin{pmatrix}\left(Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}\right) \\ ,\left(\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}\right)\end{pmatrix} K_h(|x_i - x_0|)$$

$$= \sum_{i=1}^{n}\left(Y_i^{(1)} - Y^{(1)}(x_0) - Y'^{(1)}(x_0)(x_i - x_0)\right)^2 K_h(|x_i - x_0|)$$

$$+ \sum_{i=1}^{n}\left(Y_i^{(2)} - Y^{(2)}(x_0) - Y'^{(2)}(x_0)(x_i - x_0)\right)^2 K_h(|x_i - x_0|)$$

$$+ \sum_{i=1}^{n}\left(Y_i^{(3)} - Y^{(3)}(x_0) - Y'^{(3)}(x_0)(x_i - x_0)\right)^2 K_h(|x_i - x_0|)$$

$$+ \sum_{i=1}^{n}\left(Y_i^{(4)} - Y^{(4)}(x_0) - Y'^{(4)}(x_0)(x_i - x_0)\right)^2 K_h(|x_i - x_0|) \quad (8)$$

By solving this weighted least-squares problem, the following can be obtained

$Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x), Y^{(4)}(x), Y'^{(1)}(x)$
$, Y'^{(2)}(x), Y'^{(3)}(x), Y'^{(4)}(x)$

at x_0 . So the estimation $F(x)$ at x_0 is

$$\hat{\tilde{Y}}(x_0) = (\hat{Y}^{(1)}(x_0), \hat{Y}^{(2)}(x_0), \hat{Y}^{(3)}(x_0), \hat{Y}^{(4)}(x_0)).$$

Equation (8) has eight unknown parameters

$Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x), Y^{(4)}(x), Y'^{(1)}(x_0), Y'^{(2)}(x_0), Y'^{(3)}(x_0), Y'^{(4)}$

to derive a formula for the unknown parameters nonparametric regression based on minimizing this distance, the derivatives (8) with respect to the eight unknown parameters need to be derived, set to zero and solve the eight unknown parameters.

According to the principle of the weighted least-squares and utilizing matrix notations, we can obtain

$$(\hat{Y}^{(1)}(x),\hat{Y}'^{(1)}(x))^T=(X^T(x_0)W(x_0;h)X(x_0))^{-1}X^T(x_0)\,W(x_0;h)\tilde{Y}^{(1)} \quad (9)$$

$$(\hat{Y}^{(2)}(x),\hat{Y}'^{(2)}(x))^T=(X^T(x_0)W(x_0;h)X(x_0))^{-1}X^T(x_0)\,W(x_0;h)\tilde{Y}^{(2)} \quad (10)$$

$$(\hat{Y}^{(3)}(x),\hat{Y}'^{(3)}(x))^T=(X^T(x_0)W(x_0;h)X(x_0))^{-1}X^T(x_0)\,W(x_0;h)\tilde{Y}^{(3)} \quad (11)$$

$$(\hat{Y}^{(4)}(x),\hat{Y}'^{(4)}(x))^T=(X^T(x_0)W(x_0;h)X(x_0))^{-1}X^T(x_0)\,W(x_0;h)\tilde{Y}^{(4)} \quad (12)$$

where

$$X(x_0)=\begin{pmatrix}1 & x_1-x_0\\ 1 & x_2-x_0\\ \vdots & \vdots\\ 1 & x_n-x_0\end{pmatrix},\ \tilde{Y}^{(1)}=\begin{pmatrix}Y_1^{(1)}\\ Y_2^{(1)}\\ \vdots\\ Y_n^{(1)}\end{pmatrix},\ \tilde{Y}^{(2)}=\begin{pmatrix}Y_1^{(2)}\\ Y_2^{(2)}\\ \vdots\\ Y_n^{(2)}\end{pmatrix},\ \tilde{Y}^{(3)}=\begin{pmatrix}Y_1^{(3)}\\ Y_2^{(3)}\\ \vdots\\ Y_n^{(3)}\end{pmatrix},\ \tilde{Y}^{(4)}=\begin{pmatrix}Y_1^{(4)}\\ Y_2^{(4)}\\ \vdots\\ Y_n^{(4)}\end{pmatrix}$$

and $W(x_0;h)=\mathrm{Diag}(K_h(|x_1-x_0|),K_h(|x_2-x_0|),\ldots,K_h(|x_n-x_0|))$

is a $n\times n$ diagonal matrix with its diagonal elements being $K_h(|x_i-x_o|)$ for $i=1,\ldots,n$ and symbol T is transpose of a matrix. If we suppose $e_1=(1,0)^T$ and $H(x_0;h)=(X^T(x_0)W(x_0;h)X(x_0))^{-1}X^T(x_0)W(x_0;h)$

The estimate of $F(x)$ at x_0 is

$$\hat{\tilde{Y}}(x)=\left(\hat{Y}^{(1)}(x_0),\hat{Y}^{(2)}(x_0),\hat{Y}^{(3)}(x_0),\hat{Y}^{(4)}(x_0)\right)=(e_1^TH(x_0;h)\tilde{Y}^{(1)},e_1^TH(x_0;h)\tilde{Y}^{(2)},e_1^TH(x_0;h)\tilde{Y}^{(3)},e_1^TH(x_0;h)\tilde{Y}^{(4)}) \quad (13)$$

