CREDIBILISTIC FUZZY REGRESSION

R. Guo
•
University of Cape Town, Cape Town, South Africa
e-mail: RenkuanGuo@uct.ac.za

D. Guo
•
South African National Biodiversity Institute, Cape Town, South Africa
e-mail: guo@sanbi.org

ABSTRACT

In reliability, quality control and risk analysis, fuzzy methodologies are more and more involved and inevitably introduced difficulties in seeking fuzzy functional relationship between factors. In this paper, we propose a scalar variable formation of fuzzy regression model based on the credibility measure theoretical foundation. It is expecting our scalar variable treatments on fuzzy regression models will greatly simplify the efforts to seeking fuzzy functional relationship between fuzzy factors. An *M*-estimator for the regression coefficients is obtained and accordingly the properties and the variance-covariance for the coefficient *M*-estimators are also investigated in terms of weighted least-squares arguments. Finally, we explore the asymptotic membership function for the coefficient *M*-estimators.

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 (1965, 1978). We noticed that more and more system dynamics researchers in reliability, quality and risk analysis engage into fuzzy approach. However, whenever we use fuzzy set theory for practical modeling, we will face a sequence of fundamental issues:

The first one is the self-duality in its theoretical foundation. Fuzzy mathematics initiated by Zadeh (1965) facilitated a foundation dealing with vague phenomena in fuzzy modeling. Nevertheless, the fuzzy mathematical foundation initiated by Zadeh (1965, 1978) is membership function and possibility measure based and widely used. The possibility measure was originally expected to play the role of probability measure in probability theory, but could not because it does not possess self-duality property as that in probability theory.

The second issue is the variable-orientation issue. In standard probability theory, random variable and the distribution function play important roles for converting set-based arguments into variable-based arguments, which result in great conveniences in applications. Kaufmann (1975) first proposed the concept of fuzzy variable with the intention of creating its counterpart in probability theory. Unfortunately, Kaufmann's fuzzy variable is in fact another name for a fuzzy subset, and the mathematical operations are difficult to handle.

The third one is the membership specification 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, very rich (data-oriented) statistical estimation and hypothesis testing theory have been developed. The fuzzy statistical theory developed very slowly, and its applications are difficult due to the set-oriented foundation.

To resolve the three dilemmas, Liu (2004, 2007) proposed an axiomatic foundation for modelling 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.

Many set variable oriented fuzzy regression models were proposed: Tanaka et 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 (Dubois and Prade, 1980),(Kacprzyk and Fedrizzi, 1992). It is also noticed that there is variable treatment in fuzzy regression in terms of numerical valued approaches in terms of the usage of representative values, 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 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 two is used for reviewing Liu's credibility measure theory, defining the scalar fuzzy variable and further comparing Liu's and Zadeh's fuzzy theories. In Section 3, the *M*-function utilizing maximum membership grades is introduced and therefore, the *M*-estimators for regression coefficients are derived. In Section 4, the properties of regression coefficient *M*-estimators are investigated and in Section five, we propose asymptotic membership function under the assumptions of fuzzy errors being taken normal membership function, which will lead to asymptotic credibility distribution. In section 6, we explore the fuzzy regression formation when the sample data are taken from different memberships. Section 7 offers an extension to multiple regression treatment. Finally a few concluding remarks are offered in Section 8.

2 A REVIEW ON 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 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: $\mathrm{Cr}\{\cup_i A_i\}\wedge 0.5=\sup_i\left[\mathrm{Cr}\{A_i\}\right]$ 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:

$$\begin{aligned}&\mathrm{Cr}\{\theta_1,\theta_2,\cdots,\theta_p\}\\&=\mathrm{Cr}_1\{\theta_1\}\wedge\mathrm{Cr}_2\{\theta_2\}\wedge\cdots\wedge\mathrm{Cr}_p\{\theta_p\}\end{aligned}\tag{1}$$

where $\{q_1, q_2, L\ , q_p\} O\, 2^Q$.

