THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 8, Number 2/2007, pp. 000-000

A MODEL OF MULTIPLE LINEAR REGRESSION

Ciprian POPESCU*, Marius GIUCLEA**

*Department of Mathematics, Academy of Economic Studies, 6 Piata Romana, 010374, Bucharest, Romania
**Institute of Solid Mechanics of the Romanian Academy, 15 Constantin Mile St., 70701, Bucharest, Romania
Corresponding author: Ciprian POPESCU, e-mail: popescucip@yahoo.com

In this paper, a model of multiple linear regression for fuzzy sets is given. Two cases are considered: first, for fuzzy numbers in a general form, second, for triangular numbers. Next, we give a numerical application of the proposed method.

Key words: Multiple linear regression, fuzzy numbers.

1. INTRODUCTION

The regression is one of the most useful methods for processing fuzzy data. There are many studies in this area. For example, Diamond in [3] has applied the least squares method for triangular fuzzy numbers. On the other hand, Puri and Ralescu [12], Wu and Ma [15], Ma Ming et al [9] have established an isomorphically and isometrically correspondence between a fuzzy space and a certain Banach space. On this foundation they have introduced a new norm in fuzzy spaces. Ma Ming et al have proposed a model of simple linear regression on fuzzy sets.

This paper proposes an extension of the above-mentioned works at the multiple linear regression, which has a wider applicability in many areas. At the beginning, the second section, gives the definitions regarding fuzzy numbers and the distance on the fuzzy numbers set. Next, the general method of Multiple Linear Regression (MLR) for fuzzy data is presented in section 3. The most used case of fuzzy numbers in the triangular form is considered in the fourth section and an illustrative numerical example is shown in section 5.

2. PRELIMINARIES

Let us assume a fuzzy number space denoted by F , [15]. For $w \in F$ and $r \in [0,1]$ is considered the closed interval defined by $[w]^r = \begin{cases} \{t \,/\, w(t) \geq r\}, & r \in (0,1] \\ \{t \,/\, w(t) > 0\}, & r = 0 \end{cases}$. $[w]^r$ is a closed interval which has as lower and upper bounds two functions, $\underline{w}(r)$ and $\overline{w}(r)$, with some special properties, [4]. F has embedded isomorphically and isometrically, [15] in a certain Banach space. The distance between the fuzzy numbers $u = (\underline{u}(r), \overline{u}(r))$ and $v = (\underline{v}(r), \overline{v}(r))$ is given by:

$$D_2(u, v) = \left[\int_0^1 (\underline{u}(r) - \underline{v}(r))^2 \, dr + \int_0^1 (\overline{u}(r) - \overline{v}(r))^2 \, dr \right]^{\frac{1}{2}} \tag{2.1}$$

3. THE GENERAL MODEL

The regression function depends by p independent fuzzy variables $X_1, X_2, ..., X_p \in F$:

$$g = g\left(X_1, X_2, ..., X_p\right) = b_0 + b_1 X_1 + ... + b_p X_p \tag{3.1}$$

For an observed variable $Y_i \in F$ we obtain the estimation:

$$Y_i = b_0 + b_1 X_{i1} + ... + b_p X_{ip}, i = \overline{1, n}. \tag{3.2}$$

Our task is to find the real parameters $b_0, b_1, ..., b_p$ that minimize the sum:

$$S\left(b_0, b_1, ..., b_p\right) = \sum_{i=1}^{n} D_2^2\left(b_0 + b_1 X_{i1} + ... + b_p X_{ip}, Y_i\right) \tag{3.3}$$

From relations (2.1) and (3.3) one can conclude that the explicit form of $S\left(b_0, b_1, ..., b_p\right)$ depends on the signs of parameters $b_0, b_1, ..., b_p$: totally 2^p possibilities. Here is considered the most general case: the set $B = \{b_j \in R / j = \overline{1, p}\}$ contains l positive elements and $p - l$ negative elements. Let $\{1,2,..., p\} = \{k_1, k_2, ..., k_p\}$ such that $B = \{b_1, b_2, ..., b_p\} = \{b_{k_1}, ..., b_{k_p}\}$ where $b_{k_j} > 0, j = \overline{1, l}$ and $b_{k_j} < 0, j = \overline{l + 1, p}$. Thus, we search the minimum for:

$$\begin{aligned} S\left(b_0, b_1, ..., b_p\right) &= S\left(b_0, b_{k_1}, ..., b_{k_l}, b_{k_{l+1}}, ..., b_p\right) = \sum_{i=1}^{n} D_2^2\left(b_0 + b_{k_1} X_{ik_1} + ... + b_{k_p} X_{ik_p}, Y_i\right) = \\ &= \sum_{i=1}^{n}\left[\int_0^1 \left(b_0 + b_{k_1}\underline{X}_{ik_1} + ... + b_{k_l}\underline{X}_{ik_l} + b_{k_{l+1}}\overline{X}_{ik_{l+1}} + ... + b_{k_p}\overline{X}_{ik_p} - \underline{Y}_i\right)^2 dr\right] + \\ &+ \sum_{i=1}^{n}\left[\int_0^1 \left(b_0 + b_{k_1}\overline{X}_{ik_1} + ... + b_{k_l}\overline{X}_{ik_l} + b_{k_{l+1}}\underline{X}_{ik_{l+1}} + ... + b_{k_p}\underline{X}_{ik_p} - \overline{Y}_i\right)^2 dr\right] \end{aligned} \tag{3.4}$$