3.2. K- Nearest neighbor smoothing (K-NN)

The K-NN weight sequence was introduced by Loftsgaarden and Quesenberry [14] in the related field of density estimation and has been used by Cover and Hart [15] for classification purposes. The K-NN smother is defined as:

$$\tilde{Y}_i=\sum_{j=1}^{n}\omega_j(x)\,Y_j \quad (14)$$

where

$\omega_j(x)$ *for* $j=1,\ldots,n$ is a the weight sequence at x andis defined as

$$\omega_j(x)=\begin{cases}\frac{1}{k},\ j\in J(x)\ \ j=1,\ldots,n\\ 0\qquad\qquad ,otherwise\end{cases} \quad (15)$$

where J_x is one of K-nearest observations to x and $\hat{\tilde{Y}}_i$ is the estimate of the observations $\tilde{Y}_i$ for $i=1,\ldots,n$ and $\tilde{Y}_i=(Y_i^{(1)},Y_i^{(2)},Y_i^{(3)},Y_i^{(4)})$ be asymmetric trapezoidal fuzzy numbers. So based on (14) and (15) we have

$$\hat{\tilde{Y}}_i=(\sum_{j=1}^{n}\omega_jY_j^{(1)},\sum_{j=1}^{n}\omega_jY_j^{(2)},\sum_{j=1}^{n}\omega_jY_j^{(3)},\sum_{j=1}^{n}\omega_jY_j^{(4)})=(\frac{1}{k}\sum_{j\in J(x)}Y_j^{(1)},\frac{1}{k}\sum_{j\in J(x)}Y_j^{(2)},\frac{1}{k}\sum_{j\in J(x)}Y_j^{(3)},\frac{1}{k}\sum_{j\in J(x)}Y_j^{(4)})$$

The K-NN smoothing parameter is the neighborhood size k. So if a relatively small neighborhood size is used, this will increase the regression noise and a relatively large neighborhood size is used which will increase the regression error. So section (3.4) describes leave-one-out cross-validation for finding k optimal value and this can be obtained by minimizing cross-validation criterion.

3.3. Kernel smoothing (K-S)

K-NN smoothing is a weighted averaging neighborhood and weights in neighborhood are treated equally. So in kernel smoothing $S(x)$ is defined by a fixed neighborhood around x. It is determined by kernel function and band with h. The fuzzy regression equations for kernel smoothing and K-NN smoothing are the same and so are represented by equation (14). In kernel smoothing method, $\omega_j(x_o)$ for $j=1,\ldots,n$, at x_0 is defined as

$$\omega_j(x_o)=\frac{K_h(|x_j-x_o|)}{\sum_{i=1}^{n}K_h(|x_i-x_o|)}=\frac{K(\frac{|x_j-x_o|}{h})}{\sum_{i=1}^{n}K(\frac{|x_j-x_o|}{h})}$$

and

$$\hat{\tilde{Y}}_i=(\frac{\sum_{j=1}^{n}K_h(|x_j-x_o|)Y_j^{(1)}}{\sum_{i=1}^{n}K_h(|x_i-x_o|)},\frac{\sum_{j=1}^{n}K_h(|x_j-x_o|)Y_j^{(2)}}{\sum_{i=1}^{n}K_h(|x_i-x_o|)},\frac{\sum_{j=1}^{n}K_h(|x_j-x_o|)Y_j^{(3)}}{\sum_{i=1}^{n}K_h(|x_i-x_o|)},\frac{\sum_{j=1}^{n}K_h(|x_j-x_o|)Y_j^{(4)}}{\sum_{i=1}^{n}K_h(|x_i-x_o|)})$$

the weight sequence is defined by $K_h(x)=\frac{1}{h}K\left(\frac{x}{h}\right)$

which is the kernel with scale factor. So the kernel smoothing parameter is band with h and weight depends on smoothing parameter h.

3.4. Smoothing parameters selection

The most important aspect for averaging techniques and local linear smoothing method is selecting the size of neighborhood to average k and parameter h. There are different methods for selecting parameter h such as the cross-validation method, and generalized cross validation which are used to obtain parameter h. Let

$$\hat{\tilde{Y}}(x_i,h)=\begin{pmatrix}\hat{Y}^{(1)}(x_i,h),\hat{Y}^{(2)}(x_i,h),\\ \hat{Y}^{(3)}(x_i,h),\hat{Y}^{(4)}(x_i,h)\end{pmatrix}$$

The fuzzified cross-validation procedure (CV) for selecting parameter h local linear smoothing method based on Diamond distance is defined as:

$$CV(h)=\frac{1}{n}\sum_{i=1}^{n} d^2\left(\tilde{Y}_i,\hat{\tilde{Y}}_i\right)=\frac{1}{n}\sum_{i=1}^{n}((Y_i^{(1)}-\hat{Y}_i^{(1)})^2+ (Y_i^{(2)}-\hat{Y}_i^{(2)})^2+(Y_i^{(3)}-\hat{Y}_i^{(3)})^2+((Y_i^{(4)}-\hat{Y}_i^{(4)})^2) \quad (16)$$

as its minimization gives the h optimal value.