Definition 2.1: (Liu, 2004, 2007) Any set function $\mathrm{Cr} : 2^Q$ ® $[0,1]$ satisfies **Axioms 1-4** is called a **(Ъ,Ш)**-credibility measure (or simply a credibility measure). The triple $(Q, 2^Q, \mathrm{Cr})$ 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 (2007).

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

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 $\Lambda : \mathrm{R} \to [0,1]$ of a fuzzy variable ξ on $(Q, 2^Q, \mathrm{Cr})$ is:

$$\Lambda(x) = \mathrm{Cr}\left\{\theta \in \Theta \middle| \xi(\theta) \le x\right\} \tag{2}$$

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 $x \in \mathrm{R}$. Generally speaking, the credibility distribution $\Lambda(\cdot)$ is neither left-continuous nor right-continuous. What we will deal with is 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 function $\lambda : \mathrm{R} \to [0, +\infty)$ of a fuzzy variable ξ is called a credibility density function such that,

$$\Lambda(x) = \int_{-\infty}^{x} \lambda(y)\,\mathrm{d}y, \quad \forall x \in \mathrm{R} \tag{3}$$

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 (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. Table 1 offers comparisons between the two fuzzy theories:

Table 1. Comparison between Zadeh's and Liu's Fuzzy Theories

Item	Zadeh's	Liu's
Cornerstone concept	Possibility measure $\mathrm{Poss}\{\cdot\}$	Credibility measure $\mathrm{Cr}\{\cdot\}$
Axiomatic Foundation	No	Yes, four axioms
Membership	Initial concept $\mu_A : \Theta \to [0,1]$	Induced $\mu(x) = (2\mathrm{Cr}\{\xi = x\}) \wedge 1$
Measure space	$(Q, P\%, \mathrm{Poss})$	$(Q, 2^Q, \mathrm{Cr})$
Self-duality	No, $\mathrm{Poss}\{A\} + \mathrm{Poss}\{A^c\} \ne 1$	Yes, $\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = 1$
Identical transmogrification	No. For any fuzzy sets $\tilde{A}, \tilde{B}, \tilde{C}$, $\tilde{A} = \tilde{B} + \tilde{C}$; Does not imply $\tilde{B} = \tilde{A} - \tilde{C}$	Yes. For any fuzzy variables $\tilde{\eta}, \tilde{\gamma}, \tilde{\zeta}$, $\tilde{\eta} = \tilde{\gamma} + \tilde{\zeta}$; Does imply $\tilde{\gamma} = \tilde{\eta} - \tilde{\zeta}$

Item	Zadeh's	Liu's
Link between two fuzzy theories	$\mu(x)=(2\mathrm{Cr}\{\xi=x\})\wedge 1$	$\mathrm{Cr}\{B\}=\frac{1}{2}(\mathrm{Poss}\{B\}-\mathrm{Nec}\{B\})$
Membership	Initially defined concept $\mu_A:\Theta\to[0,1]$	Secondarily defined concept $\mu(x)=(2\mathrm{Cr}\{\xi=x\})\wedge 1$

The next definition describes membership in terms of Liu's credibility measure, together with related theorems, a link between Zadeh's membership-initiated fuzzy mathematics and Liu's credibility measure-oriented fuzzy mathematics has been established in nature. The linkage definitely gives an intuitive understanding of Liu's credibility measure concept and also paves the way of applying credibility measure in practices, particularly, for those who are familiar with membership function concept.

Definition 2.5: (Liu, 2004, 2007) The (induced) membership function of a fuzzy variable ξ on $(\mathrm{Q},2^{\mathrm{Q}},\mathrm{Cr})$ is:

$$\mu(x)=(2\mathrm{Cr}\{\xi=x\})\wedge 1,\quad x\in\mathbf{R} \tag{4}$$

Conversely, for given membership function the credibility measure is determined by the credibility inversion theorem.