Next, is solved the following system of $p + 1$ equations:

$$\begin{cases} \dfrac{\partial S}{\partial b_0} = 0 \\ \dfrac{\partial S}{\partial b_{k_j}} = 0, j = \overline{1, n} \end{cases} \tag{3.5}$$

which is equivalent with:

$$\begin{cases} 2nb_0 + \sum_{j=1}^{p} \alpha_j b_{k_j} = \sum_{i=1}^{n}\left[\int_0^1 \left(\underline{Y}_i + \overline{Y}_i\right)dr\right] \\ \ldots \\ \alpha_l b_0 + \sum_{j=1}^{l-1}\gamma_{lj}b_{k_j} + \beta_l b_{k_l} + \sum_{j=l+1}^{p}\delta_{lj}b_{k_j} = \sum_{i=1}^{n}\left[\int_0^1 \left(\underline{X}_{ik_l}\underline{Y}_i + \overline{X}_{ik_l}\overline{Y}_i\right)dr\right] \\ \alpha_{l+1} b_0 + \sum_{j=1}^{l}\delta_{l+1,j}b_{k_j} + \beta_{l+1} b_{k_{l+1}} + \sum_{j=l+2}^{p}\gamma_{l+1,j}b_{k_j} = \sum_{i=1}^{n}\left[\int_0^1 \left(\overline{X}_{ik_{l+1}}\underline{Y}_i + \underline{X}_{ik_{l+1}}\overline{Y}_i\right)dr\right] \\ \ldots \\ \alpha_p b_0 + \sum_{j=1}^{l}\delta_{pj}b_{k_j} + \sum_{j=l+1}^{p-1}\gamma_{pj}b_{k_j} + \beta_p b_{k_p} = \sum_{i=1}^{n}\left[\int_0^1 \left(\overline{X}_{ik_p}\underline{Y}_i + \underline{X}_{ik_p}\overline{Y}_i\right)dr\right] \end{cases} \tag{3.6}$$

In (3.6), for $s, t = \overline{1, p}$ are used the substitutions:

$$\sum_{i=1}^{n}\left[\int_0^1 \left(\underline{X}_{ik_s} + \overline{X}_{ik_s}\right)dr\right] = \alpha_s,\ \sum_{i=1}^{n}\left[\int_0^1 \left(\underline{X}_{ik_s}{}^2 + \overline{X}_{ik_s}{}^2\right)dr\right] = \beta_s,$$

$$\sum_{i=1}^{n}\left[\int_0^1 \left(\overline{X}_{ik_s}\overline{X}_{ik_t} + \underline{X}_{ik_s}\underline{X}_{ik_t}\right)dr\right] = \gamma_{st},\ \sum_{i=1}^{n}\left[\int_0^1 \left(\overline{X}_{ik_s}\underline{X}_{ik_t} + \underline{X}_{ik_s}\overline{X}_{ik_t}\right)dr\right] = \delta_{st} \tag{3.7}$$

Obviously $\gamma_{st} = \gamma_{ts}, \delta_{st} = \delta_{ts}$ and the matrix of system (3.6) is symmetrical:

$$\mathbf{A} = \begin{pmatrix} 2n & \alpha_1 & & \alpha_l & \alpha_{l+1} & & \alpha_p \\ \alpha_1 & \beta_1 & & \gamma_{1l} & \delta_{1,l+1} & & \delta_{1p} \\ & & \ddots & \vdots & & & \\ \alpha_l & \gamma_{l1} & \ldots & \beta_l & \delta_{l,l+1} & & \delta_{lp} \\ \alpha_{l+1} & \delta_{l+1,1} & & \delta_{l+1,l} & \beta_{l+1} & \ldots & \gamma_{l+1,p} \\ & & & & \vdots & \ddots & \\ \alpha_p & \delta_{p1} & & \delta_{pl} & \gamma_{p,l+1} & & \beta_p \end{pmatrix} \tag{3.8}$$

Remark:

The determinants of the type $\Delta_{1s} = \begin{vmatrix} 2n & \alpha_s \\ \alpha_s & \beta_s \end{vmatrix}$, $\Delta_{rs} = \begin{vmatrix} \beta_r & \gamma_{rs} \\ \gamma_{sr} & \beta_s \end{vmatrix}$, $\Delta'_{rs} = \begin{vmatrix} \beta_r & \delta_{rs} \\ \delta_{sr} & \beta_s \end{vmatrix}$ *for* $r, s = \overline{1, p}$ *are strictly positive.*