$CV(h_0)=\min_{h>o} CV(h)$

In fact, we may compute $CV(h)$ for a series of value of h to search for h.

So selected optimal value of h by the $CV(h)$ nearly depends on the degree of smoothness of Y_{iL} and Y_{iR}.

Large value of h leads to lack-of-fit and small value of h makes over-fit.

Also the cross validation leave-one-out technique is used for selecting values k and h in K-NN and KS methods that are obtained by minimizing the cross-validation criterion. According to Stone [16], the CV criterion is defined as

$$CV(b)=\frac{1}{n}\sum_{i=1}^{n} L[\tilde{Y}_i,\hat{\tilde{Y}}_b(x_i,o_{\setminus i})]$$

$$=\frac{1}{n}\sum_{i=1}^{n}(D_i(b)+C_i(b)) \quad (17)$$

where

$$\hat{\tilde{Y}}_b(x_i,o_{\setminus i})=\sum_{j=1 j\neq i}^{n}\omega_j(x_i)\tilde{Y}_j$$

where $\omega_i(x_j)$ is a function with respect to x_j, $D_i(b)$ is the measure of difference and $C_i(b)$ is the measure of inclusion as in:

$$C_i(b)=P\left(\left[\hat{\tilde{Y}}_b(x_i,o_{\setminus i})\right]_\alpha^L-\left[\tilde{Y}_i\right]_\alpha^L\right)+Q\left(\left[\hat{\tilde{Y}}_b(x_i,o_{\setminus i})\right]_\alpha^R-\left[\tilde{Y}_i\right]_\alpha^R\right)$$

where P and Q are penalty terms and are defined as:

$$P=\begin{cases}1, & if\ \left[\tilde{Y}_i\right]_\alpha^L\le\left[\hat{\tilde{Y}}_b(x_i,o_{\setminus i})\right]_\alpha^L\\ 0, & otherwise\end{cases}$$

and

$$Q=\begin{cases}1, & if\ \left[\tilde{Y}_i\right]_\alpha^R\le\left[\hat{\tilde{Y}}_b(x_i,o_{\setminus i})\right]_\alpha^R\\ 0, & otherwise\end{cases}$$

To obtain $D_i(b)$, we will use difference measure of an trapezoidal fuzzy number A, by using the method of Chang and Lee [17], which is defined as:

$$\mathrm{OM}(\mathrm{A})=\int_0^1\varpi(v)\left[\chi_1(v)A_L(v)+\chi_2(v)A_R(v)\right]dv,$$

Thus $D_i(b)$ is calculated by using equation (17). For the calculation of parameter b, we minimize the CV criterion equation (16). b^* which is defined as:

$b^*=argmin\{CV(b)\}$

So b^* is the neighborhood size k in K-NN smoothing method and band with h in kernel smoothing method.

The quantity for comparison between methods of smoothing is goodness offit (GOF). It measures the difference between fuzzy regression function and its estimation. So, based on Diamond distance GOF is defined as:

$$GOF=\frac{1}{n}\sum_{i=1}^{n} d^2\left(\tilde{Y}_i,\hat{\tilde{Y}}_i\right)=\frac{1}{n}\sum_{i=1}^{n}((Y_i^{(1)}-\hat{Y}_i^{(1)})^2+ (Y_i^{(2)}-\hat{Y}_i^{(2)})^2+(Y_i^{(3)}-\hat{Y}_i^{(3)})^2+(Y_i^{(4)}-\hat{Y}_i^{(4)})^2$$

where $\hat{\tilde{Y}}_i$ is the estimation of the fuzzy regression function at all x_i s by one of the smoothing methods. So a very large value of GOF indicates lack-of-fit and a small value shows over-fit for the observed fuzzy outputs.

4. Extension the proposed method to multivariate input

It is straightforward to extent the proposed methods to the case of multivariate input. In fact, let $x=(x_1,x_2,\ldots,x_p)$ be a p-dimensional crisp input and Y be a trapezoidal fuzzy output. The fuzzy nonparametric regression model in this case is of the form

$$Y=Y^{(1)}(x_1,x_2,\ldots,x_p)+Y^{(2)}(x_1,x_2,\ldots,x_p)+Y^{(3)}(x_1,x_2,\ldots,x_p)+Y^{(4)}(x_1,x_2,\ldots,x_p)=F(x)\{+\}\varepsilon$$

4.1 K- Nearest neighbor smoothing (K-NN)

The K-NN smother is defined as:

$$\tilde{Y}_i = \sum_{j=1}^{n} \omega_j(x)\, Y_j \qquad (18)$$

where

$\omega_j(x)$ *for* $j = 1,\dots,n$ is a the weight sequence at x and is defined as

$$\omega_j(x) = \begin{cases} \frac{1}{k}, j \in J(x) & j = 1,\dots,n \\ 0 & ,otherwise \end{cases} \qquad (19)$$

where J_x is one of K-nearest observations to x and $\hat{\tilde{Y}}_i$ is the estimate of the observations $\tilde{Y}_i$ for $i = 1,\dots,n$ and $\tilde{Y}_i = (Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)})$ be asymmetric trapezoidal fuzzy numbers. So based on (14) and (15) we have

$$\hat{\tilde{Y}}_i = (\sum_{j=1}^{n} \omega_j Y_j^{(1)}, \sum_{j=1}^{n} \omega_j Y_j^{(2)}, \sum_{j=1}^{n} \omega_j Y_j^{(3)}, \sum_{j=1}^{n} \omega_j Y_j^{(4)})$$

$$= (\frac{1}{k}\sum_{j\in J(x)} Y_j^{(1)}, \frac{1}{k}\sum_{j\in J(x)} Y_j^{(2)}, \frac{1}{k}\sum_{j\in J(x)} Y_j^{(3)},$$

$$\frac{1}{k}\sum_{j\in J(x)} Y_j^{(4)})$$

The K-NN smoothing parameter is the neighborhood size k. $\hat{\tilde{Y}}_{il}(l = 1,...p)$ are computed for

