RATIO MATHEMATICA
ISSUE N. 30 (2016) pp. 45-58

ISSN (print): 1592-7415
ISSN (online): 2282-8214

Dealing with randomness and vagueness in business and management sciences: the fuzzy-probabilistic approach as a tool for the study of statistical relationships between imprecise variables

Fabrizio Maturo
Department of Management and Business Administration
University G. d'Annunzio, Chieti - Pescara
f.maturo@unich.it

Abstract

In practical applications relating to business and management sciences, there are many variables that, for their own nature, are better described by a pair of ordered values (i.e. financial data). By summarizing this measurement with a single value, there is a loss of information; thus, in these situations, data are better described by interval values rather than by single values. Interval arithmetic studies and analyzes this type of imprecision; however, if the intervals has no sharp boundaries, fuzzy set theory is the most suitable instrument. Moreover, fuzzy regression models are able to overcome some typical limitation of classical regression because they do not need the same strong assumptions. In this paper, we present a review of the main methods introduced in the literature on this topic and introduce some recent developments regarding the concept of randomness in fuzzy regression.

Keywords: fuzzy data; fuzzy regression; fuzzy random variable; tools for business and management sciences

2010 AMS subject classifications: 62J05; 62J86; 03B52; 62A86; 97M10

doi: 10.23755/rm.v30i1.8

1 Introduction

Regression analysis offers a possible solution to study the dependence between two sets of variables. Standard classical statistical linear regressions take the form [27]:

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + ... + b_j x_{ij} + ... + b_P x_{iP} + u_i \tag{1}$$

where:

- i=1,.....,N is the i-th observed unit;
- j=1,...,P is the j-th observed variable;
- y_i is the dependent variable, observed on N units;
- x_{ij} are the P independent variables observed on N units;
- b_0 is the crisp intercept and b_j are the P crisp coefficients of the P variables;
- u_i are the random error terms that indicate the deviation of Y from the model;
- y_i, x_{ij}, b_j, u_i are all crisp values.

In classical regression model it is assumed that:

- $E(u_i) = 0$
- $\sigma^2_{u_i} = \sigma^2$
- $\sigma_{u_i,u_j} = 0 \quad \forall \quad i,j \quad with \quad i \neq j$

In matrix form, the classical regression model is expressed as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{u} \tag{2}$$

where $\mathbf{y} = (\mathbf{y_1}, \mathbf{y_2}, ..., \mathbf{y_N})'$, $\mathbf{b} = (\mathbf{b_1}, \mathbf{b_2}, ..., \mathbf{b_P})'$, $\mathbf{u} = (\mathbf{u_1}, \mathbf{u_2}, ..., \mathbf{u_N})'$ are vectors and $\mathbf{X}$ is a matrix:

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & . & . & . & x_{1P} \\ 1 & x_{21} & . & & . & x_{2P} \\ 1 & . & . & . & . & . \\ 1 & . & . & . & . & . \\ 1 & x_{N1} & . & . & . & x_{NP} \end{pmatrix}$$

The aim of statistical regression is to find the set of unknown parameters so that the model gives is a good prediction of the dependent variable Y. The most widely used regression model is the Multiple Linear Regression Model (MLRM), as well as the Ordinary Least Squares (OLS) [12] is the most widespread estimation procedure. Under the OLS assumptions the estimates are BLUE (Best Linear Unbiased Estimator), as stated by the famous Gauss-Markov theorem.

OLS is based on the minimization of the sum of squared deviations:

$$\min \qquad (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} + \mathbf{X}\mathbf{b}) \tag{3}$$

The optimal solution of the minimization problem is the following vector:

$$\widehat{b} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{y} \tag{4}$$

The OLS model is comfortable but its assumptions are every restrictive. Several phenomena violate these assumptions causing biased and inefficient estimators [9]. In particular the assumptions $E(\mathbf{u}|\mathbf{X}) \approx \mathbf{N}(\mathbf{0}, \sigma^{\mathbf{2}}\mathbf{I})$ is very strong and rarely it is respected in real phenomena. Moreover in case of "quasi" multi-collinearity (many highly correlated explanatory variables), although this does not violate OLS assumption there is a bad impact on the variance of B. In these circumstance the OLS estimators are efficient and unbiased but have large variance, making estimation useless from a practical point of view.