Proof:

$$\Delta_{1s} = \begin{vmatrix} 2n & \alpha_s \\ \alpha_s & \beta_s \end{vmatrix} = \begin{vmatrix} 2n & \sum_{i=1}^{n}\left[\int_0^1 \left(\underline{X}_{ik_s} + \overline{X}_{ik_s}\right)dr\right] \\ \sum_{i=1}^{n}\left[\int_0^1 \left(\underline{X}_{ik_s} + \overline{X}_{ik_s}\right)dr\right] & \sum_{i=1}^{n}\left[\int_0^1 \left(\underline{X}_{ik_s}^2 + \overline{X}_{ik_s}^2\right)dr\right] \end{vmatrix} =$$

$$= 2n\left\{\sum_{i=1}^{n}\left[\int_0^1 \left(\underline{X}_{ik_s}^2 + \overline{X}_{ik_s}^2\right)dr\right]\right\} - \left\{\sum_{i=1}^{n}\left[\int_0^1 \left(\underline{X}_{ik_s} + \overline{X}_{ik_s}\right)dr\right]\right\}^2$$

From Cauchy-Buniakowsky-Schwarz relation is obtained:

$$2n\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_s}^2+\overline{X}_{ik_s}^2\right)dr\right]\right\}=\left[\int_0^1(2n)dr\right]\left[\int_0^1\left(\sum_{i=1}^{n}\underline{X}_{ik_s}^2+\sum_{i=1}^{n}\overline{X}_{ik_s}^2\right)dr\right]\geq$$

$$\geq\left[\int_0^1\left(\sqrt{2n}\sqrt{\sum_{i=1}^{n}\left(\underline{X}_{ik_s}^2+\overline{X}_{ik_s}^2\right)}\right)dr\right]^2\geq\left[\int_0^1\left(\sum_{i=1}^{n}\left(\underline{X}_{ik_s}+\overline{X}_{ik_s}\right)\right)dr\right]^2=\left[\sum_{i=1}^{n}\left(\int_0^1\left(\underline{X}_{ik_s}+\overline{X}_{ik_s}\right)dr\right)\right]^2\Rightarrow\Delta_{1s}\geq 0.$$

$$\Delta_{rs}=\begin{vmatrix}\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}^2+\overline{X}_{ik_r}^2\right)dr\right] & \sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}\underline{X}_{ik_s}+\overline{X}_{ik_r}\overline{X}_{ik_s}\right)\right]\\ \sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}\underline{X}_{ik_s}+\overline{X}_{ik_r}\overline{X}_{ik_s}\right)\right] & \sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_s}^2+\overline{X}_{ik_s}^2\right)\right]\end{vmatrix}=$$

$$=\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}^2+\overline{X}_{ik_r}^2\right)dr\right]\right\}\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_s}^2+\overline{X}_{ik_s}^2\right)dr\right]\right\}-\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}\underline{X}_{ik_s}+\overline{X}_{ik_r}\overline{X}_{ik_s}\right)dr\right]\right\}^2$$

Again, using Cauchy-Buniakowsky-Schwarz relation results:

$$\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}^2+\overline{X}_{ik_r}^2\right)dr\right]\right\}\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_s}^2+\overline{X}_{ik_s}^2\right)dr\right]\right\}\geq$$

$$\geq\left\{\sum_{i=1}^{n}\sqrt{\int_0^1\underline{X}_{ik_r}^2dr}\sqrt{\int_0^1\underline{X}_{ik_s}^2dr}+\sum_{i=1}^{n}\sqrt{\int_0^1\overline{X}_{ik_r}^2dr}\sqrt{\int_0^1\overline{X}_{ik_s}^2dr}\right\}^2\geq\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}\underline{X}_{ik_s}+\overline{X}_{ik_r}\overline{X}_{ik_s}\right)dr\right]\right\}^2\Rightarrow\Delta_{rs}\geq 0.$$

For $\mathbf{\Delta'}_{rs}$ the proof is analogue:

$$\Delta'_{rs}=\begin{vmatrix}\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}^2+\overline{X}_{ik_r}^2\right)dr\right] & \sum_{i=1}^{n}\left[\int_0^1\left(\overline{X}_{ik_r}\underline{X}_{ik_s}+\underline{X}_{ik_r}\overline{X}_{ik_s}\right)\right]\\ \sum_{i=0}^{n}\left[\int_0^1\left(\overline{X}_{ik_r}\underline{X}_{ik_s}+\underline{X}_{ik_r}\overline{X}_{ik_s}\right)\right] & \sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_s}^2+\overline{X}_{ik_s}^2\right)\right]\end{vmatrix}=$$

$$=\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_r}^2+\overline{X}_{ik_r}^2\right)dr\right]\right\}\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\underline{X}_{ik_s}^2+\overline{X}_{ik_s}^2\right)dr\right]\right\}-\left\{\sum_{i=1}^{n}\left[\int_0^1\left(\overline{X}_{ik_r}\underline{X}_{ik_s}+\underline{X}_{ik_r}\overline{X}_{ik_s}\right)dr\right]\right\}^2\geq 0$$

Unless all the data are collected from the same crisp X, the determinants are strictly positive. For $p=1$ we obtain the simple linear regression [9], with the unique solution ($\Delta_{11}>0$).

