

# A Formula for Fuzzy Linear Regression Analysis

Chi-Tsuen Yeh

Department of Applied Mathematics, National University of Tainan  
33, Sec. 2, Shu-Lin St., 70005, Tainan, Taiwan  
Email: ctyeh@mail.nutn.edu.tw

**Abstract**—The purpose of this paper is to deal with the problem of least-squares multiple regression with fuzzy data. The constant regression coefficient is assumed to be symmetric triangular, and other coefficients are real (crisp). By applying symmetric triangular approximations of fuzzy numbers, a new method for computing the regression coefficients is proposed. The new method is efficient and easy to determine the coefficients.

**Index Terms**—fuzzy linear regression, symmetric triangular approximation, least-squares estimates.

## I. INTRODUCTION

An arbitrary fuzzy number  $\tilde{A}$  can be represented by an ordered pair of left continuous functions  $[A^l(\alpha), A^u(\alpha)]$ ,  $0 \leq \alpha \leq 1$ , which both satisfy the following conditions:

- 1)  $A^l$  is increasing on  $[0,1]$ ,
- 2)  $A^u$  is decreasing on  $[0,1]$ ,
- 3)  $A^l(1) \leq A^u(1)$ .

Let  $\mathbb{R}$  and  $F(\mathbb{R})$  denote the sets of all real numbers and fuzzy numbers, respectively. In this paper, we always place a  $\sim$  over a letter to denote a fuzzy number, for example  $\tilde{X}_i, \tilde{A} \in F(\mathbb{R})$  and  $a, \tau, \sigma \in \mathbb{R}$ . Let  $\tilde{A}$  and  $\tilde{B} \in F(\mathbb{R})$ . The Euclidean distance between  $\tilde{A}$  and  $\tilde{B}$  is defined as

$$d(\tilde{A}, \tilde{B}) = \left[ \int_0^1 (|A^l(\alpha) - B^l(\alpha)|^2 + |A^u(\alpha) - B^u(\alpha)|^2) d\alpha \right]^{\frac{1}{2}}.$$

Given fuzzy vectors  $\vec{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)$  and  $\vec{X}_i = (\tilde{X}_{i1}, \dots, \tilde{X}_{in}) \in F(\mathbb{R})^n$ ,  $1 \leq i \leq m$ , the problem of fuzzy linear regression analysis is to find the regression coefficients  $\tilde{A}_i \in F(\mathbb{R})$ ,  $0 \leq i \leq m$ , which minimize the observation error

$$d^2(\vec{Y}, \tilde{A}_0 \vec{1} + \tilde{A}_1 \vec{X}_1 + \dots + \tilde{A}_m \vec{X}_m) := \sum_{j=1}^n d^2(\tilde{Y}_j, \tilde{A}_0 + \tilde{A}_1 \tilde{X}_{1j} + \dots + \tilde{A}_m \tilde{X}_{mj}), \quad (1)$$

where  $\vec{1} := (1, \dots, 1)$ .

In 1982, Tanaka et al. [21] first examined fuzzy regression analysis with real independent (or explanatory) and fuzzy dependent (or response) variables, in which the regression coefficients are assumed to be symmetric triangular fuzzy numbers. Afterwards, many scholars studied this type fuzzy regression analysis by generalizing the coefficients to more general fuzzy numbers [2], [6], [8], [9], [10], [14], [16], [18], [19], [22]. In 1992, Sakawa and Yano [20] considered fuzzy linear regression models with fuzzy data, that means the independent variables and dependent variable are all fuzzy numbers. Subject to the regression coefficients are all real, the least-squares estimation was further studied by Diamond [4], and the existence was

proved by Kratschmer [17]. In 1992, Bardosy et al. [11] presented normal equations systems depended on signs of regression coefficients to estimate the regression coefficients. And, in 2002 Kao and Chyu [12] proposed a two-stage method. Most of the recent approaches have to consider optimization cones, then determine the regression coefficients by comparing observation errors of these possible estimators. It will take a lot of time to find the best one. An iterative algorithm (numerical method) was proposed by Bargiela et al. [1]. In 2009, Yeh [27] showed that the multiple regression problem can be replaced with a 0-1 programming problem. Furthermore, he proposed a formula for computing the regression coefficients in one-variable models (see the following Theorem 5). For more general case that the constant coefficient is fuzzy, Diamond and Korner [5] generalized their previous results via Hukuhara difference. Wunsche and Nather [23] turned to study on the restriction that the data are of normal distributions. Chen and Hsueh [3] used the least-squares method to determine regression coefficients in multiple-variable models. There are also many other contributions to this type [2], [8], [9], [13], [15], [18].

This paper is organized as follows. In Section 2, we introduce the mean of fuzzy numbers. In Section 3, we present an formula for solving the regression model whose regression coefficients are all real. The present formula is simple and clear. In Section 4, we completely solve the regression problem whose constant coefficient is assumed to be symmetric triangular. In Section 5, we work on the example designed by Sakawa and Yano. It is shown that our estimator is better than the previous.