The effects of the quasi multi-collinearity are more evident when the sample size is small [1]. The generally proposed solution consists in removing correlated exploratory variables. This solution is unsatisfying in many applications fields where the user would keep all variables in the model.

In general, we can observe that classical statistical regression has many useful applications but presents troubles in the following situations [26]:

- Number of observations is inadequate (small data set);
- Difficulties verifying distribution assumptions;
- Vagueness in the relationship between input and output variables;
- Ambiguity of events or degree to which they occur;
- Inaccuracy and distortion introduced by linearization;

Furthermore, there are many variables that, for their own nature, are better described by a pair of ordered values, like daily temperatures or financial data. By summarizing this measurement with a single value, there is a loss of information. In these situations data are better described by interval values rather than by single values. Interval arithmetic studies and analyzes this type of imprecision; but if the intervals has no sharp boundaries, fuzzy set theory is the better tool. In particular fuzzy regression model are able to overcome some typical limitation of classical regression because they don't need the same strong assumptions. Furthermore, some nuanced concepts that exist in economic and social sciences, need to be necessarily treated with linguistic variables, which for their nature, are imprecise concepts.

2 Fuzzy Linear Regression Models (FLR)

There are two general ways, not mutually exclusive, to develop a fuzzy regression model:

- Models where the relationship of the variables is fuzzy;
- Models where the variables themselves are fuzzy;

Therefore fuzzy linear regression (FLR) can be classified in:

- Partially fuzzy linear regression (PFLR), that can be further divided into:
  - PFLR with fuzzy parameters and crisp data;
  - PFLR with fuzzy data and crisp parameters;
- Totally fuzzy linear regression (TFLR) where data and parameters are both fuzzy.

Fuzzy Least Squares Regression is more close to the traditional statistical approach. In fact, following the Least Squares line of thought [13], the aim is to minimize the distance between the observed and the estimated fuzzy data. This approach is referred as Fuzzy Least Squares Regression (FLSR).

In case of one independent variable, the model take the form:

$$\widetilde{y_i} = b_0 + b_1\widetilde{x_i} + \widetilde{u_i} \qquad \text{i=1,2,...,N} \tag{5}$$

where:

- i=1,.....,N is the i-th observed unit;
- y_i is the dependent fuzzy variable, observed on N units;
- x_i is the independent fuzzy variable, observed on N units;

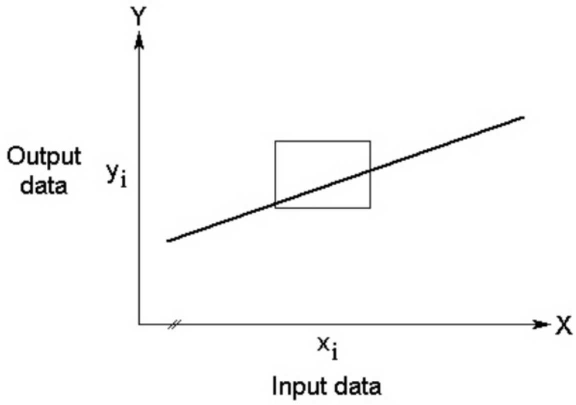

Figure 1: Relation between output and input variables

- b_0 and b_1 are the crisp intercept and the crisp regression coefficient;
- u_i are the fuzzy random error terms;

From a graphical point of view [26] the relation between output and input variables can be represented as shown in Fig.1

In case of several independent variables, the model take the form:

$$\widetilde{y_i} = b_0 + b_1\widetilde{x_{i1}} + b_2\widetilde{x_{i2}} + ... + b_j\widetilde{x_{ij}} + ... + b_P\widetilde{x_{iP}} + \widetilde{u_i} \qquad (6)$$

where:

- i=1,.....,N is the i-th observed unit;
- j=1,...,P is the j-th observed variable;
- y_i is the dependent fuzzy variable, observed on N units;
- x_{ij} are the P independent fuzzy variables, observed on N units;
- b_0 is the crisp intercept and b_j are the P crisp regression coefficients measured for the P fuzzy variables;
- u_i are the fuzzy random error terms;

Limiting the reasoning to the first model, the error term can be expressed as follows:

$$\widetilde{u_i} = \widetilde{y_i} - b_0 - b_1\widetilde{x_i} \qquad \text{i=1,2,...,N} \qquad (7)$$

Therefore, from a least square perspective, the problem becomes as follows:

$$\min \sum_{i=1}^{N} [\widetilde{y_i} - b_0 - b_1 \widetilde{x_i}]^2 \qquad \text{i=1,2,...,N} \tag{8}$$

Many criteria for measuring this distance have been proposed over the years; however, the most common are two methods:

- The Diamond's approach;
- The compatibility measures approach.

2.1 FLSR using distance measures

The Diamond's approach is also known as fuzzy least squares regression using distance measures. This is the most close approach to the traditional statistical one. Following the Least Squares line of thought, the aim is to minimize the distance between the observed and the estimated fuzzy data, by minimizing the output quadratic error of the model. Since the model contains fuzzy numbers the minimization problem considers distances between fuzzy numbers [5, 17, 20, 15, 19, 18].

Diamond defined an L^2-metric between two triangular fuzzy numbers; it measures the distance between two fuzzy numbers based on their modes, left spread and right spread as follows

$$\begin{aligned} d[(c_1, l_1, r_1), (c_2, l_2, r_2)]^2 = \\ = (c_1 - c_2)^2 + [(c_1 - l_1) - (c_2 - l_2)]^2 + [(c_1 + r_1) - (c_2 + r_2)]^2 \end{aligned} \tag{9}$$

The methods of Diamond are rigorously justified by a projection-type theorem for cones on a Banach space containing the cone of triangular fuzzy numbers, where a Banach space is a normed vector space that is complete as a metric space under the metric $d(x, y) = ||x - y||$ induced by the norm [25].

In the case of crisp coefficients and fuzzy variables, the problem is the following:

$$\min \quad \sum_{i=1}^{N} d[\widetilde{y_i}^* - \widetilde{y_i}]^2 \qquad \text{i=1,2,...,N} \tag{10}$$

where,

$$\widetilde{y_i}^* = b_0 + b_1 \widetilde{x_i} \tag{11}$$

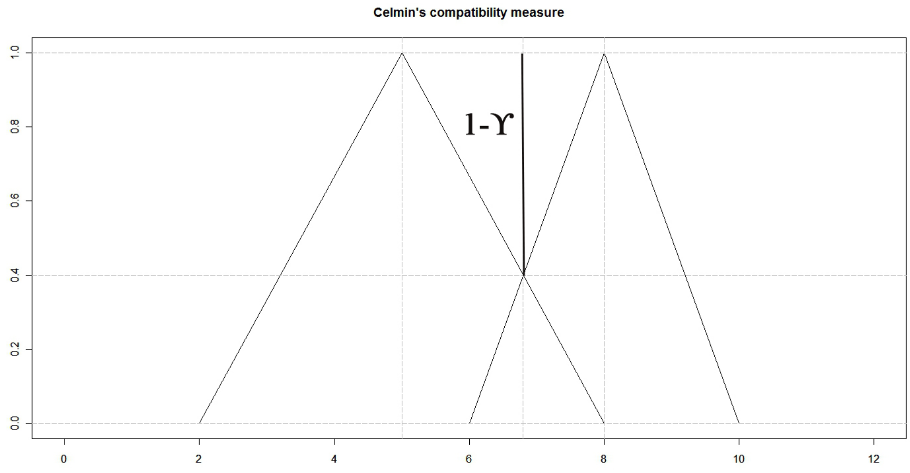

Figure 2: Compatibility measure

therefore the optimization problem can be written as follows:

$$\min \quad \sum_{i=1}^{N} d[b_0 + b_1\widetilde{x_i} - \widetilde{y_i}]^2 \qquad \text{i=1,2,...,N} \tag{12}$$

Using Diamond's difference in this minimization problem, we can obtain the parameters. If the solutions exist, it is necessary to solve a system of six equations in the same number of unknowns; of course, these equations arise from the derivatives being set equal to zero.