4. THE APPROXIMATIVE MODEL

This section studies the MLR for fuzzy numbers in the triangular form. This kind of approach is more suitable for practical applications. The input data, $X_{ij}=\left(\underline{X}_{ij}(r),\overline{X}_{ij}(r)\right), Y_i=\left(\underline{Y}_i(r),\overline{Y}_i(r)\right), i=\overline{1,n} j=\overline{1,p}$ are written in the triangular form [3]: $X_{ij}=\left(x_{ij},\underline{u}_{ij},\overline{u}_{ij}\right)$, $Y_i=\left(y_i,\underline{v}_i,\overline{v}_i\right)$ where

$$\underline{X}_{ij}(r)=x_{ij}-\underline{u}_{ij}+\underline{u}_{ij}r,\ \overline{X}_{ij}(r)=x_{ij}+\overline{u}_{ij}-\overline{u}_{ij}r,\ \underline{Y}_i(r)=y_i-\underline{v}_i+\underline{v}_ir,\ \overline{Y}_i(r)=y_i+\overline{v}_i-\overline{v}_ir.$$

Here is considered the approximation: $\int_0^1 f(r)dr\cong[f(0)+f(1)]/2$. Therefore, the sum becomes:

$$S\left(b_0, b_{k_1}, b_{k_2}, ..., b_{k_p}\right) = \sum_{i=1}^{n} D_2^2\left(b_0 + b_{k_1} X_{ik_1} ... + b_{k_p} X_{ik_p}, Y_i\right) =$$

$$= \sum_{i=1}^{n}\left[\int_0^1 \left(b_0 + b_{k_1}\underline{X}_{ik_1} + ... + b_{k_l}\underline{X}_{ik_l} + b_{k_{l+1}}\overline{X}_{ik_{l+1}} + ... + b_{k_p}\overline{X}_{ik_p} - \underline{Y}_i\right)^2 dr\right] +$$

$$+ \sum_{i=1}^{n}\left[\int_0^1 \left(b_0 + b_{k_1}\overline{X}_{ik_1} + ... + b_{k_l}\overline{X}_{ik_l} + b_{k_{l+1}}\underline{X}_{ik_{l+1}} + ... + b_{k_p}\underline{X}_{ik_p} - \overline{Y}_i\right)^2 dr\right] =$$

$$= \sum_{i=1}^{n}\left[\int_0^1 \left[b_0 + \sum_{j=1}^{l} b_{k_j}\left(x_{ik_j} - \underline{u}_{ik_j} + \underline{u}_{ik_j} r\right) + \sum_{j=l+1}^{p} b_{k_j}\left(x_{ik_j} + \overline{u}_{ik_j} - \overline{u}_{ik_j} r\right) - \left(y_i - \underline{v}_i + \underline{v}_i r\right)\right]^2 dr\right] +$$

$$+ \sum_{i=1}^{n}\left[\int_0^1 \left[b_0 + \sum_{j=1}^{l} b_{k_j}\left(x_{ik_j} + \overline{u}_{ik_j} - \overline{u}_{ik_j} r\right) + \sum_{j=l+1}^{p} b_{k_j}\left(x_{ik_j} - \underline{u}_{ik_j} + \underline{u}_{ik_j} r\right) - \left(y_i + \overline{v}_i - \overline{v}_i r\right)\right]^2 dr\right] \cong$$

$$\cong \frac{1}{2}\sum_{i=1}^{n} \left\{\left[\left(b_0 + \sum_{j=1}^{p} b_{k_j} x_{ik_j}\right) - \left(\sum_{j=1}^{l} b_{k_j}\underline{u}_{ik_j}\right) + \left(\sum_{j=l+1}^{p} b_{k_j}\overline{u}_{ik_j}\right) - \left(y_i - \underline{v}_i\right)\right]^2 + \right.$$

$$\left. + \left[\left(b_0 + \sum_{j=1}^{p} b_{k_j} x_{ik_j}\right) + \left(\sum_{j=1}^{l} b_{k_j}\overline{u}_{ik_j}\right) - \left(\sum_{j=l+1}^{p} b_{k_j}\underline{u}_{ik_j}\right) - \left(y_i + \overline{v}_i\right)\right]^2 + 2\left[b_0 + \left(\sum_{j=1}^{p} b_{k_j} x_{ik_j}\right) - y_i\right]^2\right\}.$$

Next is determined the solutions for the following system:

$$\begin{cases} \dfrac{\partial S}{\partial b_0} = 0 \\ \dfrac{\partial S}{\partial b_{k_j}} = 0, \ j = \overline{1, n} \end{cases} \tag{4.1}$$

