Annales Univ. Sci. Budapest., Sect. Comp. 12 (1991) 121-126

POSSIBILISTIC FUZZY LINEAR REGRESSION

Sándor József
Research Institute for Agricultural Economics
Zsil u. 3–5, Budapest, H–1093, Hungary

Abstract: In the paper some results on data transformation effects in the possibilistic fuzzy regression are given and a connection between possibilistic "fixed effect" and fuzzy TR-regression model is analysed.

Keywords: Possibilistic regression, linear data transformation, fuzzy TR-regression.

In the practical modelling there are new approaches and methods given by the fuzzy theory. First of all we can mention the several regression–type fuzzy models, extensions of probabilistic methods and the "clear" fuzzy approaches of this problem. In the modelling procedures there are always a lot of questions and/or problems to formalize and to solve.

Firstly we must choose what model element has fuzzy property in the real process modelled by fuzzy methods. If *basic set of observation* is a fuzzy set, all the statistical methods can be applied on this ([2]); if the *connection* between some variables is given by a fuzzy relation, we can use the fuzzy TR–regression method ([4], [5]); if *the observations* are fuzzy we can fit a nonfuzzy function to these ([6]) or if *the observations and some parameters* are fuzzy we can use a special (for example: linear) type of function in the modelling (for example: the possibilistic approaches of B. Heshmaty–A. Kandel [1] and H. Tanaka–J. Watada [8]) and so on.

Secondly there are some other basic questions to answer: What are the criterions of *the good fitness* in our fuzzy model? What *type of error* is to minimize, how can it be measured or quantitified, what type of function is used? How can the results of the several methods be comparised and chosen the "best" model, or what properties can be expected from a fuzzy regression–type model?

The answers to these questions are in general not given neither in the fuzzy theory nor in the practical fuzzy modelling. But any of these can be important in the practical modelling, for example there is a special problem, the effect of linear data transformations in the possibilistic models.

It is well–known that in the classical linear regression model based on the probability theory, the measure of the goodness of fit (R^2) is invariant in respect of linear data transformations of the independent variables x_k, $k = 1, \dots, K$. In my opinion we would expect this property in the fuzzy case, too, because the comparison of various models, computing for the same dependent variable is possible only like this. From this point of view the possibilistic approach of fuzzy linear regression problem has the same basic methodological problems.

The possibilistic fuzzy extensions of linear regression model are based on the fuzzification of the linear function $l(\mathbf{x}) = A_1x_1 + \cdots + A_Kx_K$ by the extension principle using g–type fuzzy numbers as function parameters. The observations $\widetilde{y}_i = (y_i, e_i)_g$ of the y variable are approximated by the fuzzified function values denoted by $\widetilde{l}(\mathbf{x}_i, \widetilde{\mathbf{a}})$. It is well known that in this case

$$\mu_{\widetilde{l}(\mathbf{x},\widetilde{\mathbf{a}})}(y) = \begin{cases} g^{(-1)}\left(\frac{|y-\mathbf{a}^*\mathbf{x}|}{|\mathbf{c}^*\mathbf{x}|}\right) & \text{if } |y - \mathbf{a}^*\mathbf{x}| \leq |\mathbf{c}^*\mathbf{x}| \neq 0; \\ 0 & \text{otherwise,} \end{cases}$$

where $\widetilde{a}_k = (a_k, c_k)_g$ $(k = 1, \dots, K)$ are g–type fuzzy numbers, $\mathbf{a} = (a_1, \dots, a_K)^*$ is the vector of the centers and $\mathbf{c} = (c_1, \dots, c_K)^*$ is the vector of the widths. Let $\widetilde{\mathbf{a}} = (\widetilde{a}_1, \dots, \widetilde{a}_K)^*$ be the vector of the fuzzy parameters. In the possibilistic models the "goodness of approximation" at the point $\mathbf{x}_i$ is represented by the value

$$h_i \stackrel{\text{def}}{=} \max\left\{ h : \{ y : \mu_{\widetilde{y}_i}(y) \geq h \} \subset \{ y : \mu_{\widetilde{l}(\mathbf{x}_i,\widetilde{\mathbf{a}})}(y) \geq h \} \right\}.$$

The fuzzy regression problem defined by Heshmaty and Kandel, i.e. Tanaka and Watada (in the following: HK–, i.e. TW–problem) can be formulated as follows:

Let the fuzzy observation $\widetilde{y}_i$ of the dependent variable y at the point $\mathbf{x}_i$ be given (naturally $\mathbf{x}_i$ is also given, $i = 1, \dots, N$) and let $0 \leq H < 1$ be a fixed niveau (aspiration level). In the HK–problem we try to find the fuzzy parameters $\widetilde{a}_k$ $(k = 1, \dots, K)$, so that $\mathbf{c} \geq 0$, $h_i \geq H$ $(i = 1, \dots, N)$ and $S = \sum_{k=1}^{K} c_k = \min$, i.e. the minimum of the sum of widths of fuzzy parameters is claimed; in the TW–model the minimization criterion is: $\sum_{i=1}^{N} \mathbf{c}^*|\mathbf{x}_i| \longrightarrow \min$, i.e. we minimize the sum of widths of fuzzy function values.