Theorem 2.6: (Liu, 2004, 2007) Let ξ be a fuzzy variable with membership function m, then for "B M$\mathbf{R}$,

$$\mathrm{Cr}\{\xi\in B\}=\frac{1}{2}\left(\sup_{x\in B}\mu(x)+1-\sup_{x\in B^c}\mu(x)\right),\ B\subset\mathbf{R} \tag{5}$$

As an example, if the set B is degenerated into a point x, then:

$$\mathrm{Cr}\{\xi=x\}=\frac{1}{2}\left(\mu(x)+1-\sup_{y\neq x}\mu(y)\right),\ \forall x\in\mathbf{R} \tag{6}$$

Theorem 2.7: (Liu, 2004, 2007) Let ξ be a fuzzy variable on $(\mathrm{Q},2^{\mathrm{Q}},\mathrm{Cr})$ with membership function μ. Then its credibility distribution,

$$\Lambda(x)=\frac{1}{2}\left(\sup_{y\le x}\mu(y)+1-\sup_{y>x}\mu(y)\right),\ \forall x\in\mathbf{R} \tag{7}$$

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 role of bridging Zadeh's fuzzy mathematics and the new axiomatic fuzzy theory and thus provides a conversion channel.

It is critical to emphasize again at the end of this section that different from Zehad's fuzzy set theory, the fuzzy variable on the Liu's credibility measure foundation is scalar real-valued function characterized by its credibility distribution. Therefore, the mathematical treatment of fuzzy variables on the platform is easier than that based on fuzzy sets as variable in Zadeh's fuzzy set theory.

3 AN *M*-ESTIMATOR 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:

$$Y=\alpha+\beta x+\varepsilon \tag{8}$$

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

$$\hat{Y} = a + bx \tag{9}$$

where vector (a,b) 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 random 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, $\varepsilon \sim N(0,\sigma^2)$ and ε_i and ε_j are uncorrelated. Thus, $Y_i \sim N\left(\alpha+\beta x_i,\sigma_i^2\right)$ and Y_i and Y_j are uncorrelated too.

From the credibilistic point of view, the universe should facilitate the error events. Mathematically

$$\Theta = \left\{\varepsilon_i \mid \varepsilon_i = Y_i - \left(\alpha + \beta x_i\right), i = 1,2,\cdots,n\right\} \tag{10}$$

Let $\tilde{R}$ be a fuzzy event defined on Θ, which connects to a fuzzy concept, {error is close to zero}, denoted by κ. A membership, denoted by $\mu_{\tilde{R}}(\varepsilon_i)$, represents the degree of belongingness to the fuzzy concept κ. Then the regression model fitting problem now becomes one of finding an empirical linear regression equation $\hat{Y} = a+bx$ based on the observations $\{(x_1,Y_1),(x_2,Y_2),\cdots,(x_n,Y_n)\}$ such that the membership $\mu_{\tilde{R}}(\varepsilon_i)$ is maximized.

Now, it is ready to investigate the forms of *M*-functions for regression coefficients estimation. Let the residual take a membership function of form

$$\mu_\varepsilon\left(Y_i-\left(a+bx_i\right)\right) = g\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right) \tag{11}$$

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

The basic idea underlying the searching the coefficients is to maximize the membership grade for any individual observation pair (x_i,Y_i). Then for all n pairs of observations $\{(x_1,Y_1),(x_2,Y_2),\cdots,(x_n,Y_n)\}$, The sum of the membership grades should be maximized. In mathematical language, the object function

$$J \equiv \sum_{i=1}^{n} \mu_\varepsilon\left(Y_i-\left(a+bx_i\right)\right) \tag{12}$$