## II. THE MEAN OF FUZZY NUMBERS

Let  $\tilde{Z}_i$ ,  $1 \leq i \leq n$  be fuzzy numbers. We define the *mean* of  $\tilde{Z}_i$  as follows

$$E(\tilde{Z}) = \frac{1}{n} \sum_{i=1}^n \tilde{Z}_i. \quad (2)$$

**Theorem 1.** Let  $\tilde{Z}_i$ ,  $1 \leq i \leq n$  be fuzzy numbers. Then,

$$\sum_{i=1}^n d^2(\tilde{Z}_i, \tilde{A}) = \sum_{i=1}^n d^2(\tilde{Z}_i, E(\tilde{Z})) + nd^2(E(\tilde{Z}), \tilde{A}),$$

for each  $\tilde{A} \in F(\mathbb{R})$ .

Let  $\tilde{A} = [A^l(\alpha), A^u(\alpha)]$  and  $\tilde{B} = [B^l(\alpha), B^u(\alpha)] \in F(\mathbb{R})$ . The *vector difference* between  $\tilde{A}$  and  $\tilde{B}$  is described as follows

$$\tilde{A} \ominus \tilde{B} := [A^l(\alpha) - B^l(\alpha), A^u(\alpha) - B^u(\alpha)].$$

The vector difference between  $\tilde{A}$  and  $\tilde{B}$  may not belong to  $F(\mathbb{R})$ . For example,

$$[\alpha, 2 - \alpha] \ominus [2\alpha, 2] = [-\alpha, -\alpha] \notin F(\mathbb{R}).$$

Notice that, we can still compute the distance  $d(\tilde{A}, \tilde{B})$ , defined in Section 1, even if  $\tilde{A} = [A^l(\alpha), A^u(\alpha)]$  and  $\tilde{B} = [B^l(\alpha), B^u(\alpha)] \notin F(\mathbb{R})$ .

**Theorem 2.** Let  $\vec{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)$  and  $\vec{X} = (\tilde{X}_1, \dots, \tilde{X}_n) \in F(\mathbb{R})^n$  be fuzzy vectors. Then,

$$d^2(\vec{Y}, \tilde{A}\vec{1} + b\vec{X}) = \sum_{i=1}^n d^2(\tilde{Y}_i \ominus E(\tilde{Y}), b\tilde{X}_i \ominus bE(\tilde{X})) + nd^2(E(\tilde{Y}) \ominus bE(\tilde{X}), \tilde{A}),$$

for each  $\tilde{A} \in F(\mathbb{R})$  and each  $b \in \mathbb{R}$ .

### III. A FORMULA FOR COMPUTING REAL REGRESSION COEFFICIENTS

In this section, we are going to study fuzzy regression analysis with real regression coefficients, that is to find  $a, b \in \mathbb{R}$  minimizing the observation error

$$d^2(\vec{Y}, a\vec{1} + b\vec{X}),$$

where  $\vec{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)$  and  $\vec{X} = (\tilde{X}_1, \dots, \tilde{X}_n)$  are given data.

Let  $\tilde{A} \in F(\mathbb{R})$  and  $\Omega$  be a closed convex subset of  $F(\mathbb{R})$ . Then, an element  $P_\Omega(\tilde{A}) \in \Omega$  is called the *best approximation* of  $\tilde{A}$  in  $\Omega$  if

$$d(\tilde{A}, P_\Omega(\tilde{A})) \leq d(\tilde{A}, \tilde{X}), \quad \forall \tilde{X} \in \Omega.$$

If we set  $\Omega = \mathbb{R}$  (i.e. each  $\tilde{A} \in \Omega$  is real), then

$$P_{\mathbb{R}}(\tilde{A}) := \int_0^1 A^c(\alpha) d\alpha.$$

**Proposition 3.** Let  $\tilde{A}, \tilde{B} \in F(\mathbb{R})$  and  $a, b \in \mathbb{R}$ . Then,

$$P_{\mathbb{R}}(a\tilde{A} + b\tilde{B}) = aP_{\mathbb{R}}(\tilde{A}) + bP_{\mathbb{R}}(\tilde{B}),$$

$$P_{\mathbb{R}}(a\tilde{A} \ominus b\tilde{B}) = aP_{\mathbb{R}}(\tilde{A}) - bP_{\mathbb{R}}(\tilde{B}),$$

and

$$d^2(\tilde{A}, a) = d^2(\tilde{A}, P_{\mathbb{R}}(\tilde{A})) + |P_{\mathbb{R}}(\tilde{A}) - a|^2.$$

From Theorem 2 we find

$$d^2(\vec{Y}, a\vec{1} + b\vec{X}) = \sum_{i=1}^n d^2(\tilde{Y}_i \ominus E(\tilde{Y}), b\tilde{X}_i \ominus bE(\tilde{X})) + nd^2(E(\tilde{Y}) \ominus bE(\tilde{X}), a).$$

Consequently, by Proposition 3 we have

$$\begin{aligned} & d^2(E(\tilde{Y}) \ominus bE(\tilde{X}), a) \\ & \geq d^2(E(\tilde{Y}) \ominus bE(\tilde{X}), P_{\mathbb{R}}(E(\tilde{Y}) - bP_{\mathbb{R}}(E(\tilde{X}))). \end{aligned}$$