Using the substitutions $\left(x_{ik_j} - \underline{u}_{ik_j}\right) = \underline{a}_{ij}, \left(x_{ik_j} + \overline{u}_{ik_j}\right) = \overline{a}_{ij}, \left(y_i - \underline{v}_i\right) = \underline{c}_i, \left(y_i + \overline{v}_i\right) = \overline{c}_i$ for $i = \overline{1, n}$, $j = \overline{1, p}$, (4.1) is equivalent with:

$$4nb_0 + \sum_{j=1}^{p}\left[b_{k_j}\sum_{i=1}^{n}\left(\underline{a}_{ij} + \overline{a}_{ij} + 2x_{ik_j}\right)\right] = \sum_{i=1}^{n}\left(\underline{c}_i + \overline{c}_i + 2y_i\right)$$

.....

$$\begin{aligned} & b_0\sum_{i=1}^{n}\left(\underline{a}_{il} + \overline{a}_{il} + 2x_{ik_l}\right) + \sum_{j=1}^{l-1}\left\{b_{k_j}\sum_{i=1}^{n}\left[\underline{a}_{il}\underline{a}_{ij} + \overline{a}_{il}\overline{a}_{ij} + x_{ik_j}x_{ik_l}\right]\right\} + \\ & + b_{k_l}\sum_{i=1}^{n}\left[\underline{a}_{il}^2 + \overline{a}_{il}^2 + 2x_{ik_l}^2\right] + \sum_{j=l+1}^{p}\left\{b_{k_j}\sum_{i=1}^{n}\left[\underline{a}_{il}\overline{a}_{ij} + \overline{a}_{il}\underline{a}_{ij} + x_{ik_j}x_{ik_l}\right]\right\} = \sum_{i=1}^{n}\left[\underline{a}_{il}\underline{c}_i + \overline{a}_{il}\overline{c}_i + 2y_i x_{ik_l}\right] \end{aligned} \tag{4.2}$$

$$b_0\sum_{i=1}^{n}\left(\underline{a}_{i,l+1} + \overline{a}_{i,l+1} + 2x_{ik_{l+1}}\right) + \sum_{j=1}^{l}\left\{b_{k_j}\sum_{i=1}^{n}\left[\underline{a}_{i,l+1}\overline{a}_{ij} + \overline{a}_{i,l+1}\underline{a}_{ij} + 2x_{ik_j}x_{ik_{l+1}}\right]\right\} +$$

$$+ b_{k_{l+1}} \sum_{i=1}^{n} \left[\underline{a}_{i,l+1}^2 + \overline{a}_{i,l+1}^2 + 2x_{ik_{l+1}}^2 \right] + \sum_{j=l+2}^{p} \left\{ b_{k_j} \sum_{i=1}^{n} \left[\underline{a}_{i,l+1} \underline{a}_{ij} + \overline{a}_{i,l+1} \overline{a}_{ij} + 2x_{ik_j} x_{ik_{l+1}} \right] \right\} =$$

$$= \sum_{i=1}^{n} \left[\overline{a}_{i,l+1} \underline{c}_i + \underline{a}_{i,l+1} \overline{c}_i + 2y_i x_{ik_{l+1}} \right]$$

.....

$$b_0 \sum_{i=1}^{n} \left(\underline{a}_{ip} + \overline{a}_{ip} + 2x_{ik_p} \right) + \sum_{j=1}^{l} \left\{ b_{k_j} \sum_{i=1}^{n} \left[\underline{a}_{ip} \overline{a}_{ij} + \overline{a}_{ip} \underline{a}_{ij} + 2x_{ik_j} x_{ik_p} \right] \right\} +$$

$$+ \sum_{j=l+1}^{p-1} \left\{ b_{k_j} \sum_{i=1}^{n} \left[\underline{a}_{ip} \underline{a}_{ij} + \overline{a}_{ip} \overline{a}_{ij} + 2x_{ik_j} x_{ik_p} \right] \right\} + b_{k_p} \sum_{i=1}^{n} \left[\underline{a}_{ip}^2 + \overline{a}_{ip}^2 + 2x_{ik_p}^2 \right] = \sum_{i=1}^{n} \left[\overline{a}_{ip} \underline{c}_i + \underline{a}_{ip} \overline{c}_i + 2y_i x_{ik_p} \right]$$

Theorem:

$\mathbf{b}^*$ *is an estimator in the meaning of least squares for the parameters vector* $\mathbf{b}$ *if and only if* $\mathbf{b}^*$ *is a solution for the system (4.2).*

Proof:

The matrix for the system (4.2) is:

$$\mathbf{A} = \begin{pmatrix} 4n & \sum_{i=1}^{n} \left(\underline{a}_{i1} + \overline{a}_{i1} + 2x_{ik_1} \right) & \dots & \sum_{i=1}^{n} \left(\underline{a}_{ip} + \overline{a}_{ip} + 2x_{ik_p} \right) \\ \sum_{i=1}^{n} \left(\underline{a}_{i1} + \overline{a}_{i1} + 2x_{ik_1} \right) & \sum_{i=1}^{n} \left(\underline{a}_{i1}^2 + \overline{a}_{i1}^2 + 2x_{ik_1}^2 \right) & \dots & \sum_{i=1}^{n} \left(\underline{a}_{i1} \overline{a}_{ip} + \overline{a}_{i1} \underline{a}_{ip} + 2x_{ik_1} x_{ik_p} \right) \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^{n} \left(\underline{a}_{ip} + \overline{a}_{ip} + 2x_{ik_p} \right) & \sum_{i=1}^{n} \left(\underline{a}_{ip} \overline{a}_{i1} + \overline{a}_{ip} \underline{a}_{i1} + 2x_{ik_p} x_{ik_1} \right) & \dots & \sum_{i=1}^{n} \left(\underline{a}_{ip}^2 + \overline{a}_{ip}^2 + 2x_{ik_p}^2 \right) \end{pmatrix}$$