It is obvious that for any individual observation Y_i the contribution to model goodness-of-fit is measured by the membership grade $\mu_\varepsilon(\cdot)$. The closer the observed value Y_i to the fitted value $\hat{Y}_i = \left(a+bx_i\right)$, the nearer to membership grade of the difference, i.e., the error, $\tilde{\varepsilon}_i = Y_i - \left(a+bx_i\right)$ to 1, which implies that the degree of fuzzy event $\tilde{\varepsilon}_i$ belonging to concept κ is high. Therefore, it is reasonable to use the sum of all the membership grades of the n observations $Y_1,Y_2,\cdots,Y_n$, for measuring the overall degree of belongingness of observations $Y_1,Y_2,\cdots,Y_n$. A typical membership function satisfying Equation (11) takes a normal form:

$$g(e_i) = \exp\left(-w\left(Y_i - (a+bx_i)\right)^2\right) \tag{13}$$

Now, we can define the *M*-functional equation system for fuzzy regression coefficients (a,b).

Definition 3.1: Given the differentiable membership function $g\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)$, 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

$$\begin{cases} \sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(Y_i-\left(a+bx_i\right)\right)=0 \\ \sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(Y_i-\left(a+bx_i\right)\right)x_i=0 \end{cases} \tag{14}$$

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

Theorem 3.2: Let a simple regression model $Y=\alpha+\beta x+\varepsilon$ assumes a fuzzy error ε $\mathrm{E}[\varepsilon]=0$ & $V[\varepsilon]=\sigma^2$ and membership function $g\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)$. For given n-pair 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 Equation (14). Furthermore, the M-estimator (a,b) takes a weighted least-square estimator form as

$$\begin{cases} b=\dfrac{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)\left(Y_i-\overline{Y}_h\right)}{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)^2} \\ a=\overline{Y}_h-\overline{b}x_h \end{cases} \tag{15}$$

where

$$\begin{aligned} \overline{x}_h &= \sum_{i=1}^{n} \frac{h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)} x_i \\ \overline{Y}_h &= \sum_{i=1}^{n} \frac{h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)} Y_i \end{aligned} \tag{16}$$

Theorem 3.2 is easy to prove by expanding the left side terms of Equation (14), re-arrange them, utilizing Equation (16) for obtaining Equation (15).
Let

$$\underline{Y}=\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix},\ X=\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix},\ D=\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \tag{17}$$

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

$$\underline{\Gamma}=\begin{bmatrix} a \\ b \end{bmatrix},\ W=D^{-1}=\begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix} \tag{18}$$

Then, the M-functional equation system Equation (14) can be re-written in a matrix form:

$$X^TW^{-1}X\underline{\Gamma}=X^TW^{-1}\underline{Y} \tag{19}$$

Equation (19) 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 Equation (14) will help further mathematical treatments. For example, Equation (15) can be re-expressed in matrix form:

$$\underline{\Gamma}=\left(X^TW^{-1}X\right)^{-1}X^TW^{-1}\underline{Y} \tag{20}$$

as long as the inverse matrix exists.

Example 3.3: Let then membership function g take a normal form:

$$\begin{aligned} & g\left(-\left(Y-\left(a+bx\right)\right)^2\right) \\ & =\exp\left(-w\left(Y-\left(a+bx\right)\right)^2\right) \end{aligned} \tag{21}$$

Then the derivative of membership function

$$\begin{aligned} & h\left(-w\left(Y-\left(a+bx\right)\right)^2\right) \\ & = \exp\left(-w\left(Y-\left(a+bx\right)\right)^2\right) \end{aligned} \tag{22}$$

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

$$w = \frac{2}{d_{\max} - d_{\min}} \tag{23}$$

where

$$\begin{aligned} d_{\max} &= \max_{i\in\{1,2,\cdots,n\}} \left\{\left(Y_i - \left(a+bx_i\right)\right)^2\right\} \\ d_{\min} &= \min_{i\in\{1,2,\cdots,n\}} \left\{\left(Y_i - \left(a+bx_i\right)\right)^2\right\} \end{aligned} \tag{24}$$

