Fuzzy least squares and fuzzy orthogonal least squares linear regressions

Julien Rosset[1][a], Laurent Donzé[1][b]

[1]*ASAM Group, University of Fribourg, Boulevard de Pérolles 90, 1700 Fribourg, Switzerland*
{*Julien.Rosset, Laurent.Donzé*}*@unifr.ch*

Keywords: Fuzzy least squares linear regression; Fuzzy orthogonal least squares linear regression; Fuzzy orthogonality; Fuzzy logic; Fuzzy methods in data analysis.

Abstract: We examine the well known fuzzy least squares linear regression method. We discuss the constrained and unconstrained solutions. Based on the concept of fuzzy orthogonality, we propose the fuzzy orthogonal least squares method to solve fuzzy linear regression problems. We show that, in case of (fuzzy) orthogonal regressors, an important property of the least squares method remains valid. We obtain the same estimates of the parameters of the model if we regress on all regressors, or on each regressor considered separately. An empirical application illustrates our methods.

1 INTRODUCTION

Fuzzy regression is no longer a new topic in fuzzy analysis. Indeed, for decades, all kinds of propositions have been made to perform fuzzy regression analyses. Focus has been put particularly on solving least squares problems or least absolute deviations. We intend to review some results and complete them with the so-called orthogonal least squares method. We especially investigate the fuzzy least squares linear regression and the orthogonal fuzzy least squares linear regression. We provide essentially some strategies to deal with fuzzy data in a regression context. We proposed constrained and unconstrained solutions and discussed them.

In some situations, the analyst could profit from the orthogonality property of the independent variables. Indeed, by using a proper definition of fuzzy orthogonal variables, we show an important feature of the fuzzy orthogonal least squares method to solve a linear regression problem. As in the classical case, with crisp data, in a situation of orthogonal regressors, the estimates of the model parameters are the same if the estimation is done with all regressors of the model, or if we regress the dependent variable on each regressor alone. We verify empirically that, with our fuzzy orthogonal least squares regression, this property holds in a fuzzy context.

The main contribution of this paper is the proposition of a fuzzy orthogonal linear least squares regression method preserving the crisp orthogonal linear least squares property in a fully fuzzy environment. A nice feature of the method is to respect the fuzzy arithmetic. Our work is organized as follow. We begin our study with a small literature review in section 2. Notation is fixed in the next section, 3. Section 4 is devoted to the problem of the fuzzy least squares regression, proposing two methods to deal with crisp or fuzzy input and fuzzy output. A discussion of the constrained and unconstrained solutions of the two methods is given in section 5. After briefly defining fuzzy orthogonality, we present the method of the fuzzy orthogonal least squares dealing with crisp and fuzzy input and fuzzy output in section 6. Section 7 gives us the possibility to illustrate our methods by empirical applications. Finally, section 8 concludes our study.

2 LITERATURE REVIEW

In 1982, Tanaka introduced a possibilistic approach to fuzzy regression analysis. The method consists in using possibilistic restrictions to minimize the fuzziness of the model's fuzzy parameters. In this work, the quadratic programming approach, which allows both the minimisation of the estimated deviations of the central tendency and the minimisation of the estimated deviations in the spreads of fuzzy observations, will be of concern. (Tanaka and Lee, 1997) studied the fuzzy linear regression model by means of quadratic programming to minimise the distances between the estimated output centres and the observed

[a] https://orcid.org/0000-0002-7883-1512
[b] https://orcid.org/0000-0003-3522-4672

outputs while minimising the spreads of the estimated outputs. (Tanaka and Lee, 1998) proposed an interval regression analysis based on a quadratic programming approach to deal with the problem of fuzzy coefficients becoming crisp when using linear programming in possibilistic regression analysis. (Lee and Tanaka, 1998) proposed a fuzzy regression analysis based on a quadratic programming approach to again integrate the central tendency of least squares and the possibilistic properties of fuzzy regression. (Lee and Tanaka, 1999) also explored a fuzzy linear regression model with non-symmetric fuzzy coefficients using quadratic programming and created a lower and upper approximation model. (Donoso et al., 2006) proposed two new fuzzy regression models, the quadratic possibilistic model and the quadratic non-possibilistic model, which do not focus on the minimisation of the uncertainty of the estimated results but on the minimisation of the quadratic deviations between the observations and estimated outputs. However, (Donoso et al., 2006)'s regression models were only dealing with fuzzy regressors.

To palliate this, (D'Urso and Massari, 2013) proposed a general linear regression model for studying the dependence of fuzzy response variable, on a set of crisp or fuzzy explanatory variables. They also suggested a robust fuzzy regression method, based on the Least Median Squares estimation approach in an attempt of neutralising the effects of crisp and fuzzy outliers. In this direction, (Kashani et al., 2021) proposed a penalized estimation method to estimate the coefficients of a linear regression model with a fuzzy response variable and crisp explanatory variables. (Li et al., 2023) constructed a fuzzy multiple linear least squares regression model based on two distance measures between LR-type fuzzy numbers. (Stanojevi and Stanojevi, 2022) described a fuzzy quadratic least squares regression for a fuzzy response variable and a single crisp explanatory variable which gives regression coefficients with positive spreads.

Fuzzy inner product spaces and fuzzy orthogonality have been discussed by (Ithoh, 2017) and (Mostofian et al., 2017). They proposed new definitions of a fuzzy inner product space. They also defined a suitable notion of fuzzy orthogonality in the fuzzy world and investigated some properties. (Giordani and Kiers, 2004) proposed two extensions of the classical principal component analysis dealing with fuzzy symmetric numbers. However, The lack of a properly defined fuzzy orthogonality made the derived results losing in significance. (Yabuuch and Watada, 2017) performed a principal component analysis over crisp data belonging to fuzzy groups. In their work (Yabuuch and Watada, 2017) introduced new definitions of expectation, variance and covariance to work with the concept of fuzzy groups. However, their results are limited to crisp input data.

