Journal of Uncertain Systems
Vol.2, No.3, pp.177-186, 2008
Online at: www.jus.org.uk

Scalar Fuzzy Regression Models

Renkuan Guo[1,*], Danni Guo[2], Christien Thiart[1]
[1] *Department of Statistical Sciences, University of Cape Town, Private Bag, Rondebosch 7701, Cape Town, South Africa*
[2] *Kirstenbosch Research Center, South African National Biodiversity Institute Private Bag X7, Claremont 7735, Cape Town, South Africa*

Received 10 January 2008; Accepted 20 April 2008

Abstract

In this paper, we propose a scalar variable formation of fuzzy regression model based on the axiomatic credibility measure foundation. The fuzzy estimation for fuzzy regression coefficients is investigated. A general M-estimation criterion is developed under Maximum Fuzzy Uncertainty Principle, which resulted in weighted Normal equation with adjusted term for M-estimator of the regression coefficients. Finally, we explore the fuzzy one-way classification model, the M-estimation in general and the concept of estimable function with respect to the one-way model.

Keywords: credibility measure, fuzzy regression, fuzzy one-way classification model, M-estimator, estimable function.

1 Introduction

In statistical theory, regression is an important topic for modeling the functional relationship between response variable(s) and exploratory variable(s) under random uncertainty assumptions. When data is fuzzy, fuzzy regression models were also developed although mostly on the ground of Zadeh's fuzzy mathematics. Therefore, before engaging the fuzzy regression literature review, it is necessary to mention the foundation for fuzzy regression model. Up to now, fuzzy regression models were almost all developed on the ground of Zadeh's fuzzy set theory (1965). Many set variable oriented fuzzy regression models were proposed: Tanaka *at al.* (1980, 1982) initiated an approach of fuzzy regression which minimizes the fuzziness as an optimal criterion, Diamond (1987, 1988) used least-squared errors as a decision criterion, interval regression oriented fuzzy regression model was also presented (see, Dubois and Prade, 1980, Kacprzyk and Fedrizzi, 1992 and others).

It is also noticed that there were variable styled treatments in fuzzy regression in terms of numerical valued approaches in which the representative values are utilized, say, the fuzzy mode, the fuzzy average, the fuzzy median, or the mid-range of α -cut set of fuzzy membership function to specify the fuzzy subset. However the fundamental weakness of these numerical valued treatments on fuzzy subsets lies on the utilization of the partial information of the fuzzy subsets under study. In general it is admitted that the existing fuzzy regression models are difficult to play more actual roles in industrial and business applications because of the intrinsic weakness roots in its membership-based fuzzy mathematical foundation created by Zadeh (1965).

It should be emphasized here, the contribution of Zadeh's fuzzy mathematics should not be forgotten, without him, today's mathematics may still wonder in the crisp set theoretical ground. Zadeh revealed another essential form of uncertainty – fuzziness and showed the difference from the well known and accepted essential form of uncertainty – randomness. However, on the one side, the significance of Zadeh's membership based fuzzy theory has its own intrinsic limitations. Some mathematicians, say, Walley (1991) and some Bayesian statisticians have different view on fuzzy set membership functions. It is fair to say, Zadeh's membership-based fuzzy mathematics reshaped modern mathematics and was successful in many industrial applications. No one should simply close the eyes but simply reject those facts of successful achievements. It is also true that his efforts in developing a possibility measure theory for characterizing fuzziness failed to create a true counterpart of that of probability measure theory because Zadeh's

* Corresponding author. Email: rguo@worldonline.co.za (R. Guo)

(intuitive) possibility measure does not have self-duality property (i.e., the possibility measure of the complement event A^c does not equal to one subtracts the possibility measure of event A).

Although Zadeh's possibility measure concept is intuitive and straightforward, which led a way for addressing the fuzziness, nevertheless, the non-self-duality of possibility measure prevents the scalar variable treatments of a fuzzy variable and thus Zadeh's fuzzy mathematics is often inconvenient and very complicated in mathematical manipulations. However, we still believe Zadeh's work set us a milestone of modern (or non-classical mathematics). Finally, the membership specification is still a thorny issue. During almost four decades, fuzzy researchers have to specify membership function and set up the parameter values in terms of their own working experiences. Compared to the probabilistic "counterpart", for random variable and its distribution it is very data-oriented, very rich statistical estimation and hypothesis testing theory have been developed for the parameters in specifying the distribution objectively. The fuzzy statistical theory developed very slowly, and its applications are difficult due to the set-oriented foundation and set-form variable treatments.

To resolve the difficulties facing the Zadeh's membership-based fuzzy mathematics, Liu (2004, 2007) proposed an axiomatic foundation for modeling fuzzy phenomena, credibility measure theory. The credibility measure possesses self-duality property and is able to play the role of that in probability theory. Furthermore, fuzzy variable concept and its (credibility) distribution, which are parallel to these in probability theory, are developed.