Then the M-estimators for regression coefficients are

$$\begin{cases} b = \dfrac{\sum_{i=1}^{n} \exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_\infty\right)\left(Y_i-\overline{Y}_\infty\right)}{\sum_{i=1}^{n} \exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_\infty\right)^2} \\ a = \overline{Y}_\infty - \overline{b}x_\infty \end{cases} \tag{25}$$

where

$$\begin{aligned} \overline{x}_\infty &= \sum_{i=1}^{n} \frac{\exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n}\exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)} x_i \\ \overline{Y}_\infty &= \sum_{i=1}^{n} \frac{\exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n}\exp\left(-w\left(Y_i-\left(a+bx_i\right)\right)^2\right)} Y_i \end{aligned} \tag{26}$$

4 PROPERTIES OF THE M-ESTIMATORS

It should be emphasized at the beginning of this section that the issue of properties of estimator of fuzzy regression coefficients was not really deeply explored. Furthermore, it should be also fully aware that the issue of variance-covariance structure of regression coefficient estimators in statistical linear model is a standard exercise, however, in fuzzy regression developed so far, most of the modeling exercises stopped at obtaining the estimators of regression coefficients. These difficulties rooted in the set-level variable treatments in Zadeh's fuzzy set theory.

Lemma 4.1: *M*-Estimator b is a linear function of observations $\{Y_1, Y_2, \cdots, Y_n\}$. In other words,

$$b = \sum_{i=1}^{n} \kappa_i Y_i \tag{27}$$

where

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

The proof of Lemma 4.1 is a straightforward manipulation of Equation (18).

Theorem 4.2: *M*-Estimator for the regression coefficients (α,β), denoted as (a,b), are (conditionally) unbiased. In other words,

$$\mathrm{E}[a] = \alpha, \quad \mathrm{E}[b] = \beta \tag{29}$$

The proof is a straightforward task by noticing that $\sum_{i=1}^{n} k_i = 0$ and $\sum_{i=1}^{n} k_i = 1$. Furthermore, we should notice that k_i is conditionally on the roots of *M*-function equation system.

Theorem 4.3: The estimated variance-covariance matrix of the regression coefficient *M*-estimators is given by

$$\hat{V}_\Theta[\underline{\Gamma}] = s^2\left(X^T W^{-1} X\right)^{-1} X^T V[\underline{Y}] W^{-1} X\left[\left(X^T W^{-1} X\right)^{-1}\right]^T \tag{30}$$

where

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(Y_i-\left(a+bx_i\right)\right)^2 \tag{31}$$

The proof is a matrix manipulation.
Let

$$\hat{V}_\Theta[\underline{\Gamma}] = \begin{bmatrix} \hat{\sigma}^2(\hat{\alpha}) & \hat{\sigma}\left(\hat{\alpha},\hat{\beta}\right) \\ \hat{\sigma}\left(\hat{\alpha},\hat{\beta}\right) & \hat{\sigma}^2\left(\hat{\beta}\right) \end{bmatrix} \tag{32}$$

Corollary 4.4: The estimated variances for the regression coefficient *M*-estimators a and b respectively are

$$\begin{aligned}\hat{\sigma}^2(a) &= s^2\left[\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\right]^2 \\ &\times\frac{\sum_{i=1}^{n} h^2\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(\frac{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i^2}{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i}-x_i\right)^2}{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)^2}\end{aligned} \tag{33}$$

and

$$\hat{\sigma}^2(b) = s^2\frac{\sum_{i=1}^{n} h^2\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)^2}{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)^2} \tag{34}$$

Furthermore, the correlation between the regression coefficient *M*-estimators a and b is

$$\begin{aligned}\hat{\sigma}(a,b) &= \frac{s^2}{\left[\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_g\right)^2\right]^2}\times \\ &\left[-\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i^2\right)\times\left(\sum_{i=1}^{n} h^2\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\right)\times\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i\right)\right. \\ &+\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i\right)^2\times\left(\sum_{i=1}^{n} h^2\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i\right) \\ &+\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i^2\right)\times\left(\sum_{i=1}^{n} h^2\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i\right)\times\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\right) \\ &\left.-\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i\right)\times\left(\sum_{i=1}^{n} h^2\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)x_i^2\right)\times\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\right)\right]\end{aligned} \tag{35}$$