We wanted to investigate more thoroughly the fuzzy linear least square regression methods involving both fuzzy response and explanatory variables since these have not received much attention except from (D'Urso and Massari, 2013). However, due to the complexity of their recursive approach, constructing a fuzzy orthogonal linear least squares method inspired from it seemed to be a difficult task. Because we wanted to preserve certain properties, our approach was to build a fuzzy linear least squares regression method that could be, in accordance with the concept of fuzzy orthogonality, extended to a fuzzy orthogonal linear least squares method allowing the individual computation of the estimates.

3 NOTATION

Let us define by $\tilde{x}$ a fuzzy number. We write by $\mu_{\tilde{x}}(\cdot)$, the membership function. We consider also the α-cuts of $\tilde{x}$ denoted by $\tilde{x}^{\alpha}$ or by its equivalent in interval form by $[x^{L,\alpha}, x^{R,\alpha}]$. In practice, triangular fuzzy numbers are often used. We denote them by a triplet $\tilde{x} = (x^L, x, x^R)$ with $x^L \leq x \leq x^R \in \mathbb{R}$. Indexed fuzzy triangular number will be denoted by $\tilde{x}_k = (x_k^L, x_k, x_k^R)$. If not specified otherwise, lowercase bold letters with no index will be used for n-components column vectors of fuzzy numbers, e.g. $\tilde{\boldsymbol{y}} = (\tilde{y}_1, \ldots, \tilde{y}_n)'$. A $(n \times m)$-matrix is noted in capital bold letters, e.g. $\tilde{\boldsymbol{X}}$, with the i-th line and j-th column given respectively by $\tilde{\boldsymbol{x}}_i = (\tilde{x}_{i1}, \ldots, \tilde{x}_{im})'$, $i = 1, \ldots, n$, and $\tilde{\boldsymbol{x}}^j = (\tilde{x}_{1j}, \ldots, \tilde{x}_{nj})'$, $j = 1, \ldots, m$.

The fuzzy multiplication operator between two fuzzy numbers is denoted by $\widetilde{\otimes}$. The fuzzy multiplication between two fuzzy triangular numbers $\tilde{a}$ and $\tilde{b}$ is defined as

$$\begin{aligned}\tilde{a} \widetilde{\otimes} \tilde{b} = &(\min(a^L b^L, a^L b^R, b^R a^L, a^R b^R),\\ &ab, \max(a^L b^L, a^L b^R, b^R a^L, a^R b^R)).\end{aligned} \quad (1)$$

See (Viertl, 2018) for more details.

4 FUZZY LEAST SQUARES REGRESSION

In the following, several types of fuzzy regression models will be presented. We decided to estimate the parameters of the models by the least squares technique. We will first deal with two different cases. In the first one, we will analyse a regression model with crisp independent variables and a fuzzy dependent variable. The second case will involve both fuzzy independent and dependent variables. Then we will discuss their constrained and unconstrained solutions. As we are only interested in the estimation of the parameters and not to build a statistical model, we won't incorporate random error terms in the specification of the model. The fuzziness of the variables will be modelled by triangular fuzzy numbers, and the estimated parameters too.

4.1 Case 1: fuzzy dependent variable and crisp independent variables

This case is solved with (Donoso et al., 2006, p.1305-1306) approach. Let us consider the following regression model:

$$\tilde{y}_i = \sum_{j=1}^{m} \tilde{\beta}_j x_{ij}, \quad i = 1, \ldots, n, \tag{2}$$

where $\tilde{\beta}_j$ is the j-th parameter of the model. Note that $\tilde{\beta}_j$ is assumed to be fuzzy. Our aim is to find fuzzy triangular estimators $\hat{\tilde{\beta}}_j$ that fit (2) best by minimising the following weighted sum of squares J given in (3):

$$\begin{aligned} J = & \sum_{i=1}^{n} \left(k_1 (y_i - \sum_{j=1}^{m} \beta_j x_{ij})^2 + k_2 (y_i^L - \sum_{j=1}^{m} \beta_j^L x_{ij})^2 \right. \\ & \left. + k_3 (y_i^R - \sum_{j=1}^{m} \beta_j^R x_{ij})^2 \right) \\ = & k_1 \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i' \boldsymbol{\beta})^2 + k_2 \sum_{i=1}^{n} (y_i^L - \boldsymbol{x}_i' \boldsymbol{\beta}^L)^2 \\ & + k_3 \sum_{i=1}^{n} (y_i^R - \boldsymbol{x}_i \boldsymbol{\beta}^R)^2, \end{aligned} \tag{3}$$

under the constraints

$$(\beta_j - \beta_j^L), (\beta_j^R - \beta_j) \geq 0 \text{ for } j = 1, \ldots, m. \tag{4}$$

The quantities k_1, k_2 and k_3 are tuning weights associated with the three sums of squares corresponding to the central, left and right values of the fuzzy numbers. The constraints given in (4) make sure the estimators $\hat{\tilde{\beta}}_j$ are fuzzy triangular numbers, i.e. $\hat{\tilde{\beta}}_j = (\hat{\beta}_j^L, \hat{\beta}, \hat{\beta}_j^R)$.