Hence, to minimize  $d^2(\vec{Y}, a\vec{1} + b\vec{X})$  we must choose that

$$a = P_{\mathbb{R}}(E(\tilde{Y}) - bP_{\mathbb{R}}(E(\tilde{X}))),$$

and then we have

$$\begin{aligned} & d^2(\vec{Y}, a\vec{1} + b\vec{X}) \\ & = d^2(\vec{Y}, [P_{\mathbb{R}}(E(\tilde{Y})) - bP_{\mathbb{R}}(E(\tilde{X}))]\vec{1} + b\vec{X}) \\ & = \sum_{i=1}^n d^2(\tilde{Y}_i \ominus P_{\mathbb{R}}(E(\tilde{Y})), b\tilde{X}_i \ominus bP_{\mathbb{R}}(E(\tilde{X}))). \end{aligned}$$

Recall that every closed interval can be represented by its center and radius (half width). For instance, the center of  $[A^l(\alpha), A^u(\alpha)]$  is

$$A^c(\alpha) := \frac{A^l(\alpha) + A^u(\alpha)}{2}, \quad (3)$$

and the radius of  $[A^l(\alpha), A^u(\alpha)]$  is

$$A^r(\alpha) := \frac{A^u(\alpha) - A^l(\alpha)}{2}. \quad (4)$$

That means we can use the symbol

$$[A^c(\alpha), A^r(\alpha)]_{CR}$$

to rewrite the fuzzy number  $\tilde{A} = [A^l(\alpha), A^u(\alpha)]$ , which is called the *CR-representation* of  $\tilde{A}$ , see [24].

**Proposition 4** (Proposition 1, [27]). Let

$$\tilde{A} = [A^c(\alpha), A^r(\alpha)]_{CR} \quad \text{and} \quad \tilde{B} = [B^c(\alpha), B^r(\alpha)]_{CR}$$

be fuzzy numbers with CR-representations, and  $a \in \mathbb{R}$ . Then the addition and scalar multiplication of fuzzy numbers can be described as follows:

$$\tilde{A} + \tilde{B} = [A^c(\alpha) + B^c(\alpha), A^r(\alpha) + B^r(\alpha)]_{CR}$$

and

$$a\tilde{A} = [aA^c(\alpha), |a|A^r(\alpha)]_{CR},$$

respectively. Moreover,

$$\left(\sum_i a_i \tilde{X}_i\right)^c(\alpha) = \sum_i a_i \tilde{X}_i^c(\alpha)$$

and

$$\left(\sum_i a_i \tilde{X}_i\right)^r(\alpha) = \sum_i |a_i| \tilde{X}_i^r(\alpha),$$

for all  $\tilde{X}_i \in F(\mathbb{R})$  and  $a_i \in \mathbb{R}$ .

For each fuzzy vector  $\vec{Z} = (\tilde{Z}_1, \dots, \tilde{Z}_n)$ , where

$$\tilde{Z}_j = [Z_j^c(\alpha), Z_j^r(\alpha)]_{CR}, \quad 1 \leq j \leq n,$$

are represented with CR-representations, define two real functional vectors as follows:

$$\vec{Z}^c := (Z_1^c(\alpha), \dots, Z_n^c(\alpha))$$

and

$$\vec{Z}^r := (Z_1^r(\alpha), \dots, Z_n^r(\alpha)).$$

Recall that, for any two real vector functions  $\vec{f} = (f_1, \dots, f_n)$  and  $\vec{g} = (g_1, \dots, g_n)$  the inner product of  $\vec{f}$  and  $\vec{g}$  is

$$\langle \vec{f}, \vec{g} \rangle := \sum_{i=1}^n \int_0^1 f_i(x) g_i(x) dx. \quad (5)$$

Now, let  $\vec{X} = (\tilde{X}_1, \dots, \tilde{X}_n)$  and  $\vec{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n) \in F(\mathbb{R})^n$ . Define

$$\begin{aligned} S_{xx}^c &:= \langle \vec{X}^c - P_{\mathbb{R}}(E(\tilde{X}))\vec{1}, \vec{X}^c - P_{\mathbb{R}}(E(\tilde{X}))\vec{1} \rangle, \\ S_{xy}^c &:= \langle \vec{X}^c - P_{\mathbb{R}}(E(\tilde{X}))\vec{1}, \vec{Y}^c - P_{\mathbb{R}}(E(\tilde{Y}))\vec{1} \rangle, \\ S_{xx}^r &:= \langle \vec{X}^r, \vec{X}^r \rangle, \\ S_{xy}^r &:= \langle \vec{X}^r, \vec{Y}^r \rangle. \end{aligned}$$

It is easy to verify  $S_{xx}^c \geq 0$ ,  $S_{xx}^r \geq 0$ , and  $S_{xy}^r \geq 0$ .