2.2 FLSR using compatibility measures

The second type of fuzzy least squares regression model is based on Celmins's compatibility measures [3]. A compatibility measure can defined by

$$\gamma(\widetilde{A}, \widetilde{B}) = maxmin(\mu_A(x), \mu_B(x)) \tag{13}$$

This index is included in the interval [0,1] as shown in Fig. 2. A value of ”0” means that the membership functions of the fuzzy numbers A and B are mutually exclusive as shown in Fig. 3. A value of ”1” means that the membership functions A coincides with that one of B as shown in Fig.4.

The basic idea is to maximize the overall compatibility between data and model. Thus, the objective may be reformulated in a minimization problem with the following objective function:

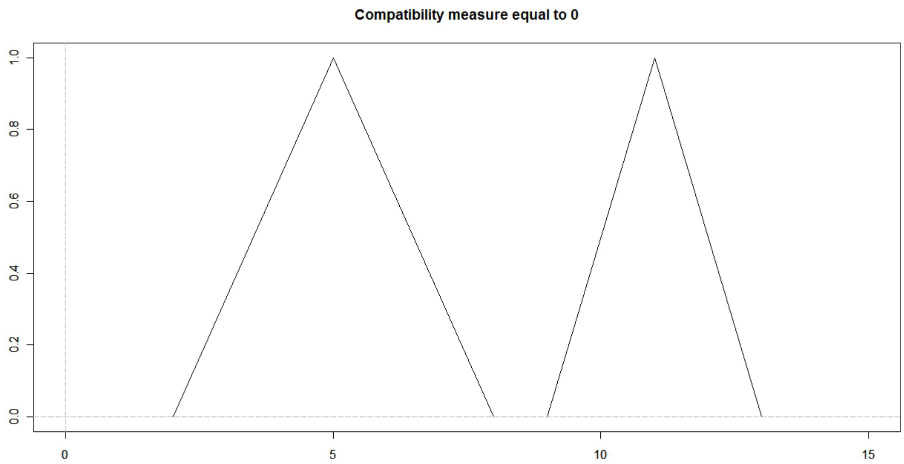

Figure 3: Zero compatibility

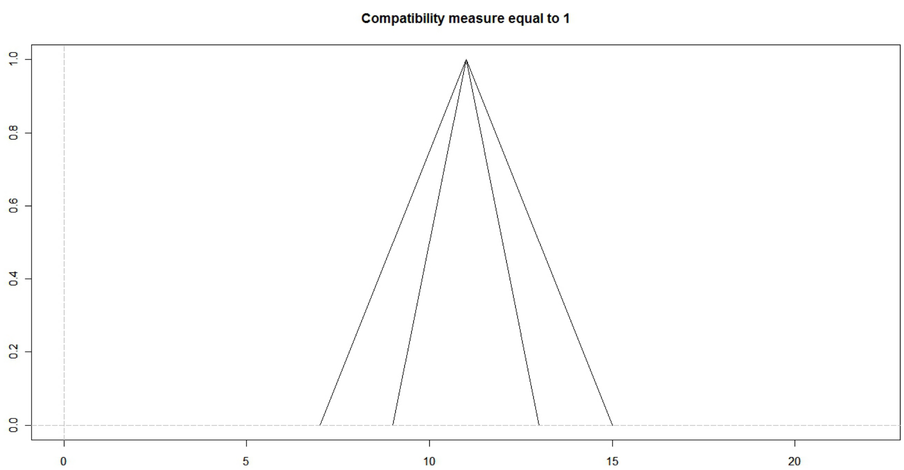

Figure 4: Max compatibility

$$\min \quad \sum_{i=1}^{N}[1-\gamma_i]^2 \qquad \text{i=1,2,...,N} \tag{14}$$

3 Fuzzy regression models with fuzzy random variables

Recent studies have reintroduced the concept of Fuzzy Random Variables (FRVs) [24] firstly introduced by Puri and Ralescu [23]. The need for FRVs arises when the data are not only affected by imprecision but also by randomness [11]. Several papers deal with this topic that it is called fuzzy-probabilistic approach. It consists in explicitly taking into account randomness for estimating the regression parameters and assessing their statistical properties [22, 7, 8].

The membership function of a fuzzy number can be expressed, in term of spreads as:

$$\mu_{\widetilde{A}}(x) = \begin{cases} L\frac{A_m - x}{A_l} & \text{for} \quad x \leq A_m, \quad A_l > 0 \\ 1 & \text{for} \quad x \leq A_m, \quad A_l = 0 \\ R\frac{x - A_m}{A_r} & \text{for} \quad x > A_m, \quad A_r > 0 \\ 0 & \text{for} \quad x > A_m, \quad A_r = 0 \end{cases} \tag{15}$$