The credibility measure (of a fuzzy event) is an average of the possibility measure and necessity measure (of a fuzzy event) and thus is self-dual. It is true that a credibility measure concept is not as intuitive as that of the possibility measure or necessity measure alone in revealing the degree of fuzziness, however, credibility measure does reveal the degree of fuzziness in a composite manner and address the weakness of possibility measure or necessity measure alone. Some researchers on possibility measure theoretical ground argued that Liu's credibility measure is too special and lack of justifications for regarding it as a measure for fuzziness.

We will have a simple argument for the justification as long as a fuzzy researcher believes Zadeh's membership function is a way for characterizing a fuzzy set, denoted as $\mu_{\tilde{A}}(\cdot)$. For this purpose, we state a trivial principle called as *Measure Equivalence Principle*. For a given uncertain event, as long as two measures (or functions) are one-one mapping each other, the two measures (or functions) are said to be equivalent and therefore they quantify the same kind of uncertainty. Recall that in probability theory, probability distribution of a random variable, denoted as $F_X(\cdot)$, is used for characterizing random uncertainty. Then any one-to-one mapping of $F_X(\cdot)$ can be also used for measuring the random uncertainty, for example, moment generating function, $m_X(\theta) = \mathrm{E}\left[e^{\theta X}\right]$, or characteristic function, $\phi_X(\theta) = \mathrm{E}\left[e^{i\theta X}\right]$. The same principle will be applicable in credibility measure case. As a matter of fact, Liu (2004, 2007) has established the two theorems between credibility measure and membership function. For any event $\{\xi = x\}, x \in B \subset \mathbb{R}$,

$$\mu(x) = \left(2\mathrm{Cr}\{\xi = x\}\right) \wedge 1, \quad x \in \mathbb{R} \ . \tag{1}$$

Conversely, for any event $\{\xi \le x\}$, the credibility distribution, $\Lambda(x) = \mathrm{Cr}\{\xi \le x\}$ links to membership function

$$\Lambda(x) = \frac{1}{2}\left(\sup_{y \le x} \mu(y) + 1 - \sup_{y > x} \mu(x)\right), \ x \in \mathbb{R} \ . \tag{2}$$

It is obvious a loose 1-1 mapping relation between membership function which implies possibility distribution and credibility distribution function exists. Then we will question these who suspect the role and value of credibility measure theory, except for accepting the fact that a credibility measure quantifies the fuzzy uncertainty, what are their choices?

In this paper, based on Liu's (2004, 2007) classical credibility measure theory, i.e., $(\vee, \wedge)$-credibility measure theory, we develop a scalar variable oriented treatment in terms of an *M*-estimation for the fuzzy regression coefficients, which leads to weighted least-squares formation.

The structure of this paper is as follows. Section 2 is used for reviewing Liu's credibility measure theory, defining the scalar fuzzy variable and its credibility distribution. In Section 3, the *M*-function utilizing Maximum Fuzzy Uncertainty Principle is introduced and therefore, a weighted normal equation with adjusted term is derived. In Section 4, the *M*-estimators for regression coefficients under maximum membership criterion is derived. In Section 5, we investigate the fuzzy one-way classification model. Finally a few concluding remarks are offered in Section 6.

2 A Review of Credibility Measure Theory

Let Θ be a nonempty set, and 2^Θ the power set on Θ. A power set is the set class containing all the possible subsets of nonempty set Θ, i.e., $2^\Theta = \{A : A \subset \Theta\}$. It is obvious that a power set 2^Θ is the largest σ-algebra on Θ. Each element of a power set, say, $A \subset \Theta$, $A \in 2^\Theta$ is called an event. A number denoted as $\mathrm{Cr}\{A\}$, $0 \le \mathrm{Cr}\{A\} \le 1$, is assigned to an arbitrary event $A \in 2^\Theta$, which indicates the credibility grade with which event $A \in 2^\Theta$ occurs. For any $A \in 2^\Theta$, set function $\mathrm{Cr}\{A\}$ satisfies the following axioms (Liu, 2004, 2007):

Axiom 1: $\mathrm{Cr}\{\Theta\} = 1$.

Axiom 2: $\mathrm{Cr}\{\cdot\}$ is non-decreasing, i.e., whenever $A \subset B$, $\mathrm{Cr}\{A\} \le \mathrm{Cr}\{B\}$.

Axiom 3: $\mathrm{Cr}\{\cdot\}$ is self-dual, i.e., for any $A \in 2^\Theta$, $\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = 1$.

Axiom 4: $\mathrm{Cr}\{\bigcup_i A_i\} \wedge 0.5 = \sup_i \left[\mathrm{Cr}\{A_i\}\right]$ for any $\{A_i\}$ with $\mathrm{Cr}\{A_i\} \le 0.5$.

Axiom 5: Let set functions $\mathrm{Cr}_k\{\cdot\} : 2^{\Theta_k} \to [0,1]$ satisfy **Axioms 1-4**, and $\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_p$, then

$$\mathrm{Cr}\left\{(\theta_1, \theta_2, \cdots, \theta_p)\right\} = \mathrm{Cr}_1\{\theta_1\} \wedge \mathrm{Cr}_2\{\theta_2\} \wedge \cdots \wedge \mathrm{Cr}_p\{\theta_p\}, \tag{3}$$

where $(\theta_1, \theta_2, \cdots, \theta_p) \in \Theta$.