4.2 Case 2: fuzzy dependent variable and fuzzy independent variables

Let us suppose now that the independent variables are also fuzzy, i.e. $\tilde{x}_{ij} = (x_{ij}^L, x_{ij}, x_{ij}^R)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. The fuzzy linear regression model becomes:

$$\tilde{y}_i = \sum_{j=1}^{m} \tilde{\beta}_j \widetilde{\otimes} \tilde{x}_{ij}, \quad i = 1, \ldots, n, \tag{5}$$

where $\widetilde{\otimes}$ denotes the fuzzy multiplication operator. We find the fuzzy triangular estimators $\hat{\tilde{\beta}}$ by minimising the sum of squares $\sum_{i=1}^{n} \left(\tilde{y}_i - \sum_{j=1}^{m} \tilde{\beta}_j \widetilde{\otimes} \tilde{x}_{ij} \right)^2$. Rewriting it using the fuzzy multiplication rule yields the following objective function J given in (6):

$$\begin{aligned} J = \sum_{i=1}^{n} & \left(k_1 (y_i - \sum_{j=1}^{m} \beta_j x_{ij})^2 \right. \\ & + k_2 (y_i^L - \sum_{j=1}^{m} (\min(\beta_j^L x_{ij}^L, \beta_j^R x_{ij}^L, \beta_j^L x_{ij}^R, \beta_j^R x_{ij}^R))^2 \\ & \left. + k_3 (y_i^R - \sum_{j=1}^{m} (\max(\beta_j^L x_{ij}^L, \beta_j^R x_{ij}^L, \beta_j^L x_{ij}^R, \beta_j^R x_{ij}^R))^2 \right). \end{aligned} \tag{6}$$

Minimising (6) is not so direct, and a proper strategy needs to be developed. If we know the sign of the independent triangular fuzzy variable $\tilde{x}_{ij}$ and the sign of the unknown fuzzy parameters $\tilde{\beta}_j$ for all $j = 1, ..., m$, we can easily know what will be the minimum and maximum values to be calculated in the objective function (6) and thus solve the regression problem under the constraints (4). In practice, the sign of the observations is evidently known, but not the sign of the parameters. However, on one hand, by using the crisp value of the fuzzy triangular observations, i.e. when $\alpha = 1$, we can perform the classical least squares to get the estimates $\hat{\beta}_j$, $j = 1, \ldots, m$, and if they are sufficiently far away from zero, the sign of $\hat{\tilde{\beta}}_j$ will be the one of $\hat{\beta}_j$. If, on the other hand, $\hat{\beta}_j$ is close to zero, one has to first estimate $\tilde{\beta}_j$, $j = 1, \ldots, m$, via (3). We note this first estimates by $\hat{\tilde{\beta}}_j^{(1)}$. We then pick their signs to fix the ones of the unknown estimators β_j^L and β_j^R and use them to solve problem (6).

Suppose we estimated the sign of the unknown fuzzy parameters as written above and then chose, according to the signs of the fuzzy observation $\tilde{x}_{ij}$ and $\tilde{\beta}_j$, $i = 1, ..., n$, $j = 0, ..., m$ the corresponding minimum $x_{ij}^{\min}$ and maximum $x_{ij}^{\max}$ that solve $\min(\beta_j^L x_{ij}^L, \beta_j^R x_{ij}^L, \beta_j^L x_{ij}^R, \beta_j^R x_{ij}^R)$ and $\max(\beta_j^L x_{ij}^L, \beta_j^R x_{ij}^L, \beta_j^L x_{ij}^R, \beta_j^R x_{ij}^R)$ respectively. The objective function (6) can then be rewritten as:

$$J = \sum_{i=1}^{n}\left(k_1(y_i - \sum_{j=1}^{m}\beta_j x_{ij})^2 + k_2(y_i^L - \sum_{j=1}^{m}(\beta_j^{\min} x_{ij}^{\min})^2 + k_3(y_i^R - \sum_{j=1}^{m}(\beta_j^{\max} x_{ij}^{\max})^2\right), \tag{7}$$

where $\beta_j^{\min}$, $\beta_j^{\max}$ are the coefficients associated to $x_{ij}^{\min}$ and $x_{ij}^{\max}$ respectively. The special case where $x_{ij} > 0$, $i = 1, \ldots, n$, deserves our attention. In this case, $\beta_j^{\min}$ and $\beta_j^{\max}$ are obviously equal to β_j^L and β_j^R respectively.
Notice that, in general, when the fuzzy observations $\tilde{x}_{ij}$ are not necessarily all positive, the coefficients $\beta_j^{\min}$ or $\beta_j^{\max}$ may appear simultaneously in both left and right squared terms in (6). This is due to the fuzzy arithmetic rule for the product that sometimes yields $\beta_j^{\min}$ or $\beta_j^{\max}$ as the coefficients, which both maximise and minimise the fuzzy product. This general case is no longer a simple quadratic problem. This is why one recommends adding a constant valued vector to each covariate to ensure their positivity and thus allowing to solve a proprer quadratic problem. Making so, only the constant of the problem is affected letting the slopes unchanged.

5 CONSTRAINED AND UNCONSTRAINED SOLUTIONS

The constraint (4) assures that the solutions are fuzzy triangular numbers. If we do not impose (4), we end up with solutions $\tilde{\beta}_j$ which are sometimes of the form (b, c, a) with $b > c > a$, $a, b, c \in \mathbb{R}$. This clearly goes against the definition of fuzzy triangular numbers. A way to bypass this not plausible solution is to consider the fuzzy parameters as fuzzy intervals.

In order to avoid confusion, it is very important to make a clear distinction between a vector of fuzzy numbers and a fuzzy vector. (Viertl, 2018, p.14) defines a fuzzy vector as:

Definition 1 (Fuzzy vector and fuzzy interval)**.**
A k-dimensional fuzzy vector $\tilde{x}^*$ with membership function $\mu_{\tilde{x}^*}$ is such that:

1. $\mu_{\tilde{x}^*} : \mathbb{R}^k \to [0,1]$.
2. *The support of $\mu_{\tilde{x}^*}$ is a bounded set.*
3. $\forall \alpha \in (0,1]$ *the so-called* α*-cut* $C_\alpha(\tilde{x}^*) := \{x \in \mathbb{R}^k | \mu_{\tilde{x}^*}(x) \geq \alpha\}$ *is non-empty, bounded, and a finite union of simply connected and closed bounded sets.*

The set of all k-dimensional fuzzy vectors is denoted by $\mathcal{F}(\mathbb{R}^k)$. A k-dimensional fuzzy vector is called a k-dimensional fuzzy interval if all α-cuts are connected compact sets.