In these models the assumption $h_i \geq H$ $(i = 1, \dots, N)$ may be rewritten in the form

$$\mathbf{a}^*\mathbf{x}_i - g(H)\left(\mathbf{c}^*|\mathbf{x}_i| - e_i\right) \leq y_i \leq \mathbf{a}^*\mathbf{x}_i + g(H)\left(\mathbf{c}^*|\mathbf{x}_i| - e_i\right) \qquad (i = 1, \dots, N)$$

and the problems are linear programming problems.

In the following we will suppose that $\tilde{a}_K$ is the fuzzy constant parameter, i.e. $x_{Ki} = 1$ for all observations. In the H–K case three lemmas are proved.

Lemma 1. *Let $\tilde{\mathbf{a}}$ be the HK-solution of the fuzzy linear regression problem. When two indices k_0, j_0 and a value $L > 1$ exist so that $|x_{k_0i}| \geq L|x_{j_0i}| \quad (i = 1, \ldots, N)$, than $c_{j_0} = 0$.* □

Proof. Let $\tilde{\tilde{\mathbf{a}}} = ((\bar{a}_1, \bar{c}_1), \ldots, (\bar{a}_K, \bar{c}_K))^*$ be a new fuzzy vector of parameters, where $\bar{\mathbf{a}} = (\bar{a}_1, \ldots, \bar{a}_K) \stackrel{\text{def}}{=} \mathbf{a}$ and

$$\bar{c}_k \stackrel{\text{def}}{=} \begin{cases} c_k & \text{if } k \neq k_0, j_0; \\ 0 & \text{if } k = j_0; \\ c_{k_0} + \frac{c_{j_0}}{L} & \text{if } k = k_0 \end{cases} \qquad (k = 1, \ldots, K).$$

It can be easily shown [3], that this parameter vector is an admissible solution of the HK–problem. The sum of widths of the new parameters is

$$\sum_{k=1}^{K} \bar{c}_k = \Big(\sum_{k \neq k_0, j_0} c_k \Big) + \Big(c_{k_0} + \frac{c_{j_0}}{L}\Big) + 0 \leq \sum_{k=1}^{K} c_k,$$

and the equality is true if and only if $c_{j_0} = 0$. The relation $\sum_{k=1}^{K} c_k \leq \sum_{k=1}^{K} \bar{c}_k$ is trivial, since $\tilde{\mathbf{a}}$ is the solution of HK–problem. So the relation $\tilde{\mathbf{a}} = \tilde{\tilde{\mathbf{a}}}$ is true. ■

This lemma says, that in the HK–model *the character of parameters, their fuzzy or not fuzzy property* depends, for example, on the current unit of measurement of each independent variable! Namely, when a variable has only positive or negative values, with the change of the unit of its measurement can a "new variable" be made in trivial way, so that this new variable has the "majority property" for all other independent variables (in the sense of the lemma) and consequently, all the other variables have nonfuzzy regression parameter!

Lemma 2. *Let us suppose that $x_{k_0i} \neq 0 \quad (i = 1, \ldots, N)$ for the same fixed index $1 \leq k_0 \leq K - 1$. In this case we can choose the unit of measurement of this variable so, that the sum of widths of the parameters should be discretionarily small.* □

Lemma 3. *The HK–model of fuzzy regression problem is solvable with discretionarily small sum of parameter widths by moving an independent variable.* □

The lemmas prove that in the HK–modell the linear data transformation has a crucial effect on the fuzzy property of model parameters. Probably this is the reason, that in the numerical examples presented in the author's work only one variable has fuzzy parameter, all the others have nonfuzzy ones. To prove these lemmas is also sufficient to construct admissible solutions with the properties given in the lemmas, and similar to the lemma 1 they are easily constructed [3].