Definition 2.1: (Liu, 2004, 2007) Any set function $\mathrm{Cr} : 2^\Theta \to [0,1]$ satisfies **Axioms 1-4** is called a $(\vee, \wedge)$**-credibility** measure (or simply a credibility measure). The triple $(\Theta, 2^\Theta, \mathrm{Cr})$ is called the credibility measure space.

A credibility measure satisfies all the properties of uncertainty measure and also many of its own. For space limitation reason, we can only review minimal materials, but more technical details can be found in Liu (2004, 2007).

Definition 2.2: (Liu, 2004, 2007) A fuzzy variable ξ is a measurable mapping from credibility space $(\Theta, 2^\Theta, \mathrm{Cr})$ to the set of real numbers, i.e., $\xi : (\Theta, 2^\Theta, \mathrm{Cr}) \to \mathbb{R}$.

We should be fully aware that on the credibility measure platform, a fuzzy variable is recorded as a real-valued number similar to that of a random variable. Definitely, similar to random variable, a real number as a realized value of a fuzzy variable has a distributional grade associated with it.

Definition 2.3: (Liu, 2004, 2007) The credibility distribution $\Lambda : \mathbb{R} \to [0,1]$ of a fuzzy variable ξ on $(\Theta, 2^\Theta, \mathrm{Cr})$ is

$$\Lambda(x) = \mathrm{Cr}\left\{\theta \in \Theta \,\middle|\, \xi(\theta) \le x\right\} \quad , \tag{4}$$

The credibility distribution $\Lambda(x)$ is the accumulated *credibility grade* that the fuzzy variable ξ takes a value less than or equal to a real-number $x \in \mathbb{R}$. Generally speaking, the credibility distribution $\Lambda(\cdot)$ is neither left-continuous nor right-continuous. What we will deal with are absolutely continuous fuzzy variables with continuous credibility density functions and thus poses no further restrictions on our developments.

Definition 2.4: (Liu, 2004, 2007) Let $\Lambda(\cdot)$ be the credibility distribution of the fuzzy variable ξ. Then a function $\lambda : \mathbb{R} \to [0, +\infty)$ is called a credibility density function of fuzzy variable ξ if

$$\Lambda(x) = \int_{-\infty}^{x} \lambda(y)\,\mathrm{d}y, \quad \forall x \in \mathbb{R}. \tag{5}$$

The axiomatic credibility measure foundation is the starting point of Liu's fuzzy theory, while the definition of a membership function is the fundamental starting point of Zadeh's fuzzy set theory. Zadeh (1965, 1978) further proposed possibility measure based theoretical framework and expected the possibility measure could be a counterpart of probability measure, nevertheless, Zadeh failed his own mission. It is necessary to emphasize here that with or without membership function fuzzy phenomena in real world can be accurately described by the credibility measure models. Linking between credibility measure and membership plays a role of bridging Zadeh's fuzzy mathematics and the new axiomatic fuzzy theory and thus provides a conversion channel.

3 The M-Estimation for Regression Coefficients

A fuzzy linear model describes a functional relationship containing fuzzy uncertainty. For simplicity, let us start with the simple fuzzy regression model

$$Y = \alpha + \beta x + \varepsilon, \tag{6}$$

where x is exploratory (or independent or controllable) variable, Y is the fuzzy response (or dependent variable), ε

is a fuzzy error term with $\mathrm{E}[\varepsilon]=0$ and $V[\varepsilon]=\sigma^2$. Note that the expectation is taken with respect to the credibility distribution. Denote the empirical (or fitted) fuzzy linear regression of Y with respect to x by

$$\hat{Y}=\hat{\alpha}+\hat{\beta}x \quad , \tag{7}$$

where vector $(\hat{\alpha},\hat{\beta})$ is the estimate of regression coefficient vector (α,β).

Let $\{(x_1,Y_1),(x_2,Y_2),\cdots,(x_n,Y_n)\}$ be a simple fuzzy sample. $Y_1,Y_2,\cdots,Y_n$ denote the *n* observed response corresponding to $x_1,x_2,\cdots,x_n$. In (probabilistic) linear model theory, the error terms are assumed to be $\varepsilon\sim N(0,\sigma^2)$ and ε_i and ε_j are uncorrelated. Thus, $Y_i\sim N(\alpha+\beta x_i,\sigma_i^2)$ and Y_i and Y_j are uncorrelated too.

The next question is to set up an objective function, under which the parameters of the fuzzy regression can be chosen such that the estimated parameters could optimize the objective function. On the ground of Zadeh's fuzzy set theory, where membership plays the critical role, Maximum Membership Principle guides the objective function setting up. However, the Maximum Membership Principle is only applicable whenever the membership is symmetric. Furthermore, in this paper we intend to explore a universal criterion on the ground of Liu's credibility measure theory.

Liu (2007) states Maximum Uncertainty Principle: "For any event, if there are multiple reasonable values that a measure may take, then the value as close as to 0.5 as possible is assigned to the event". Guiding by this principle, we state a specific criterion for the fuzzy uncertainty case.