Let the matrix:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 & 1 & 1 \\ \underline{a}_{11} & \overline{a}_{11} & x_{1k_1} & x_{1k_1} & \dots & \underline{a}_{n1} & \overline{a}_{n1} & x_{nk_1} & x_{nk_1} \\ \underline{a}_{12} & \overline{a}_{12} & x_{1k_2} & x_{1k_2} & \dots & \underline{a}_{n2} & \overline{a}_{n2} & x_{nk_2} & x_{nk_2} \\ & \vdots & & & & & \vdots & & \\ \underline{a}_{1p} & \overline{a}_{1p} & x_{1k_p} & x_{1k_p} & \dots & \underline{a}_{np} & \overline{a}_{np} & x_{nk_p} & x_{nk_p} \end{pmatrix}$$

We call $\mathbf{x}_l \in R^{4n}, l = \overline{1, p+1}$ the $l-$th row of the matrix $\mathbf{X}$. Thus $\mathbf{x}_1 = (1 \dots 1)$ and $\mathbf{x}_l = \left(\underline{a}_{1,l-1} \; \overline{a}_{1,l-1} \; x_{1k_{l-1}} \; x_{1k_{l-1}} \; \dots \; \underline{a}_{n,l-1} \; \overline{a}_{n,l-1} \; x_{nk_{l-1}} \; x_{nk_{l-1}} \right),\ l = \overline{1, p+1}$

Obviously, $\mathbf{X} \in M_{p+1,4n}(\mathbf{R}), \mathbf{X}^{\mathrm{T}} \in M_{4n,p+1}(\mathbf{R})$ and from calculus we obtain $\mathbf{X}\mathbf{X}^{\mathrm{T}} = \mathbf{A}$.

Let r the rank of $\mathbf{X}$: $r = \mathrm{rank}\mathbf{X} = \mathrm{rank}\mathbf{X}^{\mathrm{T}} \Rightarrow\ r \le \min(p+1, 4n)$. Consider the vectors: $\mathbf{b} = (b_0, b_1, \dots, b_p)^{\mathrm{T}} \in \mathbf{R}^{p+1}$, $\mathbf{y} = \left(\underline{c}_1, \overline{c}_1, y_1, y_1, \underline{c}_2, \overline{c}_2, y_2, y_2, \dots, \underline{c}_n, \overline{c}_n, y_n, y_n \right)^{\mathrm{T}} \in \mathbf{R}^{4n}$

We have $S(b_0, b_1, b_2, \dots, b_p) = \left(\mathbf{y} - \mathbf{X}^{\mathrm{T}}\mathbf{b} \right)^{\mathrm{T}} \left(\mathbf{y} - \mathbf{X}^{\mathrm{T}}\mathbf{b} \right) = \left\| \mathbf{y} - \mathbf{X}^{\mathrm{T}}\mathbf{b} \right\|^2$. The system (4.2) has the following matrix form: $\mathbf{A}\mathbf{b} = \mathbf{X}\mathbf{y} \Leftrightarrow \mathbf{X}\mathbf{X}^{\mathrm{T}}\mathbf{b} = \mathbf{X}\mathbf{y}$. We put $\mathbf{w} = \mathbf{X}^{\mathrm{T}}\mathbf{b} = \sum_{\mathbf{j}=\mathbf{0}}^{\mathbf{p}} b_j \mathbf{x}_{j+1} \in \mathbf{R}^{4n}$. Let $\mathbf{V}$ the vector subspace of $\mathbf{R}^{4n}$ generated by the columns $\mathbf{x}_j,\ j = \overline{1, p+1}$ from the matrix $\mathbf{X}^{\mathrm{T}}$.

The vectors $\mathbf{v}$ of this subspace have the form: $\mathbf{v} = \beta_0\mathbf{x}_1 + \beta_1\mathbf{x}_2 + ... + \beta_p\mathbf{x}_{p+1}$, where $\beta_0, \beta_1, ..., \beta_p$ are real numbers. Thus $\mathbf{v} = \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}$ where $\boldsymbol{\beta} = \begin{pmatrix}\beta_0 & \beta_1 & ... & \beta_p\end{pmatrix}^{\mathrm{T}}$. The vector $\boldsymbol{\beta} \in \mathbf{R}^{p+1}$ is variable then the vector $\mathbf{v}$ is also variable.