Applying the concept of fuzzy intervals, we are able to understand the main difference between the constrained solution and the unconstrained one. In the first case, when the constraints (4) are satisfied, one ends up with a vector of fuzzy parameters $\tilde{\beta}$ satisfying $\beta_k^L \leq \beta_k \leq \beta_k^R$ meaning that, the upper regression coefficients β_k^R must always give a steeper slope to the regression line, and conversely, β_k^L must always give a lower slope. When the constraints (4) are dropped, one ends up with a fuzzy interval $\tilde{\beta}_j^*$, $j = 1, \ldots, m$. In this case, the fitted solution is such that $\hat{\boldsymbol{y}}^{\text{lower}} \leq \hat{\boldsymbol{y}}^{\text{upper}}$. This allows the slopes to violate (4) so long the lower regression line is below the upper one on the interval formed by the observations $\boldsymbol{X}$ or $\tilde{\boldsymbol{X}}$.

As an example, let us consider a regression model with two crisp independent variables, the first one being the constant of the model. The unconstrained solution $\hat{\tilde{\boldsymbol{\beta}}}^*$ can be written as:

$$\hat{\tilde{\boldsymbol{\beta}}}^* = \begin{pmatrix} \hat{\tilde{\beta}}_0^* \\ \hat{\tilde{\beta}}_1^* \end{pmatrix} = \begin{pmatrix} \hat{\beta}_0^{*,\text{lower}}, \hat{\beta}_0^*, \hat{\beta}_0^{*,\text{upper}} \\ \hat{\beta}_1^{*,\text{lower}}, \hat{\beta}_1^*, \hat{\beta}_1^{*,\text{upper}} \end{pmatrix}, \tag{8}$$

where $\hat{\beta}_0^{*,\text{lower}}$ is the smaller β_0 coefficient estimated and $\hat{\beta}_0^{*,\text{upper}}$ is the greater β_0 coefficient estimated. The coefficients $\hat{\beta}_1^{*,\text{lower}}$ and $\hat{\beta}_1^{*,\text{upper}}$ are defined analogously. It can be easily shown that the estimated parameters (8) satisfy the definition of a fuzzy interval. Indeed, let C_α be the rectangle formed by the vertices $(\beta_0^{*,\text{lower},\alpha}, \beta_1^{*,\text{lower},\alpha})$, $(\beta_0^{*,\text{lower},\alpha}, \beta_1^{*,\text{upper},\alpha})$, $(\beta_0^{*,\text{upper},\alpha}, \beta_1^{*,\text{lower},\alpha})$, $(\beta_0^{*,\text{upper},\alpha}, \beta_1^{*,\text{upper},\alpha})$ where α stands for the α-cut of the fuzzy number. The more we reduce the fuzziness, that is the closer to 1 α is, the closer $\hat{\tilde{\boldsymbol{\beta}}}^{*,\alpha}$ will be to the crisp solution given when $\alpha = 1$. Thus, $C_{\alpha_i} \subseteq C_{\alpha_j}$ for $\alpha_i \geq \alpha_j$. Assuming the existence of a solution, C_α is non-empty, bounded and a finite union of simply connected and bounded rectangles. Furthermore, The vector-membership function $\mu_{\hat{\tilde{\boldsymbol{\beta}}}^*} : \mathbb{R}^2 \to [0,1]$ defined as :

$$\mu_{\hat{\tilde{\boldsymbol{\beta}}}^*} := \max\{\alpha \cdot I_{C_\alpha}(x) : \alpha \in (0,1]\}, \quad \forall x \in \mathbb{R}^2,$$

has its support bounded by the rectangle C_0.

By following the same reasoning, one can easily verify that in the case of m covariates, the solution is also a fuzzy interval, where C_α is the m-dimensional rectangle which has the form of a hypercube of 2^m vertices.

6 FUZZY ORTHOGONAL LEAST SQUARES

Let us first remind the definition of the fuzzy inner product and the concept of fuzzy orthogonality, described by (Ithoh, 2017, p.13).

Definition 2 (Fuzzy inner product)**.**

*Let X be a nonzero vector space over the field $\mathbb{R}$ and $\tilde{X} = \{x_\lambda | x \in X, \lambda \in (0,1]\}$ be the set of all fuzzy points in X. A function $\langle \cdot, \cdot \rangle : \tilde{X} \times \tilde{X} \to \mathbb{R}$ is said to be a **fuzzy inner product** on $\tilde{X}$ if*

1. *a) $\langle x_\lambda, x_\mu \rangle \geq 0$; b) $\langle x_\lambda, x_\mu \rangle = 0$ if and only if $x = 0$;*
2. *$\langle kx_\lambda, y_\mu \rangle = k\langle x_\lambda, y_\mu \rangle$;*
3. *$\langle x_\lambda + y_\lambda, z_\mu \rangle = \langle x_\lambda, z_\mu \rangle + \langle y_\lambda, z_\mu \rangle$;*
4. *$\langle x_\lambda, y_\mu \rangle = \langle y_\mu, x_\lambda \rangle$;*
5. *$\langle x_\lambda, y_\mu \rangle \leq \langle x_\lambda, y_\nu \rangle$ if $0 < \nu \leq \mu \leq 1$;*
6. *for every x_λ, $y_\mu \in \tilde{X}$ and $\varepsilon > 0$, there exists $0 < \delta < \mu$ such that $\langle x_\lambda, y_{\mu-\delta} \rangle < \langle x_\lambda, y_\mu \rangle + \varepsilon$;*
7. *for every x_λ, $y_\mu \in \tilde{X}$ and $\varepsilon > 0$, there exists $0 < \delta < 1 - \mu$ such that $\langle x_\lambda, y_{\mu+\delta} \rangle > \langle x_\lambda, y_\mu \rangle - \varepsilon$.*

*The pair $(X, \langle \cdot, \cdot \rangle)$ is called a **strong fuzzy inner product space**.*

Remark 1.