$X = (x_1, x_2, ..., x_p)$ then

$$\hat{\tilde{Y}}_i = \frac{1}{n}\hat{\tilde{Y}}_{il}$$

4.2 Kernel smoothing method

It is the same K- Nearest neighbor smoothing method but $\omega_j(X_o)$ for $j = 1,\dots,n$, at X_0 is defined as

$$\omega_j(\mathrm{x}_0) = \frac{K_h(\|\mathrm{x}_j - \mathrm{x}_0\|)}{\sum_{i=1}^{n} K_h(\|\mathrm{x}_i - \mathrm{x}_0\|)} = \frac{K(\frac{\|\mathrm{x}_j - \mathrm{x}_0\|}{h})}{\sum_{i=1}^{n} K(\frac{\|\mathrm{x}_i - \mathrm{x}_0\|}{h})}, j = 1, \dots, n,$$

4.3 Local linear smoothing

Suppose that $Y^{(1)}(X), Y^{(2)}(X), Y^{(3)}(X)$ and $Y^{(4)}(X)$ have continuous derivatives in the domain $x \in D$. Then for a given $x_0 \in D$ and Taylor's expansion, $Y^{(1)}(X), Y^{(2)}(X), Y^{(3)}(X)$ and $Y^{(4)}(X)$ can be locally approximated in neighborhood of x_0, respectively by the following linear functions:

$$Y^{(1)}(x) \simeq \hat{Y}^{(1)}(x) = Y^{(1)}(x_0) + Y^{(1)(x_1)}(x_0)(x_1 - x_{01}) + \dots$$
$$+ Y^{(1)(x_p)}(x_0)(x_p - x_{0p}) \quad (20)$$

$$Y^{(2)}(x) \simeq \hat{Y}^{(2)}(x) = Y^{(2)}(x_0) + Y'^{(2)(x_1)}(x_0)(x_1 - x_{01}) + \dots$$
$$+ Y^{(2)(x_p)}(x_0)(x_p - x_{0p}) \quad (21)$$

$$Y^{(3)}(x) \simeq \hat{Y}^{(3)}(x) = Y^{(3)}(x_0) + Y^{(3)(x_1)}(x_0)(x_1 - x_{01}) + \dots$$
$$+ Y^{(3)(x_p)}(x_0)(x_p - x_{0p}) (22)$$

$$Y^{(4)}(x) \simeq \hat{Y}^{(4)}(x) = Y^{(4)}(x_0) + Y^{(4)(x_1)}(x_0)(x_1 - x_{01}) + \dots$$
$$+ Y^{(4)(x_p)}(x_0)(x_p - x_{0p}) \quad (23)$$

where

$Y^{(1)(x_j)}(x_0), Y^{(2)(x_j)}(x_0), Y^{(3)(x_j)}(x_0)$ and $Y^{(4)(x_j)}(x_0)$ are respectively, the derivatives of $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ with respect to (x_j) based on Diamond distance (Definition 2.2) and the local linear smoothing method is estimated at x_0,

$$F(x_0) = \left(Y^{(1)}(x_0), Y^{(2)}(x_0), Y^{(3)}(x_0), Y^{(4)}(x_0)\right)$$

by minimizing

$$\sum_{i=1}^{n} d^2\left(\tilde{Y}_i, \hat{\tilde{Y}}_i\right) = \sum_{i=1}^{n} d^2\left(\left(Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}\right), \left(\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}\right)\right) K_h(\|X_i - X_0\|) \qquad (24)$$

With respect to $Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}$ and $\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}$ for the given kernel k(.) and smoothing parameter h, where

$$K_h(\|X_i - X_0\|) = k\left(\frac{\frac{\|X_i - X_0\|}{h}}{h}\right)$$

for $i = 1,\dots,n$ are a sequence of weights at X_0.