where the functions L, R : $\Re - > [0,1]$ are convex upper semi-continuous functions so that L(0) = R(0) = 1 and L(z) = R(z) = 0, for all $z \in \Re/[0,1]$ [6] and A_m is the center, A_l and A_r are the left and the right spread. Of course these functions must be chosen by the researcher in advance and must be the same for all the data.

In particular, for a triangular fuzzy number we obtain:

$$\mu_{\widetilde{A}}(x) = \begin{cases} 0 & \text{for} \quad x \leq A_m - A_l \\ 1 - \frac{A_m - x}{A_l} & \text{for} \quad A_m - A_l \leq x \leq A_m \\ 1 - \frac{x - A_m}{A_r} & \text{for} \quad A_m \leq x \leq A_m + A_r \\ 0 & \text{for} \quad x \geq A_m + A_r \end{cases} \tag{16}$$

A distance for these functions [21] could be:

$$D^2(\widetilde{A},\widetilde{B}) = (A_m - B_m)^2 + [(A_m - \lambda A_l) - (B_m - \lambda B_l)]^2 + [(A_m + \rho A_r) - (B_m + \rho B_r)]^2 \tag{17}$$

where,

$$\lambda = \int_0^1 L^{-1}(\alpha) d\alpha$$

$$\rho = \int_0^1 R^{-1}(\alpha) d\alpha$$

These functions consider the shape of the membership functions; for example, for triangular fuzzy numbers λ and ρ = 1/2.

To avoid the problem of the non-negativity of the spreads of $\widetilde{Y}$, it is possible to solve a non negative regression problem [14], or to transform the spreads of $\widetilde{Y}$ by means of the centers and the spreads of the P regressors X. In this context, we use the latter method introducing two invertible functions [7]:

$$g : (0, +\infty) \longrightarrow \Re$$

$$h : (0, +\infty) \longrightarrow \Re$$

Thus the linear regression model take the form

$$\begin{cases} Y_m = \mathbf{x}\mathbf{b}'_\mathbf{m} + \mathbf{a_m} + \mathbf{u_m} \\ g(Y_l) = \mathbf{x}\mathbf{b}'_\mathbf{l} + \mathbf{a_l} + \mathbf{u_l} \\ h(Y_r) = \mathbf{x}\mathbf{b}'_\mathbf{r} + \mathbf{a_r} + \mathbf{u_r} \end{cases} \tag{18}$$

where u_l,u_m,u_r are the real valued random variables with $E(u_l|(\mathbf{x})) = \mathbf{0}$, $E(u_m|(\mathbf{x})) = \mathbf{0}$, $E(u_r|(\mathbf{x})) = \mathbf{0}$.