Theorem 4.5: The estimated correlation coefficient for regression model $Y = \alpha + \beta x + \varepsilon$ is

$$r = \frac{\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)\left(Y_i-\overline{Y}_h\right)}{\sqrt{\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)^2\right)\times\left(\sum_{i=1}^{n} h\left(-\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(Y_i-\overline{Y}_h\right)^2\right)}} \tag{36}$$

Note here that accurately, we should say the correlation coefficient between X and Y. However, in the fuzzy regression model assumptions x_i are clearly assumed to be certain real-valued number. Therefore, in Theorem 4.5, which gives the formula for r, r reveals the association between random variate Y and exploratory variable x and thus is regarded as an inherent index to the regression.

5 ASYMPTOTIC MEMBERSHIP FOR *M*-ESTIMATOR

It would be difficult task to discuss the asymptotic membership for coefficient estimators if we assume the membership function $g\left(-\left(Y-(\alpha+\beta x)\right)^2\right)$ takes very general form. Nevertheless, if the normal membership function is assumed, then the discussions will be slightly simplified.

Theorem 5.1: Let membership function for residual error takes normal form

$$\mu(e)=\exp\left(-w\left(Y-(\alpha+\beta x)\right)^2\right) \tag{37}$$

Then the asymptotic membership function for $(a-\alpha)$ is

$$\mu_{(a-\alpha)}(v)\to\exp\left(-w\left(\frac{v^2}{\sqrt{\hat{\sigma}^2(a)}}\right)\right) \tag{38}$$

and the asymptotic membership function for $(b-\beta)$ is

$$\mu_{(b-\beta)}(u)\to\exp\left(-w\left(\frac{u^2}{\sqrt{\hat{\sigma}^2(b)}}\right)\right) \tag{39}$$

where $\sqrt{\hat{\sigma}^2(a)}$ and $\sqrt{\hat{\sigma}^2(b)}$ are given in Theorem 4.4 respectively.

Theorem 5.2: The asymptotic joint membership function for vector

$$\begin{pmatrix} a-\alpha \\ b-\beta \end{pmatrix} \tag{40}$$

is

$$\mu_{\begin{pmatrix} a-\alpha \\ b-\beta \end{pmatrix}}(u,v)\to\exp\left(-w\begin{pmatrix} u \\ v \end{pmatrix}^T\left[\hat{V}\begin{pmatrix} a \\ b \end{pmatrix}\right]^{-1}\begin{pmatrix} u \\ v \end{pmatrix}\right) \tag{41}$$

Due to the feature that a normal form membership function only requires first two moments (mean and variance), therefore, utilizing the mean-variance information from Section 4 it is reasonable to establish the asymptotic membership function for $a-\alpha$, $b-\beta$, and their bivariate joint asymptotic membership function. Once an asymptotic membership function is found the asymptotic credibility distribution can be easily derived in terms of Equation (7).

6 AN EXTENSION TO MULTI-MEMBERSHIP ERRORS

The *M*-estimation to the simple fuzzy regression lies on assuming that the errors have the same credibility distribution, or equivalently, the same membership function. However, we can extend our treatments into the case that error terms comes from multiple credibility distributions. Let a sample be $\{(x_1,Y_1),(x_2,Y_2),\cdots,(x_n,Y_n)\}$, which is assumed from multiple membership functions, i.e.,

$$\mu_i(e_i)=g\left(-w_i\left(Y_i-(\alpha+\beta x_i)\right)^2\right) \tag{42}$$

Then, the optimization criterion is still to maximize the total membership grades:

$$J=\sum_{i=1}^{n} g\left(-w_i\left(Y_i-(\alpha+\beta x_i)\right)^2\right) \tag{43}$$

and the *M*-functional equation system

$$\begin{cases} \sum_{i=1}^{n} w_i h\left(-w_i\left(Y-\left(a+bx_i\right)\right)^2\right)\left(Y-\left(a+bx_i\right)\right)=0 \\ \sum_{i=1}^{n} w_i h\left(-w_i\left(Y-\left(a+bx_i\right)\right)^2\right)\left(Y-\left(a+bx_i\right)\right)x_i=0 \end{cases} \tag{44}$$