1. *Let $(X, \langle \cdot, \cdot \rangle)$ be a strong fuzzy inner product space. If*

$$\langle x_{\lambda_1}, y_{\mu_1} \rangle = \langle x_{\lambda_2}, y_{\mu_2} \rangle \quad x, y \in X, \quad \lambda_i, \mu_i \in (0,1], \quad i = 1,2, \tag{9}$$

then $(X, \langle \cdot, \cdot \rangle)$ becomes a usual inner product space;

2. *If we use a stronger condition,*

$$\langle x_\lambda + y_\mu, z_\nu \rangle = \langle x_\lambda, y_\mu \rangle + \langle y_\mu, z_\nu \rangle, \tag{10}$$

then $(X, \langle \cdot, \cdot \rangle)$ is just a usual inner product space.

The concept of fuzzy orthogonality is defined as:

Definition 3 (Fuzzy orthogonality)**.**

*Let $\tilde{x}$ and $\tilde{y}$ be vectors in a fuzzy inner product space. One says that $\tilde{x}$ is **fuzzy orthogonal** to $\tilde{y}$ if $\langle x_\lambda, y_\mu \rangle = 0$, for some $\lambda, \mu \in (0,1]$.*

The fuzzy orthogonality between two vectors of a fuzzy inner product space $\tilde{x}$ and $\tilde{y}$ is denoted by $\tilde{x} \perp_F \tilde{y}$. In case x_λ is orthogonal to y_μ for these specific values of $\lambda, \mu \in (0,1]$, i.e $\langle x_\lambda, y_\mu \rangle = 0$, we note it $x_\lambda \perp y_\mu$.

We can now introduce the fuzzy orthogonal least squares estimators of a fuzzy regression model. Our aim is first to transform the independent variables in such a way that at the end of the procedure, the transformed variables are mutually (fuzzy) orthogonal. We then can use these transformed variables as new regressors in the regression equation and estimate the model as described in section 4. The orthogonalisation procedure will be discussed for two cases: independent crisp or fuzzy variables.

6.1 Case 3: fuzzy dependent variable and crisp independent variables

We can use a classic orthonormalisation procedure to project $\boldsymbol{X}$ onto an orthonormal basis. Let $\bar{\boldsymbol{x}} = (\bar{x}_1, ..., \bar{x}_m)'$ be the vector of the means associated to the m independent variables $X_1, ..., X_m$; $\boldsymbol{S}$ the estimated variance-covariance matrix; $\boldsymbol{G}$ the matrix of all the eigenvectors of $\boldsymbol{S}$; $\boldsymbol{\iota}_n = (1, \ldots, 1)'$ a $(n \times 1)$ vector. We project $\boldsymbol{X}$ onto an orthonormal basis using the transformation

$$\boldsymbol{X}^\perp = (\boldsymbol{X} - \boldsymbol{\iota}_n \bar{\boldsymbol{x}}')\boldsymbol{G}. \tag{11}$$

Then we solve (3) with $\boldsymbol{X}^\perp$ instead of $\boldsymbol{X}$. Due to the orthogonality of $\boldsymbol{X}^\perp$, the individual coefficients $\tilde{\beta}_j$ of the regression model can be estimated also by regressing the dependent variable on the variable X_j only, $j = 1, \ldots, m$. As a consequence, the orthogonalisation of the regressors allows us to find uncorrelated fuzzy estimators $\tilde{\beta}_j$.

Note that, sometimes, by convention, $-G$ is used instead of G in (11). One has to be careful with multiplying or not by a minus sign the orthogonal projection (11) since it can greatly affect the fuzzy triangular estimators $\tilde{\beta}_j$. This effect will be discussed more in depth in section 7.3.

6.2 Case 4: fuzzy dependent variable and fuzzy independent variables

When both the dependent and independent variables of the regression model are fuzzy, we first need to solve the min() and max() functions in (6). We perform then the orthogonalisation procedure (11) to find $\boldsymbol{X}^{\perp\min}$, $\boldsymbol{X}^\perp$ and $\boldsymbol{X}^{\perp\max}$ from $\boldsymbol{X}^{\min}$, $\boldsymbol{X}$ and $\boldsymbol{X}^{\max}$. This allows us to write the objective function (7) with orthogonal fuzzy data as

$$\begin{aligned} J = & k_1 \sum_{i=1}^{n} (y_i - \sum_{j=1}^{m} \beta_j x_{ij}^{\perp})^2 + k_2 \sum_{i=1}^{n} (y_i^L - \sum_{j=1}^{m} \beta_j^L x_{ij}^{\perp\min})^2 \\ & + k_3 \sum_{i=1}^{n} (y_i^R - \sum_{j=1}^{m} \beta_j^R x_{ij}^{\perp\max})^2. \end{aligned} \tag{12}$$

Again, the properties of the orthogonality permit us to compute the fuzzy least squares individually for each covariate X_j while giving the same solution as when computed with all covariates X_j at once.