But the expression $\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}\|^2$ has a minimum which is attained if and only if $\mathbf{v}$ is the orthogonal projection of $\mathbf{y}$ on $\mathbf{V}$, or, equivalently, $\mathbf{v} = \mathrm{pr}_{\mathbf{V}}\mathbf{y} = \mathbf{w}^*$. Since $\mathbf{w}^* \in \mathbf{V}$, there exists the real numbers $b_0^*, b_1^*, ..., b_p^* \in \mathbf{R}$ such that $\mathbf{w}^* = \sum_{j=0}^{p} b_j^*\mathbf{x}_{j+1}$. Since $\boldsymbol{\eta}^* = \boldsymbol{\eta}^*(\mathbf{y})$ depends only by $\mathbf{y}$, then the real scalars $b_0^*, b_1^*, ..., b_p^*$ depends only by $\mathbf{y}$. Thus $\boldsymbol{\beta} = \mathbf{b}^* = (b_0^*, b_1^*, ..., b_p^*)^{\mathrm{T}}$ is the estimation for $\mathbf{b}$.

Now we will prove that $\boldsymbol{\beta} = \mathbf{b}^* = (b_0^*, b_1^*, ..., b_p^*)^{\mathrm{T}}$ is a solution for the system (4.2). For this, we have $\mathbf{w}^* = \sum_{j=0}^{p} b_j^*\mathbf{x}_{j+1} = \mathbf{X}^{\mathrm{T}}\mathbf{b}^*$. Thus $\mathrm{pr}_{\mathbf{V}}\mathbf{y} = \mathbf{w}^* \Leftrightarrow \mathbf{y} - \mathbf{w}^* \perp V \Leftrightarrow \mathbf{x}_k$ is orthogonal on $\mathbf{y} - \mathbf{w}^*, k = \overline{1, p+1}$. This is equivalent with $\mathbf{x}_k^{\mathrm{T}}(\mathbf{y} - \mathbf{w}^*) = 0, k = \overline{1, p+1}$ so is obtained $\mathbf{X}(\mathbf{y} - \mathbf{X}^{\mathrm{T}}\mathbf{b}^*) = 0$. Consequently, $\mathbf{X}\mathbf{X}^{\mathrm{T}}\mathbf{b}^* = \mathbf{X}\mathbf{y}$ and $\mathbf{b}^*$ satisfies the system (4.2).

We conclude that the estimators b_j^* always exists, they verifies the relations (4.2). Reciprocally, any solution for (4.2) which depends only by $\mathbf{y}$ is an estimation for $\mathbf{b}$.

Note: A simple case appears when $\mathrm{rank}\mathbf{X} = p + 1$. Then $\mathrm{rank}\mathbf{A} = \mathrm{rank}\mathbf{X}\mathbf{X}^{\mathrm{T}} = p + 1$. Therefore $\det\mathbf{A} \neq 0$ and the matrix $\mathbf{A}$ is nonsingular. In this case, the system (4.2) has a unique solution, namely $\mathbf{b}^* = \mathbf{A}^{-1}\mathbf{X}\mathbf{y}$.

5. NUMERICAL EXAMPLE

We consider a model when the output fuzzy data, Y, depends by two input, independent fuzzy data, X_1, X_2. Suppose that we have the following three available experimental measurements:

$$X_{11} = (1, 1/2, 1/2);\ X_{12} = (1, 1/4, 1/3);\ Y_1 = (1, 3/4, 1/2)$$

$$X_{21} = (3, 1, 2);\ X_{22} = (2, 1, 1);\ Y_2 = (2, 1, 1)$$

$$X_{31} = (2, 1, 1/3);\ X_{32} = (1, 2/3, 1/2);\ Y_3 = (1, 4/5, 4/5)$$

Thus $p = 2, n = 3 : Y_i = b_0 + b_1X_{i1} + b_2X_{i2}, i = \overline{1,3}$. Next, are computed the coefficients that appear in the proposed model:

$$\begin{gathered}\underline{a}_{11} = 1/2;\ \overline{a}_{11} = 3/2;\ x_{11} = 1;\ \underline{a}_{12} = 3/4;\ \overline{a}_{12} = 4/3;\ x_{12} = 1;\ \underline{c}_1 = 1/4;\ \overline{c}_1 = 3/2;\ y_1 = 1 \\ \underline{a}_{21} = 2;\ \overline{a}_{21} = 5;\ x_{21} = 3;\ \underline{a}_{22} = 1;\ \overline{a}_{22} = 3;\ x_{22} = 2;\ \underline{c}_2 = 1;\ \overline{c}_2 = 3;\ y_2 = 2 \\ \underline{a}_{31} = 1;\ \overline{a}_{31} = 7/3;\ x_{31} = 2;\ \underline{a}_{32} = 1/3;\ \overline{a}_{32} = 3/2;\ x_{32} = 1;\ \underline{c}_3 = 1/5;\ \overline{c}_3 = 9/5;\ y_3 = 1\end{gathered} \tag{5.1}$$