**Theorem 5** (Theorem 6, [27]). *Let  $\vec{Y}$  and  $\vec{X} \in F(\mathbb{R})^n$ , and let  $(-1)^s$  denote the sign of  $S_{xy}^c$ . If  $S_{xx}^c + S_{xx}^r \neq 0$ , then the regression coefficients can be computed by the following formula*

$$\begin{aligned} b &= (-1)^s (|S_{xy}^c| + S_{xy}^r) (S_{xx}^c + S_{xx}^r)^{-1}, \\ a &= P_{\mathbb{R}}(E(\tilde{Y})) - bP_{\mathbb{R}}(E(\tilde{X})). \end{aligned} \quad (6) \quad (7)$$

#### IV. A THEOREM FOR MORE GENERAL REGRESSION MODEL

Let  $\Delta$  denote the set of all symmetric triangular fuzzy numbers. In this section, we are going to study fuzzy regression model whose constant regression coefficient is symmetric triangular, that is to find  $\tilde{A} \in \Delta$  and  $b \in \mathbb{R}$  minimizing the observation error

$$d^2(\vec{Y}, \tilde{A}\vec{1} + b\vec{X}).$$

For each  $\tilde{A} \in F(\mathbb{R})$ , an element  $P_{\Delta}(\tilde{A}) \in \Delta$  is called the symmetric triangular approximation of  $\tilde{A}$  if

$$d(\tilde{A}, P_{\Delta}(\tilde{A})) \leq d(\tilde{A}, \tilde{X}), \quad \forall \tilde{X} \in \Delta.$$

In [28], the author proved that:

**Proposition 6.** *Let  $\tilde{A} \in F(\mathbb{R})$ . Then, its the symmetric triangular approximation  $P_{\Delta}(\tilde{A}) = [\tau, \sigma]_{CR}$  can be computed by the following formula*

$$\tau = \tau(\tilde{A}) := \int_0^1 A^c(\alpha) d\alpha, \quad (8)$$

$$\sigma = \sigma(\tilde{A}) := 3 \int_0^1 A^r(\alpha)(1 - \alpha) d\alpha. \quad (9)$$

It is easy to verify that

$$\tau(a\tilde{A}) = a\tau(\tilde{A})$$

$$\sigma(a\tilde{A}) = |a|\sigma(\tilde{A})$$

for all  $\tilde{A} \in F(\mathbb{R})$  and  $a \in \mathbb{R}$ . For more general approximations of fuzzy numbers, we refer to [7], [25], [26], [29], [30].

**Proposition 7.** *Let  $\tilde{A}, \tilde{B} \in F(\mathbb{R})$  and  $a, b \in \mathbb{R}$ . Then,*

$$P_{\Delta}(a\tilde{A} + b\tilde{B}) = aP_{\Delta}(\tilde{A}) + bP_{\Delta}(\tilde{B}),$$

$$P_{\Delta}(a\tilde{A} \ominus b\tilde{B}) = aP_{\Delta}(\tilde{A}) \ominus bP_{\Delta}(\tilde{B}). \quad (10)$$

Furthermore,

$$d^2(\tilde{A}, \tilde{X}) = d^2(\tilde{A}, P_{\Delta}(\tilde{A})) + d^2(P_{\Delta}(\tilde{A}), \tilde{X}), \quad \forall \tilde{X} \in \Delta. \quad (11)$$

Let  $\tilde{A} \in \Delta$ . From Theorem 2 we find

$$\begin{aligned} d^2(\vec{Y}, \tilde{A}\vec{1} + b\vec{X}) &= \\ \sum_{i=1}^n d^2(\tilde{Y}_i \ominus E(\tilde{Y}), b\tilde{X}_i \ominus bE(\tilde{X})) + nd^2(E(\tilde{Y}) \ominus bE(\tilde{X}), \tilde{A}). \end{aligned}$$

Consequently, by Proposition 7 we have

$$\begin{aligned} &d^2(E(\tilde{Y}) \ominus bE(\tilde{X}), \tilde{A}) \quad (\text{by Eq.(10)}) \\ &\geq d^2(E(\tilde{Y}) \ominus bE(\tilde{X}), P_{\Delta}(E(\tilde{Y}) \ominus bE(\tilde{X}))) \quad (\text{by Eq.(11)}) \\ &= d^2(E(\tilde{Y}) \ominus bE(\tilde{X}), P_{\Delta}(E(\tilde{Y})) \ominus bP_{\Delta}(E(\tilde{X}))). \end{aligned}$$

Hence, if  $P_{\Delta}(E(\tilde{Y})) \ominus bP_{\Delta}(E(\tilde{X})) \in \Delta$ , we only choose

$$\tilde{A} = P_{\Delta}(E(\tilde{Y})) \ominus bP_{\Delta}(E(\tilde{X}))$$

and then we have

$$\begin{aligned} &d^2(\vec{Y}, \tilde{A}\vec{1} + b\vec{X}) \\ &= \sum_{i=1}^n d^2(\tilde{Y}_i \ominus P_{\Delta}(E(\tilde{Y})), b\tilde{X}_i \ominus bP_{\Delta}(E(\tilde{X}))). \end{aligned} \quad (12)$$