Observe that the restructuring of the initial data into $\tilde{x}_{ij} = (x_{ij}^{\min}, x_{ij}, x_{ij}^{\max})$ allowing us to compute the fuzzy products in (6) is possible if we know the signs of the fuzzy unknown parameters $\tilde{\boldsymbol{\beta}}$. This ambiguity can be solved by adequately permuting the observations to take into account the fuzzy arithmetic. After the orthogonalisation procedure, in case of negative sign occuring, the signs of the observations $\tilde{x}_{ij}^{\perp} = (x_{ij}^{\perp\min}, x_{ij}^{\perp}, x_{ij}^{\perp\max})$ may differ from the signs of $\tilde{x}_{ij} = (x_{ij}^{\min}, x_{ij}, x_{ij}^{\max})$.

To correct this and to preserve both the fuzzy arithmetic and the quadratic nature of the problem, one should shift $\tilde{\boldsymbol{x}}^{j\perp}$, $j = 1, \ldots, m$, by adding a constant $K_j = |\min(\tilde{\boldsymbol{x}}^{j\perp})|$. This operation ensures that the signs of the observations $\tilde{\boldsymbol{x}}^{j\perp}$ are positive. As we already said, this transformation has the effect of changing the intercept value however, it does not change the value of the slope coefficients. In practical applications, one may use this transformation to ensure that the signs of the observations $\tilde{\boldsymbol{X}}$ are positive as well as ensure that the orthonormalised observations $\tilde{\boldsymbol{X}}^{\perp}$ are positive too.
Lastly, note that the fuzzy orthogonal covariates $\tilde{\boldsymbol{x}}^{j\perp}$ meet the definition of fuzzy orthogonality of (Ithoh, 2017). Indeed, by construction their left part, center and right part are orthogonal at any given α-cut,

$$\begin{aligned}
\langle \tilde{\boldsymbol{x}}^{j\perp L}, \tilde{\boldsymbol{x}}^{k\perp L} \rangle &= 0, \\
\langle \boldsymbol{x}^{j\perp}, \boldsymbol{x}^{k\perp} \rangle &= 0, \\
\langle \tilde{\boldsymbol{x}}^{j\perp R}, \tilde{\boldsymbol{x}}^{k\perp R} \rangle &= 0, \quad \forall j,k = 1, \ldots, m, j \neq k,
\end{aligned} \tag{13}$$

$\tilde{\boldsymbol{x}}^{j\perp}$ satisfying Masuo Itoh's definition of fuzzy orthogonaly.

7 APPLICATION

Let us apply the above fuzzy least squares regression methods and discuss them. We considered the fuzzy data set given in Table 1 inspired by (Tanaka and Lee, 1998).

7.1 Fuzzy dependent variable and crisp independent variables

Let us postulate the following fuzzy regression model (14):

$$\tilde{Y}_i = \tilde{\beta}_0 X_0 + \tilde{\beta}_1 X_i + \tilde{\beta}_2 X_i^2, \quad i = 1, \ldots, 22. \tag{14}$$

Using $k_1 = k_2 = k_3 = 1$ in (3), we find, following the methodology described in section 4.1 the fuzzy triangular estimates given in Table 2. In Table 2, The studied model is displayed in the first column while the second column features the fuzzy regression coefficients obtained with the use of a given model under constraints 4. These fuzzy estimates are fuzzy triangular numbers since they are solutions of constrained models. The third column of Table 2 depicts the fuzzy estimates given by solving the models in the first column under no constraint, thus these fuzzy estimates are of the form of a fuzzy interval, as explained in definition 1.

Table 3 gives the sum of squared residuals. The first column tells which model has been used and if the fuzzy estimates have been obtained under constraints or not. Then, the four remaining columns give the sum of squared residuals with respect to the left, center and right values and their addition computed using the objective function of the respective models.

7.2 Fuzzy dependent variable and fuzzy independent variables

We consider again the data in Table 1. We fuzzify the covariates X and X^2 by triangular fuzzy numbers. Let be $\tilde{X} = (X^L, X, X^R)$, $\tilde{X}^2 = \tilde{X} \tilde{\otimes} \tilde{X} = (X^{L^2}, X^2, X^{R^2})$ with $X^L = X - u_1$, $X^R = X + u_2$ where $u_1 \sim \mathrm{U}(0, 0.5)$ and $u_2 \sim \mathrm{U}(0, 0.7)$. Note that $\tilde{X}^2$ is easily computed since the observations are positive. The problem is now given by the fuzzy regression equation (15):

$$\tilde{Y}_i = \tilde{\beta}_0 X_0 + \tilde{\beta}_1 \tilde{\otimes} \tilde{X}_i + \tilde{\beta}_2 \tilde{\otimes} \tilde{X}_i^2 \quad i = 1, \ldots, 22. \tag{15}$$

Using the methodology of section 4.2 the constrained and unconstrained solutions are given in Table 2 and the sum of squared residuals in Table 3.

7.3 Fuzzy dependent variable and crisp orthogonal independent variables

By means of an SVD decomposition and using (11), we orthonormalise $\boldsymbol{X}$ onto $\boldsymbol{X}^{\perp}$. The orthonormalised covariates $X_1^{\perp}$ and $X_2^{\perp}$ of X and X^2 are given in Table 1. The model is expressed by the regression equation (16):

$$\tilde{Y}_i = \tilde{\beta}_0 X_{0,i} + \tilde{\beta}_1 X_{1,i}^{\perp} + \tilde{\beta}_{2,i} X_2^{\perp}, \quad i = 1, \ldots, 22. \tag{16}$$

Following the methodology discussed in section 6.1, the estimated parameters, without and with a multiplication of a minus sign, are given in Table 2. Note that for the unconstrained solution (Fuzzy intervals), the solutions are the same.