Maximum Fuzzy Uncertainty Principle: If there are multiple reasonable fuzzy values that a credibility measure may take, then the credibility measure for an event should be a value as close to 0.5 as possible.

Let us assume that events $\{A_i,\ i=1,\cdots,n\}$ are defined on a credibility space $(\Theta,2^{\Theta},\mathrm{Cr})$. Then the objective function can be

$$J=\sum_{i=1}^{n}\left(\mathrm{Cr}\{A_i\}-0.5\right)^2 \quad . \tag{8}$$

For the simple regression defined by Eq. (8), the credibility distribution of error term ε is assumed to take a form

$$\Lambda_{\varepsilon}(u)=\mathrm{Cr}\{\theta:\varepsilon(\theta)\le u\} \quad . \tag{9}$$

In other words, the credibility distribution of ε is defined on the set

$$\Theta=\{\theta\mid\varepsilon(\theta)=Y-(\alpha+\beta x)\} . \tag{10}$$

Then the objective function under the Maximum Fuzzy Uncertainty Principle takes a form

$$J(\alpha,\beta)=\sum_{i=1}^{n}\left(\Lambda_{\varepsilon}\left(Y_i-(\alpha+\beta x_i)\right)-0.5\right)^2 . \tag{11}$$

Note that

$$\Lambda_{\varepsilon}\left(Y_i-(\alpha+\beta x_i)\right)=\mathrm{Cr}\{\theta:\varepsilon(\theta)\le Y_i-(\alpha+\beta x_i)\} . \tag{12}$$

It is obvious that for a symmetric type membership family, the objective function defined in Eq. (11) is similar to the Maximum Membership Principle, although the solution may not be the same. Nevertheless, our criterion for setting up objective function is universal no matter what kind of membership families we might face.

Now, it is ready to return back to the (scalar) fuzzy simple regression model fitting problem, which now becomes one of finding an empirical linear regression equation $\hat{Y}=\hat{a}+\hat{b}x$ based on the observations $\{(x_1,Y_1),(x_2,Y_2),\cdots,(x_n,Y_n)\}$ such that the credibility distribution at $\hat{Y}_i=\hat{a}+\hat{b}x_i$ should be a value as close to 0.5 as possible. Then minimizing the objective function J defined in Eq. (11) with respect to α and β, the M-equation system for coefficients (α,β) is

$$\begin{cases}\sum_{i=1}^{n}\left(\Lambda_{\varepsilon}\left(Y_i-(\alpha+\beta x_i)\right)-0.5\right)\lambda_{\varepsilon}\left(Y_i-(\alpha+\beta x_i)\right)=0\\ \sum_{i=1}^{n}\left(\Lambda_{\varepsilon}\left(Y_i-(\alpha+\beta x_i)\right)-0.5\right)\lambda_{\varepsilon}\left(Y_i-(\alpha+\beta x_i)\right)x_i=0 \quad .\end{cases} \tag{13}$$

It should be aware that, if the characterization of credibility distribution of error term ε may involve other parameters, denoted as $\underline{\gamma}^T=(\gamma_1,\gamma_2,\cdots,\gamma_l)$, then full M-equation system should take the form

$$\begin{cases} \sum_{i=1}^{n}\left(\Lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)-0.5\right)\lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)=0 \\ \sum_{i=1}^{n}\left(\Lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)-0.5\right)\lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)x_i=0 \\ \sum_{i=1}^{n}\left(\Lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)-0.5\right)\frac{\partial\Lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)}{\partial\gamma_1}=0 \\ \vdots \\ \sum_{i=1}^{n}\left(\Lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)-0.5\right)\frac{\partial\Lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)}{\partial\gamma_l}=0. \end{cases} \tag{14}$$

Eq. (13) is merely taking the first two equations from Eq. (14) with the assumption that all the parameters, denoted as $(\alpha,\beta,\gamma_1,\cdots,\gamma_l)$, satisfy or be determined by Eq. (14) only. Eq. (13) cannot determine coefficients (α,β) alone if $\underline{\gamma}^T=(\gamma_1,\gamma_2,\cdots,\gamma_l)$ participate in credibility distribution.

Now, for non ambiguity, we rewrite Eq. (13) as

$$\begin{cases} \sum_{i=1}^{n}\left(\Lambda_\varepsilon\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)-0.5\right)\lambda_\varepsilon\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)=0 \\ \sum_{i=1}^{n}\left(\Lambda_\varepsilon\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)-0.5\right)\lambda_\varepsilon\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)x_i=0. \end{cases} \tag{15}$$

Notation $\left(\hat\alpha,\hat\beta\right)$ implies that $(\alpha,\beta,\gamma_1,\cdots,\gamma_l)$ are M-estimator which solves the M-equation system Eq. (14).