Thus, the rest is to find  $b \in \mathbb{R}$  which minimizes the above value (i.e. Eq.(12)). But in general case, it may happen

$$P_{\Delta}(E(\tilde{Y})) \ominus bP_{\Delta}(E(\tilde{X})) \notin \Delta.$$

Note that, Proposition 7 implies

$$\begin{aligned} P_{\Delta}(E(\tilde{Y})) \ominus bP_{\Delta}(E(\tilde{X})) &= [\tau(E(\tilde{Y})) - b\tau(E(\tilde{X})), \\ &(\sigma(E(\tilde{Y})) - |b|\sigma(E(\tilde{X}))(1 - \alpha))]_{CR} \end{aligned} \quad (13)$$

Hence, we obtain  $P_{\Delta}(E(\tilde{Y})) \ominus bP_{\Delta}(E(\tilde{X})) \notin \Delta$  if and only if

$$\sigma(E(\tilde{Y})) < |b|\sigma(E(\tilde{X})).$$

Let  $\vec{X} = (\tilde{X}_1, \dots, \tilde{X}_n)$  and  $\vec{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)$ . Define

$$\begin{aligned} T_{xx}^c &:= \langle \vec{X}^c - \tau(E(\tilde{X}))\vec{1}, \vec{X}^c - \tau(E(\tilde{X}))\vec{1} \rangle, \\ T_{xy}^c &:= \langle \vec{X}^c - \tau(E(\tilde{X}))\vec{1}, \vec{Y}^c - \tau(E(\tilde{Y}))\vec{1} \rangle, \\ T_{xx}^r &:= \langle \vec{X}^r - \sigma(E(\tilde{X}))(1 - \alpha)\vec{1}, \vec{X}^r - \sigma(E(\tilde{X}))(1 - \alpha)\vec{1} \rangle, \\ T_{xy}^r &:= \langle \vec{X}^r - \sigma(E(\tilde{X}))(1 - \alpha)\vec{1}, \vec{Y}^r - \sigma(E(\tilde{Y}))(1 - \alpha)\vec{1} \rangle. \end{aligned}$$

Note that, by Eq.(5), (8), and (9) it is easy to verify that

$$\begin{aligned} \langle \vec{X}^c, \tau(E(\tilde{X}))\vec{1} \rangle &= n\tau(E(\tilde{X}))^2, \\ \langle \vec{X}^c, \tau(E(\tilde{Y}))\vec{1} \rangle &= \langle \vec{Y}^c, \tau(E(\tilde{X}))\vec{1} \rangle = n\tau(E(\tilde{X}))\tau(E(\tilde{Y})), \\ \langle \vec{X}^r, \sigma(E(\tilde{X}))(1 - \alpha)\vec{1} \rangle &= \frac{n}{3}\sigma(E(\tilde{X}))^2, \\ \langle \vec{X}^r, \sigma(E(\tilde{Y}))(1 - \alpha)\vec{1} \rangle &= \langle \vec{Y}^r, \sigma(E(\tilde{X}))(1 - \alpha)\vec{1} \rangle \\ &= \frac{n}{3}\sigma(E(\tilde{X}))\sigma(E(\tilde{Y})). \end{aligned}$$

Hence, we have

$$T_{xx}^c := \langle \vec{X}^c, \vec{X}^c \rangle - n\tau(E(\tilde{X}))^2, \quad (14)$$

$$T_{xy}^c := \langle \vec{X}^c, \vec{Y}^c \rangle - n\tau(E(\tilde{X}))\tau(E(\tilde{Y})), \quad (15)$$

$$T_{xx}^r := \langle \vec{X}^r, \vec{X}^r \rangle - \frac{n}{3}\sigma(E(\tilde{X}))^2, \quad (16)$$

$$T_{xy}^r := \langle \vec{X}^r, \vec{Y}^r \rangle - \frac{n}{3}\sigma(E(\tilde{X}))\sigma(E(\tilde{Y})). \quad (17)$$

**Theorem 8.** Input  $\vec{Y}$  and  $\vec{X} \in F(\mathbb{R})^n$ . Suppose that

$$T_{xx}^c + T_{xx}^r \neq 0.$$

Let  $(-1)^t$  denote the sign of  $T_{xy}^c$  and

$$\delta^+ := \frac{|T_{xy}^c| + T_{xy}^r}{T_{xx}^c + T_{xx}^r} \quad (18)$$

and

$$\delta^- := \frac{-|T_{xy}^c| + T_{xy}^r}{T_{xx}^c + T_{xx}^r}. \quad (19)$$

Then the regression coefficients can be computed in the following cases:

- 1) If  $\delta^+ \geq 0$  and  $\sigma(E(\tilde{Y})) \geq \delta^+\sigma(E(\tilde{X}))$ , then

$$b = (-1)^t \delta^+, \quad (20)$$

$$\tilde{A} = P_{\Delta}(E(\tilde{Y})) \ominus bP_{\Delta}(E(\tilde{X})). \quad (21)$$

- 2) If  $\delta^+ \geq 0$  and

$$\delta^+\sigma(E(\tilde{X})) > \sigma(E(\tilde{Y})) \geq \delta^-\sigma(E(\tilde{X})),$$

then set

$$b^{(1)} = (-1)^t \delta^-, \quad (22)$$

$$\tilde{A}^{(1)} = P_{\Delta}(E(\tilde{Y})) \ominus b^{(1)}P_{\Delta}(E(\tilde{X})), \quad (23)$$

and applying Theorem 5, by Eqs.(6) and (7) compute real regression coefficients, say  $a^{(2)}$  and  $b^{(2)}$ . Finally, compare the two observation errors

$$d^2(\vec{Y}, \tilde{A}^{(1)}\vec{1} + b^{(1)}\vec{X}) \quad \text{and} \quad d^2(\vec{Y}, a^{(2)}\vec{1} + b^{(2)}\vec{X})$$

and output the best estimators.