Also, by substituting (20), (21), (22) and (23) at (24), the following can be obtained

$$\sum_{i=1}^{n} d^2\left(\tilde{Y}_i, \hat{\tilde{Y}}_i\right) = \sum_{i=1}^{n} d^2\left(\left(Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}\right), \left(\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}\right)\right) K_h(\|X_i - X_0\|)$$

$$=\sum_{i=1}^{n}\left(Y_i^{(1)}-Y^{(1)}(x_0)-\sum_{j=1}^{p}Y^{(1)(x)_j}(X_0)(x_{ij}-x_{0j})\right)^2 K_h(\|X_i-X_0\|)$$

$$+\sum_{i=1}^{n}\left(Y_i^{(2)}-Y^{(2)}(x_0)-\sum_{j=1}^{p}Y^{(2)(x)_j}(X_0)(x_{ij}-x_{0j})\right)^2 K_h(\|X_i-X_0\|)$$

$$+\sum_{i=1}^{n}\left(Y_i^{(3)}-Y^{(3)}(x_0)-\sum_{j=1}^{p}Y^{(3)(x)_j}(X_0)(x_{ij}-x_{0j})\right)^2 K_h(\|X_i-X_0\|)$$

$$+\sum_{i=1}^{n}\left(Y_i^{(4)}-Y^{(4)}(x_0)-\sum_{j=1}^{p}Y^{(4)(x)_j}(X_0)(x_{ij}-x_{0j})\right)^2 K_h(\|X_i-X_0\|) \quad (25)$$

$\|X_i-X_0\|$ is Euclidean distance between X_i and X_0.

$$\hat{\tilde{Y}}(x)=\left(\hat{Y}^{(1)}(x_0),\hat{Y}^{(2)}(x_0),\hat{Y}^{(3)}(x_0),\hat{Y}^{(4)}(x_0)\right)$$
$$=(e_1^T H(X_0;h)\tilde{Y}^{(1)}, e_1^T H(X_0;h)\tilde{Y}^{(2)}, e_1^T H(X_0;h)\tilde{Y}^{(3)}$$
$$, e_1^T H(X_0;h)\tilde{Y}^{(4)}) \quad (26)$$

where

$$X(x_0)=\begin{pmatrix} 1 & x_{11}-x_{01}\dots x_{1p}-x_{0p} \\ 1 & x_{21}-x_{01}\dots x_{2p}-x_{0p} \\ \vdots & \vdots \\ 1 & x_{n1}-x_{01}\dots x_{np}-x_{0p}\end{pmatrix},$$

$$\tilde{Y}^{(1)}=\begin{pmatrix}Y_1^{(1)}\\ Y_2^{(1)}\\ \vdots \\ Y_n^{(1)}\end{pmatrix}, \tilde{Y}^{(2)}=\begin{pmatrix}Y_1^{(2)}\\ Y_2^{(2)}\\ \vdots \\ Y_n^{(2)}\end{pmatrix}, \tilde{Y}^{(3)}=\begin{pmatrix}Y_1^{(3)}\\ Y_2^{(3)}\\ \vdots \\ Y_n^{(3)}\end{pmatrix}, \tilde{Y}^{(4)}=\begin{pmatrix}Y_1^{(4)}\\ Y_2^{(4)}\\ \vdots \\ Y_n^{(4)}\end{pmatrix}$$

and $W(X_0;h)=\mathrm{Diag}(K_h(\|X_1-X_0\|), K_h(\|X_2-X_0\|),\dots,K_h(\|X_p-X_0\|))$

is a $n\times n$ diagonal matrix with its diagonal elements being $K_h(\|x_i-x_o\|)$ for $i=1,\dots,n$ and symbol

T is transpose of a matrix. If we suppose $e_1=(1,0)^T$ and $H(x_0;h)=(X^T(X_0)W(X_0;h)X(X_0))^{-1}X^T(X_0)W(X_0;h)$

The estimate of $F(x)$ at x_0 is

$$\hat{\tilde{Y}}(x)=\left(\hat{Y}^{(1)}(x_0),\hat{Y}^{(2)}(x_0),\hat{Y}^{(3)}(x_0),\hat{Y}^{(4)}(x_0)\right)$$
$$=(e_1^T H(x_0;h)\tilde{Y}^{(1)}, e_1^T H(x_0;h)\tilde{Y}^{(2)},$$
$$e_1^T H(x_0;h)\tilde{Y}^{(3)}, e_1^T H(x_0;h)\tilde{Y}^{(4)}) \quad (27)$$

5. Numerical Examples and Conclusion

In this section, there are two examples in which the input is a crisp number and the output is a trapezoidal fuzzy number. We estimate the values by using three smoothing methods. Then these methods can be compared with each other and for this purpose, their GOF and their charts are used.

Example 1: This example is a generated dataset in the same way as that in Cheng and Lee [4] The following function is considered $f(x)=\frac{x^2}{5}+2e^{\frac{x}{10}}$

So x_i is uniformly generated within the interval [0, 1] and i=1,…,100,

$\tilde{Y}_i=\left(Y_i^{(1)},Y_i^{(2)},Y_i^{(3)},Y_i^{(4)}\right)$, so

$=(y_i-e_i, y_i+\frac{1}{3}e_i, y_i+\frac{2}{3}e_i, y_i+e_i)$

$y_i=f(X_i)+rand[-0.5,0.5]$ and

$e_i=1/4f(X_i)+rand[0,1]$.

Local Linear smoothing method, K-NN and kernel smoothing are applied to the fitting model. So Gauss and Parabolic shape kernel are used to produce the weight sequence for local linear smoothing and kernel smoothing methods. Table 1 shows smoothing parameter selected by cross-validation procedure results from different methods. Figures 1, 2 and 6 show the results of three methods. These results can be compared using figure 3 and table 4. Like the previous example, L-L-S method is better than K-NN, and K- S methods. In table 3, GOF of L-L-S method is lower than K-NN, K- S methods.