Theorem 6.1: 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 Equation (43). Furthermore, the *M*-estimator (a,b) takes a weighted least-square estimator form as

$$\begin{cases} b=\dfrac{\sum_{i=1}^{n} h\left(-w_i\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)\left(Y_i-\overline{Y}_h\right)}{\sum_{i=1}^{n} h\left(-w_i\left(Y_i-\left(a+bx_i\right)\right)^2\right)\left(x_i-\overline{x}_h\right)^2} \\ a=\overline{Y}_h-\overline{b}x_h \end{cases} \tag{45}$$

where

$$\begin{aligned} \overline{x}_h &= \sum_{i=1}^{n} \frac{h\left(-w_i\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n} h\left(-w_i\left(Y_i-\left(a+bx_i\right)\right)^2\right)} x_i \\ \overline{Y}_h &= \sum_{i=1}^{n} \frac{h\left(-w_i\left(Y_i-\left(a+bx_i\right)\right)^2\right)}{\sum_{i=1}^{n} h\left(-w_i\left(Y_i-\left(a+bx_i\right)\right)^2\right)} Y_i \end{aligned} \tag{46}$$

However, we should notice that the matrix form of *M*-functional equation system, i.e., the weighted least-squares formed normal equation will be similar but we need take care of the variance-covariance matrix since the model assumption is changed to,

$$\mathrm{E}\left(e_i\right)=0 \ \& \ V\left(e_i\right)=\sigma_i^2,\ \mathrm{E}\left(e_i e_j\right)=0 \tag{47}$$

The remaining investigations can be carried on in a similar way but the error variance-covariance matrix needs care.

7 AN EXTENSION TO MULTIPLE FUZZY REGRESSION

Let us assume that the response variable Y functionally related to *p* exploratory variables, $x_1,x_2,\ldots,x_p$. Let $\left\{\left(Y_i,x_{1i},\ldots,x_{pi}\right),\ i=1,\ldots,n\right\}$ be the sample observations have model assumptions:

$$\begin{aligned} &\mathrm{E}\left(Y_i-\sum_{k=0}^{p} b x_{ik}\right)=0 \\ &V\left(Y_i-\sum_{k=0}^{p} b x_{ik}\right)=\sigma_i^2 \\ &\mathrm{E}\left(\left(Y_i-\sum_{k=0}^{p} b x_{ik}\right)\left(Y_j-\sum_{k=0}^{p} b x_{jk}\right)\right)=0 \end{aligned} \tag{48}$$

Then the optimization criterion is

$$\mathrm{J}=\sum_{i=1}^{n} g\left(-w_i\left(Y_i-\sum_{k=1}^{p} b_k x_{ki}\right)^2\right) \tag{49}$$

Let $\underline{\Gamma}=\left(\beta_0,\beta_1,\ldots,\beta_p\right)^T$, then $\hat{\underline{\Gamma}}=\left(\hat{\beta}_0,\hat{\beta}_1,\ldots,\hat{\beta}_p\right)^T$ denotes the *M*-estimator of regression coefficients. The *M*-functional equation system is thus

$$\begin{cases} \sum_{i=1}^{n} w_i h\left(-w_i\left(Y_i-\sum_{k=1}^{p}\beta_k x_{ki}\right)^2\right)\left(Y_i-\sum_{k=1}^{p}\beta_k x_{ki}\right)=0 \\ \sum_{i=1}^{n} w_i h\left(-w_i\left(Y_i-\sum_{k=1}^{p}\beta_k x_{ki}\right)^2\right)\left(Y_i-\sum_{k=1}^{p}\beta_k x_{ki}\right)x_{1i}=0 \\ \vdots \\ \sum_{i=1}^{n} w_i h\left(-w_i\left(-w_i\left(Y_i-\sum_{k=1}^{p}\beta_k x_{ki}\right)^2\right)^2\right)\left(Y_i-\sum_{k=1}^{p}\beta_k x_{ki}\right)x_{pi}=0 \end{cases} \tag{50}$$