We emphasize that this property gives us no basis for comparison of various models, computing for the same dependent variable: based on the sum of parameter widths we can not choose the "best" model. With the use of special "standardized" independent variables this effect may be eliminated in a certain sense but not correctly:

In the TW-model the objective function is $\sum_{i=1}^{N} \mathbf{c}^*|\mathbf{x}_i|$ and when we use the function $g(u) = 1 - u$ and the "standardized" independent variables defined by

$$\bar{x}_{ki} \stackrel{\text{def}}{=} \frac{x_{ki}}{\sum_{i=1}^{N} |x_{ki}|} \quad (k = 1, \dots, K),$$

so we have a HK-model, because with these variables $\sum_{i=1}^{N} \mathbf{c}^*|\overline{\mathbf{x}}_i| = \mathbf{c}^* \sum_{i=1}^{N} |\overline{\mathbf{x}}_i| = \mathbf{c}^*\mathbf{1} = \sum_{k=1}^{K} c_k$. The standardized variables $\overline{\mathbf{x}}_k$ are independent from the unit of measurement, moreover, variable with "majority property" does not exist. However, the linear data transformation "moving" has an effect in the TW-model, too, but this effect is smaller than in the HK-model. Let be $\mathbf{d} = (d_1, \dots, d_K = 0)^*$. So the following lemma is true [3]:

Lemma 4. *Let us use the variables $x_k - d_k$ in the TW-model! In this case the sum of the widths of the function values $S(\mathbf{d}) = \mathbf{c}^* \sum_{i=1}^{N} |\overline{\mathbf{x}}_i - \mathbf{d}|$ will be the smallest, when*

$$\min_i x_{ki} \le d_k \le \max_i x_{ki} \quad \forall k \in \{1, \dots, K-1\}. \qquad \square$$

This lemma says that when a variable has only positiv values, i.e. its increase causes the increase of the "error" (the width of the fuzzificated function value), then we can construct a not worse solution, when we move the minimum value of this variable to zero; and if a variable has only negative values, i.e. its increase causes the decrease of the approximation error, then it is possible to construct a not worse solution with the moving its maximum value to zero. The best solution can be found in the case of such moving d_k of the variables, for which the inequalities $\min_i x_{ki} \le d_k \le \max_i x_{ki}$ are true. Consequently, in this case we try to find a model with first decreasing then increasing width of fitted value. As marginal cases ($d_k = \min_i x_{ki}$, i.e. $d_k = \max_i x_{ki}$) we get back the models with only increasing, i.e. decreasing width of fitted value. If we also claim to find the best moving of variable, this means the transform of TW linear programming model into a parametrical LP model or into a nonlinear programming problem. But in this way the goodness of the various solutions characterized by $S(\mathbf{d}) = \mathbf{c}^* \sum_{i=1}^{N} |\mathbf{x}_i - \mathbf{d}|$ may be compared.

We note that there is a very important special possibilistic model class. Namely the proved bad effects don't occur in these models, when only the constant of the linear function is a fuzzy number and all the other parameters are nonfuzzy, because in this case the width of the fuzzified function value is the same for all observations. This case can be called "fixed effect" case.

What connection can be found between the "practical" possibilistic approach and the "theoretical" fuzzy TR–regression which is defined as the solution(s) of the equation

$$R = R_X \cap r, \qquad R, r \in \mathcal{F}(\mathbb{R}^K \times \mathbb{R}) \quad \text{i.e.}$$

$$\mu_R(\mathbf{x}, y) = T(\mu_{R_X}(\mathbf{x}), \mu(\mathbf{x}, y)) \quad \forall (\mathbf{x}, y) \in \mathbb{R}^{K+1}$$

where the (given) fuzzy relation R characterizes the fuzzy connectivity of the observed variables and R_X is the projection of R? We present a relation class for which the fuzzy TR–regression leads uniquely to the possibilistic fuzzy "fixed effect" model. Let T be a strong t–norm with the generator function g^2 and let be

$$\begin{pmatrix} \mathbf{u} \\ v \end{pmatrix} = \begin{pmatrix} C & \mathbf{d} \\ \mathbf{e}^* & f \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ y \end{pmatrix},$$

moreover $\tilde{v}$, $\tilde{u}_j \in \mathcal{F}(\mathbb{R})$, $j = 1, \ldots, K$ g–type fuzzy numbers. We suppose that our basic relation R is given by the T–intersection of these fuzzy numbers:

$$\mu_R(\mathbf{x}, y) = \mu_{\tilde{R}}(\mathbf{u}, v) = T_{g^2}(\mu_{\tilde{u}_1}(u_1), \ldots, \mu_{\tilde{u}_M}(u_M), \mu_{\tilde{v}}(v)).$$