Notice that the function form of commonly used credibility distributions (induced by membership functions) may be classified into three basic categories: (a) step function; (b) linear function containing factor like $w(u-c)$ (c) Nonlinear function, containing factor like $w(u-c)^d$ or $\exp\left(-w(u-c)^d\right)$. Accordingly, (error) term $Y_i-\left(\hat\alpha+\hat\beta x_i\right)$ will either appear in $\Lambda_\varepsilon(\cdot)$ (linear category) or in $\lambda_\varepsilon(\cdot)$ (nonlinear category) and therefore an appropriate weighted normal equation will be derived from Eq. (15).

For linear category, we take triangular fuzzy error term with parameter (a,b,c) , the credibility distribution is

$$\Lambda_{\varepsilon_i}\left(Y_i-(\alpha+\beta x_i)\right)=\begin{cases} 0 & Y_i-(\alpha+\beta x_i)\le a \\ \frac{Y_i-(\alpha+\beta x_i)-a}{2(b-a)} & a<Y_i-(\alpha+\beta x_i)\le b \\ \frac{Y_i-(\alpha+\beta x_i)+c-2b}{2(c-b)} & b<Y_i-(\alpha+\beta x_i)\le c \\ 1 & Y_i-(\alpha+\beta x_i)>c. \end{cases} \tag{16}$$

It is ready to write

$$\begin{aligned} \Lambda_{\varepsilon_i}\left(Y_i-(\alpha+\beta x_i)\right)&=\left(1-\vartheta_{(-\infty,a]}\left(Y_i-(\alpha+\beta x_i)\right)\right)\vartheta_{(-\infty,a]}\left(Y_i-(\alpha+\beta x_i)\right) \\ &+\frac{\vartheta_{(a,b]}\left(Y_i-(\alpha+\beta x_i)\right)}{2(b-a)}\left(Y_i-(\alpha+\beta x_i)-a\right) \\ &+\frac{\vartheta_{(b,c]}\left(Y_i-(\alpha+\beta x_i)\right)}{2(c-b)}\left(Y_i-(\alpha+\beta x_i)+c-2b\right)+\vartheta_{(c,+\infty)}\left(Y_i-(\alpha+\beta x_i)\right). \end{aligned} \tag{17}$$

Notice that the optimal $\Lambda_\varepsilon\left(Y_i-(\alpha+\beta x_i)\right)$ should approach 0.5 and therefore $Y_i-(\alpha+\beta x_i)$ should falls in $(b-\delta,b+\delta)$ and thus the first term and fourth term in Eq. (17) disappears for small $\delta>0$. In other words, term $Y_i-(\alpha+\beta x_i)$ can be isolated and then Eq. (17) can be written as

$$\begin{cases} \sum_{i=1}^{n}w_i\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)=\sum_{i=1}^{n}w_ir_i \\ \sum_{i=1}^{n}w_i\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)x_i=\sum_{i=1}^{n}w_ir_ix_i, \end{cases} \tag{18}$$

where r_i is the remaining terms and w_i is the appropriate weight term from the rearrangements in Eq. (17). This will result in a weighted normal equation with adjusted term.

As to the credibility distribution containing nonlinear factor case, for illustrative purpose, let the credibility distribution take a form

$$\Lambda_\varepsilon\left(Y_i-\left(\alpha+\beta x_i\right)\right)=\begin{cases}\dfrac{1}{2}e^{-\tau\left(Y_i-\left(\alpha+\beta x_i\right)\right)^2}, & Y_i-\left(\alpha+\beta x_i\right)<0\\ \dfrac{1}{2}\left(2-e^{-\tau\left(Y_i-\left(\alpha+\beta x_i\right)\right)^2}\right), & Y_i-\left(\alpha+\beta x_i\right)\ge 0.\end{cases} \tag{19}$$

It will be very clear that

$$\Lambda_{\varepsilon_i}\left(Y_i-\left(\alpha+\beta x_i\right)\right)=g\left(-\tau\left(Y_i-\left(\alpha+\beta x_i\right)\right)^2\right). \tag{20}$$

Therefore, Eq. (17) will take a form

$$\begin{cases}\sum_{i=1}^{n}\left(\Lambda_{\varepsilon_i}\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)-0.5\right)g'\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)=0\\ \sum_{i=1}^{n}\left(\Lambda_{\varepsilon_i}\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)-0.5\right)g'\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)x_i=0.\end{cases} \tag{21}$$

Define

$$w_i^{-1}=\left(\Lambda_{\varepsilon_i}\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right)-0.5\right)g'\left(Y_i-\left(\hat\alpha+\hat\beta x_i\right)\right). \tag{22}$$

and

$$\underline{Y}=\begin{bmatrix}Y_1\\Y_2\\ \vdots\\Y_n\end{bmatrix},\ X=\begin{bmatrix}1 & x_1\\1 & x_2\\ \vdots & \vdots\\1 & x_n\end{bmatrix},\ \underline{\Gamma}=\begin{bmatrix}\alpha\\ \beta\end{bmatrix},\ W=\begin{bmatrix}w_1 & 0 & \cdots & 0\\0 & w_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\0 & 0 & \cdots & w_n\end{bmatrix}. \tag{23}$$

Then the weighted normal equation is obtained as

$$X^TW^{-1}X\underline{\Gamma}=W^{-1}X^T\underline{Y}. \tag{24}$$

4 M-estimator for Coefficients with Normal Formed Membership

Let the residual take a membership function of form

$$\mu_\varepsilon\left(Y_i-\left(a+bx_i\right)\right)=g\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right). \tag{25}$$

where $g(\cdot)$ is a differentiable function with $g(0)=1$, and $\lim_{x\to-\infty} g(x)=0$.