Table 1: Fuzzy Dataset

$\tilde{Y}$	X_0	X	X^2	$X_1^{\perp}$	$X_2^{\perp}$
(15,22.5,30)	1	1	1	-1.93184373	-0.35491140
(20,28.75,37.5)	1	2	4	-1.80915821	-0.25981071
(15,25,35)	1	3	9	-1.67727776	-0.17390496
(25,42.5,60)	1	4	16	-1.53620237	-0.09719415
(25,40,55)	1	5	25	-1.38593203	-0.02967828
(40,52.5,65)	1	6	36	-1.22646676	0.02864264
(55,75,95)	1	7	49	-1.05780654	0.07776863
(70,85,100)	1	8	64	-0.87995139	0.11769968
(80,105,130)	1	9	81	-0.69290129	0.14843578
(90,120,150)	1	10	100	-0.49665625	0.16997695
(115,145,175)	1	11	121	-0.29121627	0.18232317
(140,167.5,195)	1	12	144	-0.07658135	0.18547446
(155,187.5,220)	1	13	169	0.14724851	0.17943080
(175,212.5,250)	1	14	196	0.38027331	0.16419220
(200,240,280)	1	15	225	0.62249305	0.13975866
(240,275,310)	1	16	256	0.87390773	0.10613018
(270,305,340)	1	17	289	1.13451735	0.06330676
(300,342.5,385)	1	18	324	1.40432191	0.01128840
(340,380,420)	1	19	361	1.68332142	-0.04992490
(380,420,460)	1	20	400	1.97151586	-0.12033314
(420,460,500)	1	21	441	2.26890525	-0.19993632
(465,507.5,550)	1	22	484	2.57548957	-0.28873445

As the regressors of model (16) are mutually orthogonal, we can find the estimates of the parameter $\tilde{\beta}_1$ and $\tilde{\beta}_2$ by estimating successively the models:

$$\tilde{Y}_i = \tilde{\beta}_0 X_{0,i} + \tilde{\beta}_1 X_{1,i}^{\perp}, \quad i = 1,\dots,22, \tag{17}$$

and

$$\tilde{Y}_i = \tilde{\beta}_0 X_{0,i} + \tilde{\beta}_{2,i} X_2^{\perp}, \quad i = 1,\dots,22. \tag{18}$$

Finally, remark that one can come back to the estimates computed with observations $\boldsymbol{X}$ using a linear transformation. Let $\boldsymbol{G}$ be the matrix of the eigenvectors of the variance-covariance matrix of the regressors,

$$\begin{aligned} \boldsymbol{G}\boldsymbol{\beta}^{L,\perp} &= \boldsymbol{\beta}^{L}, \\ \boldsymbol{G}\boldsymbol{\beta}^{\perp} &= \boldsymbol{\beta}, \\ \boldsymbol{G}\boldsymbol{\beta}^{R,\perp} &= \boldsymbol{\beta}^{R}. \end{aligned} \tag{19}$$

This transformation only works for the unconstrained solutions. In the constrained case, this is in general not true.

At this point, an important observation has to be made. When using orthogonal observations $\boldsymbol{X}^{\perp}$ we find a different fuzziness for the fuzzy estimators $\hat{\tilde{\beta}}_2$ and $\hat{\tilde{\beta}}_2^{\perp}$:

$$\hat{\tilde{\beta}}_2 = (1.010993, 1.010993, 1.010993)$$

is crisp while

$$\hat{\tilde{\beta}}_2^{\perp} = (-130.0262, -111.6804, -93.33466)$$

is fuzzy. This phenomenon is due to the constraints (4). They force the fuzzy "slope" represented by $\tilde{\beta}_2$ to be steeper for the upper β_2^R and lower for β_2^L. Depending on the reference frame used, this constrains more or less the solutions. Because of this, one has to be careful with the convention of multiplying or not a minus sign after having projected the observations onto an orthonormal basis. The solution $\hat{\tilde{\beta}}^{\perp}$ of model (16) would have become

$$\hat{\tilde{\beta}}^{\perp} = \begin{bmatrix} (165.2273, 192.6705, 220.11364) \\ (-108.2216, -108.2216, -108.2216) \\ (111.6804, 111.6804, 111.6804) \end{bmatrix}, \tag{20}$$

if we choose to multiply by a minus sign, and thus $\hat{\tilde{\beta}}_2^{\perp}$ would also be crisp. Notice, on the other hand, that the unconstrained solution is unaffected by the choice of sign.

7.4 Fuzzy dependent variable and fuzzy orthogonal independent variables

The model considered is given by the regression equation (21):

$$\tilde{Y}_i = \tilde{\beta}_0 X_{0,i} + \tilde{\beta}_1 \tilde{X}^{\perp}_{1,i} + \tilde{\beta}_{2,i} \tilde{X}^{\perp}_{2,i}, \quad i = 1, \ldots, 22. \quad (21)$$

The orthogonal covariates $X_1^{\perp}$ and $X_2^{\perp}$ are obtained by orthogonalisation of $\tilde{X}_1$ and $\tilde{X}_2$ as explained in section 6.2. Then, by application of the method discussed in section 6.2 the resulting estimated parameters are given in Table 2 and the associated sum of squared residuals in Table 3. Note that the estimated parameters can be computed individually for each covariate $\tilde{X}_j$ and yielding the same results. Moreover, using the transformation (19) with $\boldsymbol{X}^{\min}$ and $\boldsymbol{X}^{\max}$ instead of $\boldsymbol{X}$ for the left, respectively right fuzzy estimates $\beta^{L,\perp}$, $\beta^{R,\perp}$ will return the original estimates computed with $\boldsymbol{X}^{\min}$, $\boldsymbol{X}$ and $\boldsymbol{X}^{\max}$.

7.5 General discussion

Notice that, in table 2, in case 1, 2 and 3 (with minus sign), the fuzzy interval estimates seem to better preserve the fuzziness than the fuzzy triangular ones. This show that, empirically constrained models tend to have crispier estimates than unconstrained ones. Moreover, as depicted in case 3, we can see how the freedom of multiplying or not by a minus sign the orthogonal projection (11) affects the fuzziness of the estimates.