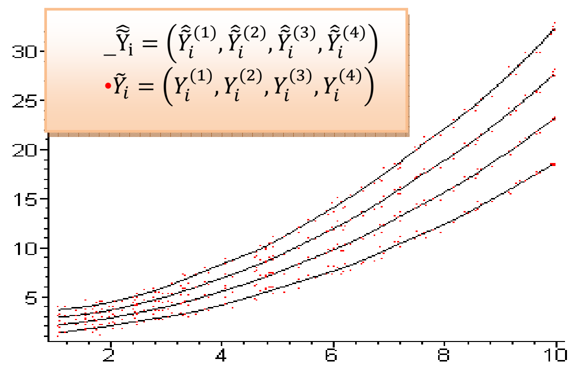

Fig.1 Obtained results by L-L-S method with Gausian kernel for h=0.43

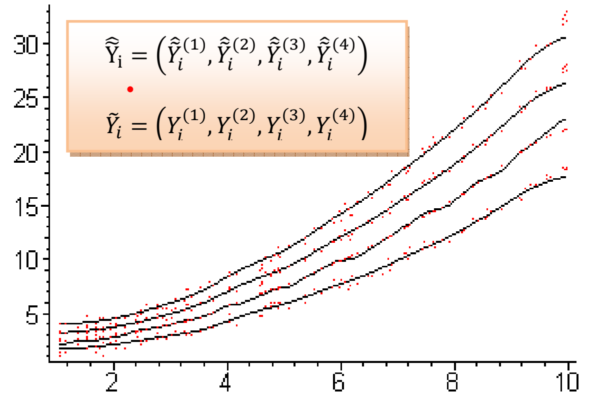

Fig.2 Obtained results by K-S method with Gausian kernel for h=0.12

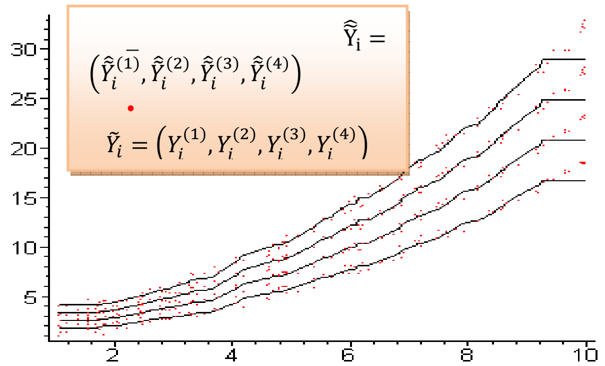

Fig.3 Obtained results by KNN method with K=19

Table 1 The obtained results of different methods for example 1

method	kernel	Smoothing parameter	GOF
KNN	-	19	0.328
KS	Gauss Parabolic shape	0.12 1.72	0.30 0.0085
LLS	Gauss Parabolic shape	0.43 1.2	0.0045 0.0046

Example 2: Consider the following function:

$$f(x_1, x_2) = 24.23r^2(0.75 - r^2) + 5,$$
$$r^2 = (x_1/10 - 0.5)^2 + (x_2/10 - 0.5)^2.$$

where the domain of $X = (x_1, x_2)$ is $D = [0,10]^2$. A set of data is generated the same way as that in [18] and in the following manner.

The crisp inputs of the independent variables x_1 and x_2 are randomly taken from 0 to 10. Let output $\tilde{Y}_i$ is a trapezoidal fuzzy number and it is generated by:

$$\tilde{Y}_i = (Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}) = (y_i - e_i, y_i + \frac{1}{3}e_i, y_i + \frac{2}{3}e_i, y_i + e_i), \text{ so}$$

$$\begin{cases} y_i = f(x_{i1}, x_{i2}), \\ e_i = (1/4)f(x_i) + rand[0,1], \end{cases} i = 1, \dots, 30,$$

where rand $[a, b]$ denotes a random number between a and b for each i. the different methods are applied to fit regression model. The error value of GOF are numerically used to evaluate the performance of the different methods. So Gauss kernel is used to produce the weight sequence for local linear smoothing and kernel smoothing methods.

Tables 2 and 3 show the obtained results from different methods. These results can be compared each other. Like

the previous examples, L-L-S method is better than K-NN, and K- S methods.

Table 2 The input data and the fitted fuzzy outputs by three smoothing methods

x_1	x_2	$\tilde{Y}_i = (Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)})$	K-NN	K-S	L-L-S
0.5160	**5.9760**	**(2.1750, 2.8910, 3.6070, 4.3230)**	**(2.7107, 3.6238, 4.5369, 5.4500)**	**(2.36,3.15,3.94,4.729)**	**(2.148,2.85,3.55,4.25)**
0.7250	**2.1290**	**(2.6730, 3.6457, 4.6183, 5.5910)**	**(2.7107, 3.6238, 4.5369, 5.4500)**	**(2.67,3.64,4.62,5.59)**	**(2.67,3.64,4.61,5.59)**
0.8070	**9.8370**	**(3.2840, 4.3347, 5.3853, 6.4360)**	**(2.7543, 3.7417, 4.7290, 5.7163)**	**3.28,4.33,5.38,6.43)(**	**(3.28,4.32,5.38,6.43)**
0.8910	**7.1500**	**(2.3060, 3.2447, 4.1833, 5.1220)**	**(2.7843, 3.8252, 4.8661, 5.9070)**	**(2.52,3.51,4.5,5.48)**	**(2.41,3.28,4.45,5.46)**
1.0710	**7.4820**	**(2.7630, 3.8963, 5.0297, 6.1630)**	**(2.5983, 3.5637, 4.5290, 5.4943)**	**(2.57,3.598,4.62,5.66)**	**(2.55,3.47,4.69,5.77)**
1.1940	**6.2100**	**(2.7260, 3.5500, 4.3740, 5.1980)**	**(2.6270, 3.6628, 4.6986, 5.7343)**	**(2.58,3.43,4.28,5.13)**	**(2.41,3.23,3.92,4.68)**