It is obvious that for this case, Maximum Membership Principle works, and accordingly we can define the *M*-functional equation system for fuzzy regression coefficients (α,β).

Definition 4.1: Given the differentiable membership function $g\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)$, which measures the degree of belongingness to empirical linear regression line $\hat Y=a+bx$ at observation pair (Y_i,x_i), then the normal formed M-functional system based on the *n* observations $\{(x_1,Y_1),(x_2,Y_2),\cdots,(x_n,Y_n)\}$ takes the form

$$\begin{cases}\sum_{i=1}^{n}h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(Y_i-\left(a+bx_i\right)\right)=0\\ \sum_{i=1}^{n}h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(Y_i-\left(a+bx_i\right)\right)x_i=0,\end{cases} \tag{26}$$

where $h(x)=g'(x)=dg/dx$.

Theorem 4.2: Let a simple regression model $Y=\alpha+\beta x+\varepsilon$ assumes a fuzzy error ε $E[\varepsilon]=0$, $V[\varepsilon]=\sigma^2$ and membership function $g\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)$. For given *n* pair of independent observations $\{(x_1,Y_1),(x_2,Y_2),\cdots,(x_n,Y_n)\}$, a general M-estimator of the coefficients for fitted regression $\hat Y=a+bx$, (a,b) is the solution to the general M-function equation system Eq (26). Furthermore, the *M*-estimator (a,b) takes a weighted least-square estimator form as

$$\begin{cases} b = \dfrac{\sum_{i=1}^{n} h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)\left(Y_i-\overline{Y}_h\right)}{\sum_{i=1}^{n} h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)^2} \\ a=\overline{Y}_h-\overline{b}x_h, \end{cases} \tag{27}$$

where

$$\overline{x}_h=\sum_{i=1}^{n}\frac{h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n} h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)}x_i, \overline{Y}_h=\sum_{i=1}^{n}\frac{h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n} h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right)}Y_i\ . \tag{28}$$

Theorem 4.2 is easy to prove by expanding the left side terms of Eq. (26), re-arrange them, for obtaining Eq. (27). Let

$$\underline{Y}=\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix},\ X=\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix},\ D=\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}, \tag{29}$$

where $d_i=h\left(-\tau\left(Y_i-\left(a+bx_i\right)\right)^2\right),\ i=1,\cdots,n$. Further, let

$$\underline{\Gamma}=\begin{bmatrix} a \\ b \end{bmatrix},\ W=D^{-1}=\begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix}. \tag{30}$$

Then, the M-functional equation system Eq. (26) can be re-written in a matrix form

$$X^T W^{-1} X \underline{\Gamma} = W^{-1} X^T \underline{Y}. \tag{31}$$

Eq. (31) takes the weighted least-squares normal equation form in statistical linear model theory. However, the weighted least-squares formation of the *M*-functional equation system of Eq. (26) will help further mathematical treatments. For example, Eq. (31) can be re-expressed in matrix form

$$\hat{\underline{\Gamma}}=\left(X^T W^{-1} X\right)^{-1} W^{-1} X^T \underline{Y}. \tag{32}$$

as long as the inverse matrix exists.

Example 4.3: Let the membership function g take a normal form

$$g\left(-\left(Y-\left(a+bx\right)\right)^2\right)=\exp\left(-\tau\left(Y-\left(a+bx\right)\right)^2\right). \tag{33}$$

Then the derivative of membership function is

$$h\left(-w\left(Y-\left(a+bx\right)\right)^2\right)=\exp\left(-w\left(Y-\left(a+bx\right)\right)^2\right). \tag{34}$$

The factor τ is sample-dependent and can be defined by

$$\tau=\frac{2}{d_{\max}-d_{\min}}\quad , \tag{35}$$

where

$$d_{\max}=\max_{i\in\{1,2,\cdots,n\}}\left\{\left(Y_i-\left(a+bx_i\right)\right)^2\right\}, d_{\min}=\min_{i\in\{1,2,\cdots,n\}}\left\{\left(Y_i-\left(a+bx_i\right)\right)^2\right\}. \tag{36}$$

Then the *M*-estimators for regression coefficients are

$$\begin{cases} b = \dfrac{\sum_{i=1}^{n} \exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_\infty\right)\left(Y_i-\overline{Y}_\infty\right)}{\sum_{i=1}^{n} \exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_\infty\right)^2} \\ a=\overline{Y}_\infty-\overline{b}x_\infty, \end{cases} \tag{37}$$

where

$$\overline{x}_{\infty}=\sum_{i=1}^{n}\frac{\exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n}\exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)}x_i,\overline{Y}_{\infty}=\sum_{i=1}^{n}\frac{\exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n}\exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)}Y_i \quad . \tag{38}$$