It can be rewritten like this:

$$\mu_R(\mathbf{x}, y) = g^{(-1)}((\mathbf{u}^*\mathbf{u} + v^2)^{\frac{1}{2}}) =$$

$$= g^{(-1)}\left(((\mathbf{x}^*C^* + y\mathbf{d}^*)(C\mathbf{x} + \mathbf{d}y) + (\mathbf{x}^*\mathbf{e} + fy)^2)^{\frac{1}{2}}\right) =$$

$$= g^{(-1)}\left((\mathbf{x}^*C^*C\mathbf{x} + 2\mathbf{d}^*C\mathbf{x}y + \mathbf{d}^*\mathbf{d}y^2 + \mathbf{x}^*\mathbf{e}\mathbf{e}^*\mathbf{x} + 2f\mathbf{e}^*\mathbf{x}y + f^2y^2)^{\frac{1}{2}}\right) =$$

$$= g^{(-1)}\left((\mathbf{x}^*(C^*C + \mathbf{e}\mathbf{e}^*)\mathbf{x} + 2(\mathbf{d}^*C + f\mathbf{e}^*)\mathbf{x}y + (\mathbf{d}^*\mathbf{d} + f^2)y^2)^{\frac{1}{2}}\right) =$$

$$= g^{(-1)}\left((\mathbf{x}^*(C^*C + \mathbf{e}\mathbf{e}^*)\mathbf{x} + (by + \mathbf{a}^*\mathbf{x})^2 - (\mathbf{a}^*\mathbf{x})^2)^{\frac{1}{2}}\right),$$

where

$$b \stackrel{\text{def}}{=} (\mathbf{d}^*\mathbf{d} + f^2)^{\frac{1}{2}} \ (\geq 0), \qquad \mathbf{a} \stackrel{\text{def}}{=} \frac{\mathbf{d}^*C + f\mathbf{e}^*}{(\mathbf{d}^*\mathbf{d} + f^2)^{\frac{1}{2}}}.$$

It is easy to show that

$$\mu_{R_X}(\mathbf{x}) = g^{(-1)}\left((x^*(C^*C + \mathbf{e}\mathbf{e}^* - \mathbf{a}^*\mathbf{a})\mathbf{x})^{\frac{1}{2}}\right)$$

and so our basic equation has the following form

$$\mu_R(\mathbf{x}, y) = T_{g^2}\left(\mu_{R_X}(\mathbf{x}), g^{(-1)}(|\mathbf{a}^*\mathbf{x} + by|)\right).$$

It can be easily shown that the solution is unique and its membership function is:

$$\mu_r(\mathbf{x}, y) = g^{(-1)}(|\mathbf{a}^*\mathbf{x} + by|),$$

i.e. the fuzzy TR–regression r for all fixed $\mathbf{x}$ values is a g–type fuzzy number. The centers belong to the line $\frac{-\mathbf{a}^*\mathbf{x}}{b}$ and the widths are constant $(\frac{1}{b})$. The possibilistic model form of this result is

$$r = \frac{-\mathbf{a}^*\mathbf{x}}{b} + \left(y; 0, \frac{1}{b}\right)_g = \mathbf{a}'^*\mathbf{x} + \widetilde{a}'_0 = \widetilde{l}(\mathbf{x}, \widetilde{\mathbf{a}}'),$$

where $\mathbf{a}'$ is a vector of crisp numbers, $\widetilde{a}'_0$ is a g–type fuzzy number and $\widetilde{\mathbf{a}}' = (\mathbf{a}', \widetilde{a}'_0)^*$.

REFERENCES

[1] B. Heshmaty – A. Kandel, Fuzzy linear regression and its applications to forecasting in uncertain environment, *Fuzzy Sets and Systems* **15**(1985) 159–191.

[2] K. Jajuga, Linear Fuzzy Regression, *Fuzzy Sets and Systems* **20**(1986) 343–353.

[3] S. József, On the effect of linear data transformations in the possibilistic fuzzy linear regression, Working Paper, University of Economics, Budapest, 1989.

[4] S. József, Über das Fuzzy–Regressionsproblem, *Annales Univ. Sci. Budapest, Sect. Comp.* 9.(1988) 171–184.

[5] S. József, On the fuzzy TR–regression, BUSEFAL 36(1988), 92–97.

[6] S. József et al, Soft Econometrical Modelling. The example of fuzzy methods, Working Paper,University of Economics, Budapest, 1989.

[7] H. Tanaka, Fuzzy data analysis by possibilistic linear models, *Fuzzy Sets and Systems* **24**(1987) 363–375.

[8] H. Tanaka – J. Watada, Possibilistic linear systems and their application to the linear regressionmodel, *Fuzzy Sets and Systems* **27**(1988) 275–289.