- 3) If  $\delta^+ < 0$  or  $\sigma(E(\tilde{Y})) < \delta^-\sigma(E(\tilde{X}))$ , then compute real regression coefficients by Theorem 5.

## V. AN ILLUSTRATIVE EXAMPLE

**Example 1** ([1], [2], [8], [13], [18], [20]). In order to cross-reference to the previous contributions let's consider the following example designed by Sakawa and Yano [20]:

| $\vec{Y}_i$                             | $\vec{X}_i$                            |
|-----------------------------------------|----------------------------------------|
| 1. $[3.5 + 0.5\alpha, 4.5 - 0.5\alpha]$ | $[1.5 + 0.5\alpha, 2.5 - 0.5\alpha]$   |
| 2. $[5 + 0.5\alpha, 6 - 0.5\alpha]$     | $[3 + 0.5\alpha, 4 - 0.5\alpha]$       |
| 3. $[6.5 + \alpha, 8.5 - \alpha]$       | $[4.5 + \alpha, 6.5 - \alpha]$         |
| 4. $[6 + 0.5\alpha, 7 - 0.5\alpha]$     | $[6.5 + 0.5\alpha, 7.5 - 0.5\alpha]$   |
| 5. $[8 + 0.5\alpha, 9 - 0.5\alpha]$     | $[8 + 0.5\alpha, 9 - 0.5\alpha]$       |
| 6. $[7 + \alpha, 9 - \alpha]$           | $[9.5 + \alpha, 11.5 - \alpha]$        |
| 7. $[10 + 0.5\alpha, 11 - 0.5\alpha]$   | $[10.5 + 0.5\alpha, 11.5 - 0.5\alpha]$ |
| 8. $[9 + 0.5\alpha, 10 - 0.5\alpha]$    | $[12 + 0.5\alpha, 13 - 0.5\alpha]$     |

By applying Eqs.(3) and (4), represent all fuzzy numbers with CR-representations, as follows:

| $\vec{Y}_i$                       | $\vec{X}_i$                    |
|-----------------------------------|--------------------------------|
| 1. $[4.0, 0.5(1 - \alpha)]_{CR}$  | $[2.0, 0.5(1 - \alpha)]_{CR}$  |
| 2. $[5.5, 0.5(1 - \alpha)]_{CR}$  | $[3.5, 0.5(1 - \alpha)]_{CR}$  |
| 3. $[7.5, 1.0(1 - \alpha)]_{CR}$  | $[5.5, 1.0(1 - \alpha)]_{CR}$  |
| 4. $[6.5, 0.5(1 - \alpha)]_{CR}$  | $[7.0, 0.5(1 - \alpha)]_{CR}$  |
| 5. $[8.5, 0.5(1 - \alpha)]_{CR}$  | $[8.5, 0.5(1 - \alpha)]_{CR}$  |
| 6. $[8.0, 1.0(1 - \alpha)]_{CR}$  | $[10.5, 1.0(1 - \alpha)]_{CR}$ |
| 7. $[10.5, 0.5(1 - \alpha)]_{CR}$ | $[11.0, 0.5(1 - \alpha)]_{CR}$ |
| 8. $[9.5, 0.5(1 - \alpha)]_{CR}$  | $[12.5, 0.5(1 - \alpha)]_{CR}$ |

By Eq.(2) we obtain the mean of  $\tilde{Y}_i$  is

$$E(\tilde{Y}) = [6.875 + 0.625\alpha, 8.125 - 0.625\alpha].$$

Consequently, by applying Eqs.(3) and (4) the CR-representation of  $E(\tilde{Y})$  is

$$E(\tilde{Y}) = [7.5, 0.625(1 - \alpha)]_{CR}.$$

In the same way, we obtain

$$\begin{aligned} E(\tilde{X}) &= [6.9375 + 0.625\alpha, 8.1875 - 0.625\alpha] \\ &= [7.5625, 0.625(1 - \alpha)]_{CR}. \end{aligned}$$

Now, by Eqs.(8) and (9) we obtain

$$\begin{aligned} \tau(E(\tilde{X})) &= 7.5625, \\ \sigma(E(\tilde{X})) &= 0.625, \\ \tau(E(\tilde{Y})) &= 7.5, \\ \sigma(E(\tilde{Y})) &= 0.625. \end{aligned}$$

Hence, by Eq.(5) and (14)-(17) we obtain

$$\begin{aligned} T_{xx}^c &:= 97.71875, \\ T_{xy}^c &:= 50.75, \\ T_{xx}^r &:= 0.125, \\ T_{xy}^r &:= 0.125. \end{aligned}$$