5 Fuzzy One-Way Classification Models

It is not difficult to extend the simple fuzzy regression model to multiple regression case, i.e., the error term

$$\varepsilon_i=Y_i-\left(\beta_0+\beta_1x_{1i}+\beta_2x_{2i}+\beta_lx_{li}\right). \tag{39}$$

The mathematical treatments are essential the same as the simple regression case. However, we may face the problem to evaluate the effects of factors. Let us to examine a simple example. Let Y_{ij} denote the response of j^{th} individual of i^{th} type of factor (crisp-valued level), where $i=1,2,\cdots,l$ and $j=1,2,\cdots,m_i$. The problem is to estimate the fuzzy effect of the factor on individual. The fuzzy model takes a form

$$Y_{ij}=\nu+\alpha_i+\varepsilon_{ij}\ . \tag{40}$$

where ν represent the overall mean response of individual, α_i represents the effect of type i, and ε_{ij} is the fuzzy residual term peculiar to the response Y_{ij}. The feature of fuzzy error ε_{ij} will be discussed later. First, we assume the membership of the error term takes a form

$$\Lambda_\varepsilon\left(Y_{ij}-\left(\nu+\alpha_1+\cdots+\alpha_l\right)\right)=g\left(-\tau\left(Y_{ij}-\left(\nu+\alpha_1+\cdots+\alpha_l\right)\right)^2\right). \tag{41}$$

Denote

$$\underline{Y}=\begin{bmatrix}Y_{11}\\ \vdots\\ Y_{1n_1}\\ Y_{21}\\ \vdots\\ \vdots\\ Y_{ln_l}\end{bmatrix},\ X=\begin{bmatrix}1&1&0&\cdots&0\\ \vdots&\vdots&\vdots&\cdots&\vdots\\ 1&1&0&\cdots&0\\ 1&0&1&\ddots&\vdots\\ \vdots&\vdots&\vdots&\vdots&1\\ \vdots&\vdots&\vdots&\vdots&\vdots\\ 1&0&0&0&1\end{bmatrix},\ \underline{\varepsilon}=\begin{bmatrix}\varepsilon_{11}\\ \vdots\\ \varepsilon_{1n_1}\\ \varepsilon_{21}\\ \vdots\\ \vdots\\ \varepsilon_{ln_l}\end{bmatrix}\ \underline{\Gamma}=\begin{bmatrix}\nu\\ \alpha_1\\ \vdots\\ \alpha_l\end{bmatrix} \tag{42}$$

and

$$W^{-1}=\begin{bmatrix}w_1^{-1}&0&\cdots&0\\ 0&w_2^{-1}&\cdots&0\\ \vdots&\vdots&\ddots&\\ 0&0&\cdots&w_n^{-1}\end{bmatrix}, \tag{43}$$

where $w_i^{-1}=g'\left(-\tau\left(Y_{ij}-\left(\nu+\alpha_1+\cdots+\alpha_l\right)\right)\right)$.

Then the Maximum Membership Principle results in the weighted normal equation

$$X^TW^{-1}X\underline{\Gamma}=W^{-1}X\underline{Y}\ . \tag{44}$$

Note that the matrix is not of full rank and thus the inverse does not exist and we can only engage the generalized inverse of matrix G such that

$$\left(W^{-1/2}X\right)^TG\left(W^{-1/2}X\right)=X^TW^{-1}X\ ,W^{-1/2}XGX^TW^{-1}X=W^{-1/2}X \tag{45}$$

and furthermore $\left(W^{-1/2}X\right)^TG\left(W^{-1/2}X\right)$ is invariant to G and symmetric. Then

$$\hat{\underline{\Gamma}}=GW^{-1}X\underline{Y}\ . \tag{46}$$

The solution is not unique. However, the estimable functions are unique. For further discussions, define matrix H such that

$$H=GX^TW^{-1}X\ . \tag{47}$$

Then

$$\hat{\underline{\Gamma}}=GW^{-1}X\underline{Y}+\left(H-I\right)\underline{z}, \tag{48}$$

where $\underline{z}$ is an arbitrary vector.

Definition 5.1: A linear function $\underline{\mathbb{k}}^T \underline{\Gamma}$ is estimable with respect to fuzzy linear model $Y_{ij} = \nu + \alpha_i + \varepsilon_{ij}$ if

$$\mathrm{E}\left[\underline{\mathbb{k}}^T \hat{\underline{\Gamma}}\right] = \underline{\mathbb{k}}^T \underline{\Gamma}, \tag{49}$$

Example of estimable functions, say, $\alpha_j - \alpha_1$, $j = 2, 3, \cdots, l$.

Remark 5.2: For given observations $\underline{Y}$, the fuzzy estimator of estimable function is just $\underline{\mathbb{k}}^T \hat{\underline{\Gamma}}$. For example,

$$\hat{\alpha}_j - \hat{\alpha}_1 = \frac{1}{m_j} \sum_{i=1}^{m_j} Y_{ji} - \frac{1}{m_1} \sum_{i=1}^{m_1} Y_{1i}, \tag{50}$$

which can be interpreted as the effect of j^{th} factor. Nevertheless, the significance level can not be justified because the distribution of relevant ratio of fuzzy quadratic forms has not been developed yet.