The row vector of length 3p of all the components of the regressors is:

$$\mathbf{x} = (\mathbf{x_{m1}}, \mathbf{x_{l1}}, \mathbf{x_{r1}}, ..., \mathbf{x_{mP}}, \mathbf{x_{lP}}, \mathbf{x_{rP}})$$

The row vectors of length 3p of the parameters related to $\mathbf{x}$ are:

$$\mathbf{b_m} = (\mathbf{b_{mm1}}, \mathbf{b_{ml1}}, \mathbf{b_{mr1}}, ..., \mathbf{b_{mmP}}, \mathbf{b_{mlP}}, \mathbf{b_{mrP}})$$

$$\mathbf{b_l} = (\mathbf{b_{lm1}}, \mathbf{b_{ll1}}, \mathbf{b_{lr1}}, ..., \mathbf{b_{lmP)}}, \mathbf{b_{llP}}, \mathbf{b_{lrP}})$$

$$\mathbf{b_r} = (\mathbf{b_{rm1}}, \mathbf{b_{rl1}}, \mathbf{b_{rr1}}, ..., \mathbf{b_{rmP}}, \mathbf{b_{rlP}}, \mathbf{b_{rrP}})$$

The generic element b_{ijt} is the regression coefficient between the component $i\epsilon[m, l, r]$ of $\widetilde{Y}$, where m,l,r refer to center and the transformed spread of $\widetilde{Y}$, and the component $j\epsilon[m, l, r]$ of the regressor $\widetilde{x}_t$ with t=1,....,P, where m,l,r refer to the corresponding center, left spread and right spread. For example b_{mr2} is the relationship between the right spread of $\widetilde{x}_2$ and the center of Y. Of course, a_m, a_l, and a_r are the intercepts.

The covariance matrix of $\mathbf{x}$ is denoted by:

$$\Sigma_{(\mathbf{x})} = E[(\mathbf{x} - \mathbf{E}(\mathbf{x}))'(\mathbf{x} - \mathbf{E}(\mathbf{x}))] \tag{19}$$

The covariance matrix of u_m, u_l, u_r is indicated with Σ and contains the variances $\sigma^2_{u_m}$, $\sigma^2_{u_l}$ and $\sigma^2_{u_r}$.

The regression parameters can be expressed as:

$$\begin{aligned}
&\mathbf{b_m}' = [\mathbf{\Sigma}_{(\mathbf{x})}]^{-\mathbf{1}}\mathbf{E}[(\mathbf{x} - \mathbf{E}(\mathbf{x}))'(\mathbf{Y_m} - \mathbf{E}(\mathbf{Y_m})]\\
&\mathbf{b_l}' = [\mathbf{\Sigma}_{(\mathbf{x})}]^{-\mathbf{1}}\mathbf{E}[(\mathbf{x} - \mathbf{E}(\mathbf{x}))'(\mathbf{g}(\mathbf{Y_l}) - \mathbf{E}(\mathbf{g}(\mathbf{Y_l}))]\\
&\mathbf{b_r}' = [\mathbf{\Sigma}_{(\mathbf{x})}]^{-\mathbf{1}}\mathbf{E}[(\mathbf{x} - \mathbf{E}(\mathbf{x}))'(\mathbf{h}(\mathbf{Y_r}) - \mathbf{E}(\mathbf{h}(\mathbf{Y_r}))]\\
&a_m = E(Y_m|\mathbf{x}) - [\mathbf{\Sigma}_{(\mathbf{x})}]^{-\mathbf{1}}\mathbf{E}[(\mathbf{x} - \mathbf{E}(\mathbf{x}))'(\mathbf{Y_m} - \mathbf{E}(\mathbf{Y_m})]\\
&a_l = E(g(Y_l)|\mathbf{x}) - [\mathbf{\Sigma}_{(\mathbf{x})}]^{-\mathbf{1}}\mathbf{E}[(\mathbf{x} - \mathbf{E}(\mathbf{x}))'(\mathbf{g}(\mathbf{Y_l}) - \mathbf{E}(\mathbf{g}(\mathbf{Y_l}))]\\
&a_r = E(h(Y_r)|\mathbf{x}) - [\mathbf{\Sigma}_{(\mathbf{x})}]^{-\mathbf{1}}\mathbf{E}[(\mathbf{x} - \mathbf{E}(\mathbf{x}))'(\mathbf{h}(\mathbf{Y_r}) - \mathbf{E}(\mathbf{h}(\mathbf{Y_r}))]
\end{aligned} \tag{20}$$