$T_{xy}^c$  is positive, so  $(-1)^t = 1$ . Consequently, by Eqs.(18) and (19) we obtain

$$\delta^+ = 0.520$$

and

$$\delta^- = -0.517.$$

Since  $\delta^+ \geq 0$  and  $\sigma(E(\tilde{Y})) \geq \delta^+\sigma(E(\tilde{X}))$ , by Eqs.(20) and (21) (Theorem 8, Case 1) we obtain

$$b = (-1)^t \delta^+ = 0.520$$

and

$$\begin{aligned} \tilde{A} &= P_{\Delta}(E(\tilde{Y})) \ominus bP_{\Delta}(E(\tilde{X})) \quad (\text{by Eq.(14)}) \\ &= [\tau(E(\tilde{Y})) - b\tau(E(\tilde{X})), \\ &\quad (\sigma(E(\tilde{Y})) - |b|\sigma(E(\tilde{X}))(1 - \alpha))]_{CR} \\ &= [3.5675, 0.3(1 - \alpha)]_{CR} \\ &= [3.2675 + 0.3\alpha, 3.8675 - 0.3\alpha] \end{aligned}$$

**Table 1.** Comparison of the observation errors for previous estimators

|                    | $\vec{X}_i$       | $\vec{Y}_i$     | $d^2(\vec{Y}_i, \tilde{A} + b \odot \vec{X}_i)$ |         |         |
|--------------------|-------------------|-----------------|-------------------------------------------------|---------|---------|
|                    |                   |                 | D.                                              | C.-H.   | Thm. 8  |
| 1.                 | (1.5, 2, 2.5)     | (3.5, 4, 4.5)   | 0.7344                                          | 0.7466  | 0.7417  |
| 2.                 | (3, 3.5, 4)       | (5, 5.5, 6)     | 0.0282                                          | 0.0273  | 0.0275  |
| 3.                 | (4.5, 5.5, 6.5)   | (6.5, 7.5, 8.5) | 2.3177                                          | 2.3265  | 2.3200  |
| 4.                 | (6.5, 7, 7.5)     | (6, 6.5, 7)     | 1.0106                                          | 0.9965  | 1.0049  |
| 5.                 | (8, 8.5, 9)       | (8, 8.5, 9)     | 0.5195                                          | 0.5359  | 0.5267  |
| 6.                 | (9.5, 10.5, 11.5) | (7, 8, 9)       | 2.1578                                          | 2.1086  | 2.1352  |
| 7.                 | (10.5, 11, 11.5)  | (10, 10.5, 11)  | 2.9113                                          | 2.9743  | 2.9403  |
| 8.                 | (12, 12.5, 13)    | (9, 9.5, 10)    | 0.6648                                          | 0.6285  | 0.6476  |
| Observation errors |                   |                 | 10.3443                                         | 10.3442 | 10.3439 |

Using the data set in Example 1, Diamond [4] obtained the following regression coefficients

$$\tilde{A} = [3.263 + 0.3\alpha, 3.863 - 0.3\alpha],$$

$$b = 0.521.$$

And, Chen and Hsueh [3] obtained

$$\tilde{A} = [3.272 + 0.3\alpha, 3.872 - 0.3\alpha],$$

$$b = 0.521.$$

Table 1 shows that the observation errors corresponding to the above three estimators, where the observation errors are computed by Eq.(1). It is easy to see that our estimators are the best, as follows

$$\tilde{A} = [3.2675 + 0.3\alpha, 3.8675 - 0.3\alpha],$$

$$b = 0.520.$$

## VI. CONCLUSION

In the paper, we apply the CR-representations of fuzzy numbers and symmetric triangular approximations to solve the regression model whose constant coefficient is assumed to be symmetric triangular and other coefficients to be real. The proposed formulae are simple and clear. In the future, one may consider applying more general approximations (such as triangular approximations [30], trapezoidal approximations [7], [26], and symmetric trapezoidal approximations [28]) to solve the following regression model

$$d^2(\vec{Y}, \tilde{A}\vec{1} + b\vec{X}),$$

where  $b \in \mathbb{R}$  and  $\tilde{A}$  is triangular (trapezoidal or symmetric trapezoidal).

## ACKNOWLEDGMENT

This work is supported by National Science Council of Taiwan (99-2115-M-024-002). The author also gratefully acknowledges the helpful comments and suggestions of the reviewers, which have improved the presentation.