1.3000	4.8520	(2.3920, 3.5420, 4.6920, 5.8420)	(3.5160, 4.5847, 5.6533, 6.7220)	(2.48,3.59,4.69,5.79)	(2.40,3.42,4.68,5.82)
2.6390	5.7270	(5.4300, 6.6620, 7.8940, 9.1260)	(3.7950, 5.0861, 6.3772, 7.6683)	(5.24,6.50,7.75,9.01)	(5.43,6.86,7.9,9.136)
2.8300	3.6310	(3.5630, 5.0543, 6.5457, 8.0370)	(4.5013, 5.9142, 7.3271, 8.7400)	(4.1,5.58,7.07,8.55)	(3.70,5.219,6.71,8.22)
2.9670	8.8280	(4.5110, 6.0263, 7.5417, 9.0570)	(4.5137, 6.0977, 7.6817, 9.2657)	(5.51,7.23,8.94,10.68)	(4.77,6.32,7.91,9.94)
3.1610	7.1270	(5.4670, 7.2123, 8.9577, 10.7030)	(5.0813, 6.7167, 8.3520, 9.9873)	(5.37,7.05,8.73,10.41)	(5.36,7.09,8.74,10.43)
3.1780	7.1450	(5.2660, 6.9113, 8.5567, 10.2020)	(5.2347, 6.9227, 8.6107, 10.2987)	(5.37,7.051,8.74,10.42)	(5.35,7.08,8.73,10.43)
3.5530	4.2610	(4.9710, 6.6443, 8.3177, 9.9910)	(5.4497, 7.1930, 8.9363, 10.6797)	(4.79,6.31,7.82,9.34)	(4.87,6.259,8.12,9.74)
3.6090	8.8930	(6.1120, 8.0233, 9.9347, 11.8460)	(5.3610, 6.9981, 8.6352, 10.2723)	(6.17,8.02,9.86,11.72)	6.37,8.27,10.17,12.06)(
37110	3.8310	(5.0000, 6.3267, 7.6533, 8.9800)	(5.9820, 7.6284, 9.2749, 10.9213)	(4.77,6.24,7.72,9.19)	(4.94,6.54,7.72,9.05)
3.9190	8.9900	(6.8340, 8.5353, 10.2367, 11.9380)	(6.1337, 7.8408, 9.5479, 11.2550)	(6.38,8.26,10.16,12.053)	(6.620,8.559,10.46,12.38)
3.9220	8.9840	(6.5670, 8.6603, 10.7537, 12.8470)	(6.6517, 8.5981, 10.5446, 12.4910)	(6.38,8.27,10.16,12.053)	(6.622,8.562,10.467,12.388)
4.4860	9.4330	(6.5540, 8.5987, 10.6433, 12.6880)	(6.6747, 8.7682, 10.8618, 12.9553)	(6.67,8.64,10.61,12.57)	(6.77,8.54,10.88,12.93)
5.2480	6.6770	(6.9030, 9.0457, 11.1883, 13.3310)	(7.4800, 9.5636, 11.6471, 13.7307)	(6.97,9.13 ,11.29,13.45)	(6.94,9.1,11.25,13.40)
5.6960	9.6320	(8.9830, 11.0463, 13.1097, 5.1730)	(7.3267, 9.3551, 11.3836, 13.4120)	(8.69, 10.74,12.81,14.87)	(8.97,11.02,13.099,15.16)
5.8020	0.2170	(6.0940, 7.9733, 9.8527, 11.7320)	(7.4383, 9.3919, 11.3454, 13.2990)	(6.10, 7.98,9.86,11.74)	(6.09,7.97,9.85,11.73)
6.4390	2.6600	(7.2380, 9.1560, 11.0740, 12.9920)	(6.5437, 8.5377, 10.5317, 12.5257)	(7.08, 9.13,11.18,13.23)	(7.104,8.83,10.825,12.95)
6.3490	2.3490	(6.2990, 8.4837, 10.6683, 12.8530)	(7.1963, 9.3701, 11.5439, 13.7177)	7.12, 9.18,11.24,13.31)(	(6.67,8.806,10.825,13.904)
6.6280	6.6090	(8.0520, 10.4707, 12.889, 15.3080)	(7.5373, 9.7349, 11.9324, 14.1300)	(7.97, 10.37,12.777,15.178)	(8.042,10.45,12.85,15.29)
7.1060	2.0840	(8.2610, 10.2503, 12.2397, 4.2290)	(8.1450, 10.3594, 12.5739, 14.7883)	(7.81, 9.91,12.01,14.106)	(8.12,10.41,12.35,14.39)
7.1880	1.7980	(8.1220, 10.3573, 12.5927, 4.8280)	(8.5080, 10.6344, 12.7609, 14.8873)	(7.94, 10.05,12.17,14.28)	(8.11,10.43,12.425,12.583)
7.5790	3.5590	(9.1410, 11.2957, 13.4503, 5.6050)	(9.4363, 11.8377, 14.2390, 16.6403)	(8.99,11.13,13.27,15.41)	(9.14,11.28,13.45,15.603)
8.9310	7.5400	(11.0460, 13.860, 16.6740, 9.4880)	(10.7697, 13.5014, 16.2332, 8.9650)	(11.53, 14.5,17.49,20.48)	(11.198,14.06,16.98,19.87)
9.2970	7.1810	(12.1220, 15.3487, 18.5753, .8020)	(12.1727, 15.1822, 18.1918, 21.2013)	(11.697,14.75,17.81,20.87)	(12.066,15.29,18.48,21.69)
9.7360	8.8430	(13.3500, 16.3380, 19.3260, .3140)	(12.1727, 15.1822, 18.1918, 21.2013)	(13.25,16.23,19.22,22.20)	(13.35,16.35,19.33,22.31)