Since the total variation of the response can be written in terms of variances and covariances of real random variables, it can be decomposed in the variation not depending on the model and that explained by the model. Thus, we can obtain a determination coefficient for the fuzzy model based on the decomposition of the total variance given by:

$$\begin{aligned}
&E[D^2(Y_t, E(Y_t)] = E[D^2(Y_t, E(Y_t|\mathbf{x})]+\\
&+ E[D^2(E(Y_t|\mathbf{x}, \mathbf{E}(\mathbf{Y_t}))]
\end{aligned} \tag{21}$$

Therefore, the linear determination coefficient R^2 can be defined as:

$$\begin{aligned}
R^2 &= \frac{E[D^2(E((Y_t|\mathbf{x}), \mathbf{E}(\mathbf{Y_t}))]}{E[D^2(Y_t, E(Y_t)]} =\\
&= 1 - \frac{E[D^2(Y_t, E(Y_t|\mathbf{x})]}{E[D^2(Y_t, E(Y_t)]}
\end{aligned} \tag{22}$$

The meaning of this index is the same of the classical regression model. The estimation problem of the regression parameters is faced by means of the LS criterion. As shown in [6], applying the appropriate substitutions and using the concept of distance between two fuzzy numbers, like in the Diamond's approach, it is possible to find the equation of the estimators of all parameters.

4 Conclusions

Fuzzy regression models are able to overcome some limitations of classical regression because they do not need the same strong assumptions. In this paper, we have presented a review of the main methods introduced in the literature on this topic and some recent developments regarding the concept of randomness in fuzzy regression. In practical applications relating to business and management sciences, fuzzy regression models with fuzzy random variables are more suitable for the characteristics of the data. However, some of the main issues of Zadeh's operations with these models are the following: the addition and the multiplication between fuzzy numbers lead to a considerable increase of the spreads; the multiplication of two symmetric fuzzy numbers does not provide a symmetric fuzzy number or at least a fuzzy number with equal spreads; spreads of Zadeh's product depend heavily on the modes of the numbers; some important algebraic properties, such as the distributive property, are valid only in particular circumstances; the product of two triangular fuzzy numbers does not provide a triangular fuzzy number. Therefore, alternative operations in order to overcome some problems connected to the addition and the product between fuzzy numbers in fuzzy linear regression models are strongly necessary. Moreover, our research prospects include considering finite geometric spaces [16, 2], multivalued functions [4] and algebraic hyperoperations [10] in fuzzy regression models.

References

[1] Achen, C., 1982. Interpreting and using regression. Sage Publications, California.

[2] Ameri, R., Nozari, T., 2012. Fuzzy hyperalgebras and direct product, Ratio Mathematica. 24.

[3] Celmins, A., 1987. Least squares model fitting to fuzzy vector data. Fuzzy Sets and Systems 22(3), 245–269.

[4] Corsini, P., Mahjoob, R., 2010. Multivalued functions, Fuzzy subsets and Join spaces. Ratio Mathematica 20, 1–41.

[5] Diamond, P., 1988. Fuzzy least squares. Information Sciences 46 (3), 141–157.

[6] Ferrraro, M., Colubi, A., Gonzales-Rodriguez, A., Coppi, R., 2010. A linear regression model for imprecise response. Int J Approx Reason 21, 759–770.

[7] Ferrraro, M., Colubi, A., Gonzales-Rodriguez, A., Coppi, R., 2011. A determination coefficient for a linear regression model with imprecise response. Environmetrics 22, 516–529.

[8] Gonzales-Rodriguez, A., Blanco, A., Colubi, A., Lubiano, M., 2009. Estimation of a simple linear regression model for fuzzy random variables. Fuzzy Sets Syst 160, 357–370.