## REFERENCES

- [1] A. Bargiela, W. Pedrycz, T. Nakashima, "Multiple regression with fuzzy data," *Fuzzy Sets and Systems*, vol.158, pp.2169-2188, 2007.
- [2] L.-H. Chen and C.-C. Hsueh, "A mathematical programming method for formulating a fuzzy regression model based on distance criterion," *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol.37, no.3, pp.705-712, 2007.
- [3] L.-H. Chen and C.-C. Hsueh, "Fuzzy Regression Models Using the Least-Squares Method Based on the Concept of Distance," *IEEE Transactions on Fuzzy Systems*, vol.17, no.6, pp.1259-1272, 2009.
- [4] P. Diamond, "Fuzzy least squares," *Information Science*, vol.46, pp.141-157, 1988.
- [5] P. Diamond and R. Korner, "Extended fuzzy linear models and least squares estimates," *Computers and Mathematics with Applications*, vol.33, no.9, pp.15-32, 1997.
- [6] P. D'Urso and A. Santoro, "Goodness of fit and variable selection in the fuzzy multiple linear regression," *Fuzzy Sets and Systems*, vol.157, pp.2627-2647, 2006.
- [7] P. Grzegorzewski and E. Mrowka, "Trapezoidal approximations of fuzzy numbers preserving the expected interval - algorithms and properties," *Fuzzy Sets and Systems*, vol.159, pp.1354-1364, 2008.
- [8] M. Hojati, C.R. Bector and K. Smimou, "A simple method for computation of fuzzy linear regression," *European Journal of Operational Research*, vol.166, pp.172-184, 2005.
- [9] D.H. Hong and C. Hwang, "Extended fuzzy regression models using regularization method," *Information Science*, vol.164, pp.31-46, 2004.
- [10] D.H. Hong and C. Hwang, "Interval regression analysis using quadratic loss support vector machine," *IEEE Transactions on Fuzzy Systems*, vol.13, no.2, pp.229-237, 2005.
- [11] J. Kacprzyk and M. Fedrizzi, Eds., "Fuzzy least squares regression: theory and application," in *Fuzzy Regression Analysis*, Omnitech Press Warsaw and Physica-Verlag, Heidelberg, 1992, pp.181-193.
- [12] C. Kao and C.L. Chyu, "A fuzzy linear regression model with better explanatory power," *Fuzzy Sets and Systems*, vol.126, pp.401-409, 2002.
- [13] C. Kao and C.L. Chyu, "Least-squares estimates in fuzzy regression analysis," *European Journal of Operational Research*, vol.148, pp.426-435, 2003.
- [14] B. Kim and R.R. Bishu, "Evaluation of fuzzy linear regression models by comparing membership functions," *Fuzzy Sets and Systems*, vol.100, pp.343-352, 1998.
- [15] H.K. Kim, J.H. Yoon, Y. Li, "Asymptotic properties of least squares estimation with fuzzy observations," *Information Science*, vol.178, pp.439-451, 2008.
- [16] R. Korner and W. Nather, "Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates," *Information Science*, vol.109, pp.95-118, 1998.
- [17] V. Kratshmer, "Least-squares estimation in linear regression models with vague concepts," *Fuzzy Sets and Systems*, vol.157, pp.2579-2592, 2006.
- [18] M.H. Mashinchi, M.A. Orgun, M. Mashinchi and W. Pedrycz, "A tabu-harmony search based approach to fuzzy linear regression," *IEEE Transactions on Fuzzy Systems*, accepted.
- [19] M. Modarres, E. Nasrabadi and M.M. Nasrabadi, "Fuzzy linear regression models with least square errors," *Applied Mathematics and Computation*, vol.163, pp.977-989, 2005.

- [20] M. Sakawa and H. Yano, "Multiobjective fuzzy linear regression analysis for fuzzy input-output data," *Fuzzy Sets and Systems*, vol.47, pp.173-181, 1992.
- [21] H. Tanaka, S. Uejima, K. Asai, "Linear regression analysis with fuzzy model," *IEEE Transactions on Systems Man Cybernetics*, vol.12, No.6, pp.903-907, 1982.
- [22] N. Wang, W.X. Zhang, C.L. Mei, "Fuzzy nonparametric regression based on local linear smoothing technique," *Information Science*, vol.177, pp.3882-3900, 2007.
- [23] A. Wunsche and W. Nather, "Least-squares fuzzy regression with fuzzy random variables," *Fuzzy Sets and Systems*, vol.130, pp.43-50, 2002.
- [24] C.-T. Yeh, "Reduction of fuzzy linear systems of dual equations," *International Journal of Fuzzy Systems*, Vol.9, No.3, pp.173-178, 2007.
- [25] C.-T. Yeh, "On improving trapezoidal and triangular approximations of fuzzy numbers," *International Journal of Approximate Reasoning*, vol.48, no.1, pp.297-313, 2008.
- [26] C.-T. Yeh, "Weighted trapezoidal and triangular approximations of fuzzy numbers," *Fuzzy Sets and Systems*, vol.160, pp.3059-3079, 2009.
- [27] C.-T. Yeh, "Reduction to least-squares estimates in multiple fuzzy regression analysis," *IEEE Transactions on Fuzzy Systems*, vol.17, no.4, pp.935-948, 2009.
- [28] C.-T. Yeh, "Approximations by interval, triangular and trapezoidal fuzzy numbers," *Proceeding of 2009 International Fuzzy Systems Association World Congress*, pp.143-148, 2009.
- [29] C.-T. Yeh, "Weighted semi-trapezoidal of fuzzy numbers," *Fuzzy Sets and Systems*, vol.165, pp.61-80, 2011.
- [30] W. Zeng and H. Li, "Weighted triangular approximation of fuzzy numbers," *International Journal of Approximate Reasoning*, vol.46, no.1, pp.137-150, 2007.