In the one-way classification model is assuming that the factors, α_i, $i = 1, 2, \cdots, l$, are fuzzy-valued level with given membership functions, $\mu_{\alpha_i}(\cdot)$, $i = 1, 2, \cdots, l$. The mathematical treatments are similar to the crisp-factor level case.

6 Conclusions

In this paper, we investigate the estimation problem of the scalar fuzzy regression model on the ground of Liu's credibility measure theory. Furthermore, we propose the estimation methodology in fuzzy one-way classification model parallel to Searle's work (1982). Definitely, the idea of a general fuzzy linear model theory roots in this paper. The open question is the hypothesis testing problems which will rely on the distributional theory of fuzzy quadratic forms.

Furthermore, fuzzy data may be sampled from different fuzzy error structures, i.e., fuzzy errors are assuming different membership functions. Then additional information on fuzzy memberships should be given.

As to some researchers in fuzzy regression may think and query whether our approach can generate the same modeling results as the conventional ones, say, Tanaka *at al.* (1980, 1982). The answer to this query is no because the observations in Tanaka's fuzzy regression model are fuzzy sets with memberships fully assumed. The comparison of Tanaka's fuzzy regression model to our scalar fuzzy regression model is just similar to that in statistics, regression for given random variables as observations and regression for observations from a population of given distribution are not the same. In our future paper we will address the regression modeling with Tanaka's data.

We should strongly emphasize here, Liu's credibility measure theory is mathematically rigorous and sound. The application of Liu's theory will gain a great convenience in mathematical treatments. We have noticed that there appear some voices within fuzzy research communities. Some researchers feel the credibility measure is "too special" and the credibility measure theory is too new to be justified. They demand a justification with probability measure. They used the book of Walley (1991) as a cornerstone to fire bullets towards credibility measure theory. However, our reading on the book of Walley (1991) leads us to have opposite conclusion since Walley treated membership function as a second-order probability. Even so, Walley (1991) is very negative on membership function just as he said, "If fuzzy sets have a useful role to play, it is in modeling the ambiguity of ordinary language. But ambiguity is only one potential source of imprecision in probabilities and, unlike some other sources of imprecision, it can be eliminated through careful elicitation. The membership μ_p that are chosen to model ambiguous probability judgments seem both arbitrary and inappropriately precise. No clear interpretation of μ_p has been established. Its assessment requires substantial input from a user, in addition to that needed for sensitivity analysis. (The extra assessments are comparable to those needed to determine second-order probabilities.)" If we accept Walley's standing point, we should be serious on the modification on fuzzy theory based on membership function, including possibility measure theory. Liu's work (2004, 2007) just reflect such a demand because his axiomatic foundation does not requires membership at all although he is still establishing certain link between his foundation to old fuzzy school.

7 Acknowledgements

This research is partially supported by South African National Research Foundation Grant FA2006042800024. Authors thank deeply for Professor B. D. Liu's comments on my previous drafted version.

References

[1] Diamond, P., Least squares fitting of several fuzzy variables, in: J.C. Bezdek (Ed.), *Analysis of Fuzzy Information*, CRC Press, Tokyo, pp.329-331, 1987.
[2] Diamond, P., Fuzzy least squares, *Inform. Sci.*, vol.46, pp.141-157, 1988.
[3] Dubois, D., and H. Prade, *Theory and Applications, Fuzzy Sets and Systems*, Academic Press, New York, 1980.
[4] Kacprzyk, J., and M. Fedrizzi, *Fuzzy Regression Analysis*. Physica-Verlag, Heidelberg, 1992.
[5] Kaufmann, A., *Introduction to the Theory of Fuzzy Subsets*, Academic Press, New York, 1975.
[6] Liu, B., *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*, Springer-Verlag Heidelberg, Berlin, 2004.
[7] Liu, B., *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*, 2nd Edition, Springer-Verlag Heidelberg, Berlin, 2007.
[8] Searle, S.R., *Matrix Algebra Useful For Statistics*, John Wiley and Sons, New York, 1982.
[9] Tanaka, H., S. Uejima, and K. Asia, Fuzzy linear regression model, *Internat. Congress on Applied Systems Research and Cybernetics*, Aculpoco, Mexico, 1980.
[10] Tanaka, H., S. Uejima, and K. Asia, Linear regression analysis with fuzzy model, *IEEE Trans. Systems Man. Cybernet*, vol.12, no.6, pp.903-907, 1982.
[11] Walley, P., *Statistical Reasoning with Imprecise Probabilities*, Campman and Hall, London, 1991.
[12] Zadeh, L.A., Fuzzy sets. *Information and Control*, vol.8, pp.338-353, 1965.
[13] Zadeh, L.A., Fuzzy sets as basis for a theory of possibility, *Fuzzy Sets and Systems*, vol.1, pp.3-28, 1978.