[9] Gujarati, D., 2003. Basic Econometrics. McGraw-Hill, New York.

[10] Hošková-Mayerová, Š., Maturo, A., 2016. Fuzzy sets and algebraic hyperoperations to model interpersonal relations, in: Recent Trends in Social Systems: Quantitative Theories and Quantitative Models. Springer Nature, pp. 211–221. URL: `http://dx.doi.org/10.1007/978-3-319-40585-8_19`, doi:10.1007/978-3-319-40585-8_19.

[11] Klir, G., 2006. Uncertainty and information: foundations of generalized information theory. Wiley, New York.

[12] Kratschmer, V., 2006a. Strong consistency of least-squares estimation in linear regression models with vague concepts. J Multivar Anal 97, 633–654.

[13] Kratschmer, V., 2006b. Strong consistency of least-squares estimation in linear regression models with vague concepts. J Multivar Anal 97, 1044–1069.

[14] Lawson, C., R.J. Hanson, 1995. Solving least squares problems. Classics in applied mathematics 15.

[15] Maturo, A., Maturo, F., 2013. Research in social sciences: Fuzzy regression and causal complexity, in: Multicriteria and Multiagent Decision Making with Applications to Economics and Social Sciences. Springer, pp. 237–249. URL: `http://dx.doi.org/10.1007/978-3-642-35635-3_18`, doi:10.1007/978-3-642-35635-3_18.

[16] Maturo, A., Maturo, F., 2014. Finite geometric spaces, steiner systems and cooperative games. Analele Universitatii "Ovidius" Constanta - Seria Matematica 22. URL: `http://dx.doi.org/10.2478/auom-2014-0015`, doi:10.2478/auom-2014-0015.

[17] Maturo, A., Maturo, F., 2016. Fuzzy events, fuzzy probability and applications in economic and social sciences, in: Recent Trends in Social Systems: Quantitative Theories and Quantitative Models. Springer Nature, pp. 223–233. URL: `http://dx.doi.org/10.1007/978-3-319-40585-8_20`, doi:10.1007/978-3-319-40585-8_20.

[18] Maturo, F., 2016. FUZZINESS: TEORIE E APPLICAZIONI. Aracne Editrice, Roma, Italy. chapter LA REGRESSIONE FUZZY. pp. 99–110.

[19] Maturo, F., Fortuna, F., 2016. Bell-shaped fuzzy numbers associated with the normal curve, in: Topics on Methodological and Applied Statistical Inference. Springer. URL: `http://dx.doi.org/10.1007/978-3-319-44093-4_13`, doi:10.1007/978-3-319-44093-4_13.

[20] Maturo, F., Hošková-Mayerová, Š., 2016. Fuzzy regression models and alternative operations for economic and social sciences, in: Recent Trends in Social Systems: Quantitative Theories and Quantitative Models. Springer Nature, pp. 235–247. URL: `http://dx.doi.org/10.1007/978-3-319-40585-8_21`, doi:10.1007/978-3-319-40585-8_21.

[21] M.S. Yang, C.K., 1996. On a class of fuzzy c-numbers clustering procedures for fuzzy data. Fuzzy Sets and Syst 84, 49–60.

[22] Nather, W., 2006. Regression with fuzzy random data. Comput Stat Data Anal 51, 235–252.

[23] Puri, M., Ralescu, D., 1986. Fuzzy random variables. J Math Anal Appl 114, 409–422.

[24] Ramos-Guajardo, A., Colubi, A., Gonzales-Rodriguez, A., 2010. One-sample tests for a generalized frÃ¨chet variance of a fuzzy random variable. Metrika 71, 185–202.

[25] Shapiro, A., 2004. Fuzzy regression and the term structure of interest rates revisited. Proceedings of the 14th International AFIR Colloquium Vol.1, 29–45.

[26] Shapiro, A., 2005. Fuzzy regression models. ARC .

[27] Stock, J., Watson, M., 2009. Introduction to econometrics. Pearson Addison Wesley.