Four cases are possible. From the general form (4.2) we obtain, in each case, the following system of equations with their solutions:

$$1)\, b_1 > 0, b_2 > 0 : \begin{cases} 12b_0 + 24.33b_1 + 15.91b_2 = 15.75 \\ 24.33b_0 + 65.38b_1 + 41.2b_2 = 30.77 \Rightarrow b_0 = -0.65; b_1 = -6.07 < 0; b_2 = 10.76; \\ 15.91b_0 + 41.2b_1 + 26.7b_2 = 26.95 \end{cases}$$

$$2)\, b_1 > 0, b_2 < 0 : \begin{cases} 12b_0 + 24.33b_1 + 15.91b_2 = 15.75 \\ 24.33b_0 + 65.38b_1 + 33.06b_2 = 30.77 \\ 15.91b_0 + 33.06b_1 + 26.7b_2 = 20.35 \end{cases} \Rightarrow b_0 = 1.56; b_1 = -0.06 < 0; b_2 = 0.08;$$

$$3)\, b_1 < 0, b_2 > 0 : \begin{cases} 12b_0 + 15.91b_2 + 24.33b_1 = 15.75 \\ 15.91b_0 + 26.7b_2 + 33.06b_1 = 26.95 \\ 24.33b_0 + 33.06b_2 + 65.38b_1 = 32.39 \end{cases} \Rightarrow b_0 = -0.07; b_1 = -0.02; b_2 = 1.08;$$

$$4)\, b_1 < 0, b_2 < 0 : \begin{cases} 12b_0 + 24.33b_1 + 15.91b_2 = 15.75 \\ 24.33b_0 + 65.38b_1 + 41.2b_2 = 32.39 \\ 15.91b_0 + 41.2b_1 + 26.7b_2 = 20.35 \end{cases} \Rightarrow b_0 = 1.5; b_1 = 0.73 > 0; b_2 = -1.26;$$

The unique solution which fulfill the initial conditions concerning the sign of the parameters b_0, b_1, b_2 appears in the third case: $b_0 = -0.07; b_1 = -0.02; b_2 = 1.08.$ Thus Y depends by X through the relation: $Y = -0.07 - 0.02X_1 + 1.08X_2$.

REFERENCES

1. ARTHANARY T. S., YADOLAH DODGE, *Mathematical Programming in Statistics*, John Wiley and Sons, New York, 1980.
2. DIAMOND P., KLOEDEN P., *Metric spaces of fuzzy sets*, Fuzzy Sets and Systems **35**, pp. 241-249, 1990
3. DIAMOND P., *Fuzzy least squares*, Inform. Sci. **46**, pp. 141-157**,** 1988.
4. DIAMOND P., KLOEDEN P., *Metric spaces of fuzzy sets, Corrigendum*, Fuzzy Sets and Systems **45,** pp. 123, 1992.
5. GOETSCHEL R., VOXMAN W., *Elementary Calculus*, Fuzzy Sets and Systems **18,** pp. 31-43, 1986.
6. HAMMERBACHER I. M., YAGER R. R., *Predicting television revenues using fuzzy subsets*, TIMS Stud. Management Sci. **20**, pp. 469-477, 1984.
7. KATSARAS A., LIU D. B., *Fuzzy vector spaces and fuzzy topological vector spaces*, J. Math. Anal. Appl. **58**, pp. 135-146, 1977.
8. MIHOC G., URSEANU V., URSIANU E., *Modele de analiză statistică*, Ed. Ştiinţifică şi enciclopedică, Bucureşti, 1982.
9. MA MING, FRIEDMAN M., KANDEL A., *General fuzzy least squares*, Fuzzy Sets and Systems **88**, pp. 107-118, 1997.
10. NEGOIŢĂ C. V., RALESCU D. A., *Applications of Fuzzy Sets to Systems Analysis*, Wiley, New York, 1975.
11. PRADE H., Operations research with fuzzy data, in: P. P. Wang and S. K. Chang, Eds., Fuzzy sets: *Theory and Application to Policy Analysis and Information Systems*, plenum, New York,pp. 115-169, 1980.
12. PURI L.M., RALESCU D. A., *Differentials for fuzzy functions*, *J. Math. Anal. Appl.* **91**, pp. 552-558, 1983.
13. TANAKA H., ISIBUCHI H., YOSHIKAWA S., *Exponential possibility regression analysis*, Fuzzy Sets and Systems **69**, pp. 305-318, 1995.
14. TANAKA H., UEJIMA S., ASAI K., *Linear regression analysis with fuzzy model*, IEEE Trans. Systems Man Cybernet. **SMC-12**, pp. 903-907, 1992.
15. WU CONG-XIN, MA MING, *Embedding problem of fuzzy number space: Part I*, Fuzzy Sets and Systems **44**, pp. 33-38,1991.
16. WU CONG-XIN, MA MING, *Embedding problem of fuzzy number space: Part III*, Fuzzy Sets and Systems **46**, pp. 281-286, 1992
17. YAGER R. R., *Fuzzy prediction based upon regression models*, Inform. Sci. **26**, pp. 45-63, 1982.
18. ZIMMERMANN H. J., *Fuzzy programming and linear programming with several objective functions*, Fuzzy Sets and Systems **1**, pp. 45-55, 1978.

Received October 3, 2006