Notice that the error variance-covariance matrix

$$V=\begin{bmatrix} \sigma_1^2 & 0 & \vdots & 0 \\ 0 & \sigma_1^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \tag{51}$$

The matrix form – weighted least-squares normal equation is

$$X^T W^{-1} X\underline{\Gamma} = X^T W^{-1}\underline{Y} \tag{52}$$

where

$$\underline{Y}=\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix},\ X=\begin{bmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{bmatrix},\ D=\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \tag{53}$$

and

$$W^{-1}=\mathrm{diag}\left(\sigma_i^2 h\left(-w_i\left(Y_i-\sum_{k=1}^{p}\beta_k x_{ki}\right)^2\right)\right) \tag{54}$$

Then the *M*-estimator for the multiple regression coefficients are

$$\hat{\underline{\Gamma}}=\left(X^T W^{-1} X\right)^{-1} X^T W^{-1}\underline{Y} \tag{55}$$

as long as the inverse matrix exists.

The properties and the variance-covariance structure of $\hat{\underline{\Gamma}}$ could be investigated as that of weighted least-squares regression formality.

8 CONCLUSION

The major advantage of this paper is its scalar variable treatments of fuzzy observations because fuzzy variable concept is established on the credibility measure foundation. Therefore, we are able to propose an *M*-estimation approach for simple fuzzy regression model. The optimization criterion is minimizing the fuzzy uncertainty by seeking the fitted errors with the membership grades as large as possible. In this sense, our simple fuzzy regression has the similar optimization criterion as Tanaka et al. (1980, 1982). We notice that the *M*-functional equation system can be re-written in weighted least-squares normal equation formation, which enables heavily to borrow arguments similar to statistical linear model theory. Finally, it is necessary to point out that in similar manner, the *M*-functional equation system could be defined for fuzzy multivariate regression model.

9 ACKNOWLEDGEMENT

This research is partially supported by South African National Research Foundation Grant FA2006042800024.

10 REFERENCES

Diamond, P. (1987). Least squares fitting of several fuzzy variables, in: J.C. Bezdek (Ed.), *Analysis of Fuzzy Information*, CRC Press, Tokyo, 329-331.
Diamond, P. (1988). Fuzzy least squares, Inform. Sci. 46, 141-157.
Dubois, D. and Prade, H. (1980). *Theory and Applications, Fuzzy Sets and Systems*. New York: Academic Press.
Kacprzyk, J. and Fedrizzi, M. (1992). *Fuzzy Regression Analysis*. Physica-Verlag, Heidelberg.
Kaufmann, A. (1975). *Introduction to the Theory of Fuzzy Subsets*; Vol. I. New York: Academic Press.
Liu, B.D. (2004). *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*. Berlin: Springer-Verlag Heidelberg.
Liu, B.D. (2007). *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*, 2^{nd} Edition. Berlin: Springer-Verlag Heidelberg.
Tanaka, H., Uejima, S. and Asia, K. (1980). Fuzzy linear regression model. *Internat. Congress on Applied Systems Research and Cybernetics*, Aculpoco, Mexico.
Tanaka, H., Uejima, S. and Asia, K. (1982). Linear regression analysis with Fuzzy model, *IEEE Trans. Systems Man. Cybernet*. 12 (6): 903–907.
Zadeh, L.A. (1965). Fuzzy Sets. *Information and Control*; Vol. 8: 338-353.
Zadeh, L.A. (1978). Fuzzy sets as basis for a theory of possibility. *Fuzzy Sets and Systems*; Vol. 1: 3-28.