Table 3 The obtained results of different methods for example 2

Method	kernel	Smoothing parameter	GOF
K-NN	-	3	2.7217
K-S	Gauss	0.6	0.6683
L-L-S	Gauss	0.75	0.1394

References

[1] H. Tanaka, S. Uejima, K. Asia, "Linear regression analysis with fuzzy model", IEEE Transactions on Systems, Man, and Cybernetics,Vol12,P 903-907, 1982.

[2] Danesh, S., Farnoosh, R., Razzaghnia, T. (2015). "Fuzzy nonparametric regression based on adaptive neuro fuzzy inference system", Neurocomputing, Vol173, P1450-1460, 2015.

[3] Naderkhani, R., Behzad, M.H., Razzaghnia, T. et al. "Fuzzy Regression Analysis Based on Fuzzy Neural Networks Using Trapezoidal Data". Int. J. Fuzzy

Syst. Vol23, P1267–1280, 2021.

[4]Cheng, C. B., Lee, E. S ,. “Nonparametric fuzzy regression K-NN and Kernel Smoothing techniques”, Computers and Mathematics with Applications, Vo 38, P 239-251, 1999.

[5] R. Farnoosh, J. Ghasemian and o. SolaymaniFard, “A modification on ridge estimation for fuzzy nonparametric regression”, Iraninan Journal of Fuzzy systems,Vol 9,P 75-88, 2012.

[6] H. Ishibushi, H. Tanaka, “Fuzzy regression analysis using neural networks”, Fuzzy Sets and Systems, Vol50,P257-265,1992.

[7] H. Ishibushi, H. Tanaka, “Fuzzy neural networks with interval weights and its application to fuzzy regression analysis”, Fuzzy Sets and Systems, Vol 57,P 27-39,1993.

[8] J. Fan, I. Gijbels, Local polynomial modeling and its applications, Chapman & Hall, London, 1996.

[9] W. Hardle, Applied Nonparametric Regression, Cambridge University Press, New York, 1990.

[10] M.Danesh ,S.Danesh , T.Razzaghnia, A.Maleki, “Prediction of fuzzy nonparametric regression function: a comparative study of a new hybrid method and smoothing methods”, Global Analysis and Discrete Mathematics, Vol 6, Issue 1, P 143–177,2021.

[11] N.Wang, W.X. Zhang and C.L Mei, “Fuzzy nonparametric regression based on local linear smoothing technique”, Information Sciences, Vol177, P 3882-3900, 2007.

[12] P. Diamond, “Fuzzy least squares”, Information Sciences,Vol46, P 141-157, 1988.

[13] T. Razzaghnia, E. Pasha, E. Khorram, A. Razzaghnia, “Fuzzy linear regression analysis with trapezoidal coefficients”, First Joint Congress On Fuzzy And Intelligent Systems, 2007, Aug. 29-31, Mashhad, Iran.

[14] D. O. Loftsgaarden and G.P. Quesenberry, “A nonparametric estimate of a multivariate density function”, Annals of Mathematical Statistics,Vol36, P1049-1051, 1965.

[15] R. Coppi, P.D'Urso, P. Giordani, A Santoro, “Least squares estimation of a linear regression model with LR fuzzy response”, Computational Statistics and Data Analysis, Vol 51,P 267-286,2006.

[16] M. Stone, “Cross-validatory choice and assessment of statistical predictions”, Journal of the Royal Statistical Society, VOl 36 (Series B), P111-147, 1994.

[17] Chang, P. T., Lee, E. S., “A generalized fuzzy weighted least-squares regression”, Fuzzy Sets and Systems,Vol82, P289-298, 1996.

[18] C.-B. Cheng, E. S. Lee, “Applying Fuzzy Adoptive Network to Fuzzy Regression Analysis”,Computers and Mathematics with Applications,Vol38, P 123-140, 1999.