7.6 Sums of squared residuals

Lastly, in Table 3, notice that models (14) and its orthonormalised version (16) share the same sum of squared residuals. Moroever, this is also true for models 15 and 21. This could in fact be resulting from the transformation 19 which allows one to retrieve the estimates found with non orthogonal regressors with the orthonormalised ones.

8 CONCLUSION

We reexamine the fuzzy least squares method to solve the so-called fuzzy linear regression problems. We deal with two cases in particular. First, we considered that the independent variables are crisp, and second, we treat the case of fuzzy independent variables. In both situation, the dependent variable is fuzzy. We develop a proper strategy to efficiently deal with the fuzziness appearing in the observations. Moreover, we present and discuss two different types of solutions arising from constrained and unconstrained fuzzy least squares regression problems which are respectively fuzzy triangular valued and fuzzy interval valued.

Then, the extension of the method to orthogonal fuzzy least squares regression methods has been investigated. In case of (fuzzy) orthogonal independent variables, an important property of the classical least squares method has been preserved. Due to the orthogonality of the regressors, the individual coefficients $\tilde{\beta}_j$ of the regression model can be estimated also by regressing the dependent variable on the $j-$th covariate only. As a consequence, the orthogonalisation of the regressors allows us to find uncorrelated fuzzy estimators $\tilde{\beta}_j$. Moreover, in the unconstrained case, we showed that there exist a linear transformation to find the coefficients associated to the original regression model, i.e. the model before the orthogonalisation of the regressors. We also highlighted that the resulting sum of squares residuals of a model and its orthonormal counterpart are the same. These discoveries seem to be very promising to make progress in statistical inference with fuzzy data, in particular in the study of the fuzzy distributions of the estimates.

REFERENCES

Donoso, S., Marín, N., and Villa, M. A. (2006). Quadratic programming models for fuzzy regression. *Intelligent Data Engineering and Automated Learning – IDEAL 2006, 7th International Conference, Burgos, Spain, September 20-23, 2006. Proceedings. Lecture Notes in Computer Science*, pages 1304–1311.

D'Urso, P. and Massari, R. (2013). Weighted Least Squares and Least Median Squares estimation for the fuzzy linear regression analysis. *Metron*, 71(3):279–306.

Giordani, P. and Kiers, H. A. (2004). Principal Component Analysis of symmetric fuzzy data. *Computational Statistics & Data Analysis*, 45(3):519–548.

Ithoh, M. (2017). Fuzzy inner product spaces and fuzzy orthogonality. *Advances in Fuzzy Sets and Systems*, 5(1-10).

Kashani, M., Arashi, M., Rabiei, M., D'Urso, P., and Giovanni, L. D. (2021). A fuzzy penalized regression model with variable selection. *Expert Systems with Applications*, 175:114696.

Lee, H. and Tanaka, H. (1998). Fuzzy regression analysis by quadratic programming reflecting central tendency. *Metrika*, 25(1):65–80.

Lee, H. and Tanaka, H. (1999). Fuzzy approximations with non-symmetric fuzzy parameters in fuzzy regression analysis. *Journal of the Operations Research Society of Japan*, 42(1):98–112.

Li, Y., He, X., and Liu, X. (2023). Fuzzy multiple linear least squares regression analysis. *Fuzzy Sets and Systems*, 459:118–143.

Mostofian, E., Azhini, M., and Bodaghi, A. (2017). Fuzzy inner product spaces and fuzzy orthogonality. *Tbilisi Mathematical Journal*, 10(2):157–171.

Stanojevi, B. and Stanojevi, M. (2022). Quadratic least square regression in fuzzy environment. *Elsevier B.V.*

Tanaka, H. and Lee, H. (1997). Fuzzy linear regression combining central tendency and possibilistic properties. *Proceedings of 6th International Fuzzy Systems Conference*, 1:63–68.

Tanaka, H. and Lee, H. (1998). Interval regression analysis by quadratic programming approach. *IEEE Transactions on Fuzzy Systems*, 6(1):473–481.

Viertl, R. (2018). *Statistical Methods for Fuzzy Data*. Wiley.

Yabuuch, Y. and Watada, J. (2017). Fuzzy Principal Component Analysis and Its Application. *International Journal of Biomedical Soft Computing and Human Sciences: the official journal of the Biomedical Fuzzy Systems Association*, 3(1):83–92.

Table 2: Models estimations [1]

Models	Fuzzy triangular estimates [2]	Fuzzy intervals [3]
Case 1 (14)	$(15.673, 22.605, 29.537)$ $(-2.1602, -0.3766, 1.4069)$ $(1.0109, 1.0109, 1.0109)$	$(20.032, 22.605, 25.178)$ $(-3.2499, -0.3766, 2.4966)$ $(1.0583, 1.0109, 0.9636)$
Case 2 (15)	$(14.473, 21.298, 30.283)$ $(-1.3693, -0.0499, 0.7347)$ $(0.9967, 0.9967, 0.9967)$	$(21.861, 22.605, 23.275)$ $(-3.4034, -0.3766, 2.4123)$ $(1.0849, 1.0109, 0.9259)$
Case 3 (16) (without (-) [4])	$(165.22, 192.67, 220.11)$ $(100.18, 108.22, 116.26)$ $(-130.02, -111.68, -93.334)$	$(165.22, 192.67, 220.11)$ $(100.18, 108.22, 116.26)$ $(-130.02, -111.68, -93.334)$
Case 3 (16) (with (-) [5])	$(165.22, 192.67, 220.11)$ $(-108.22, -108.22, -108.22)$ $(111.68, 111.68, 111.68)$	$(165.22, 192.67, 220.11)$ $(100.18, 108.22, 116.26)$ $(-130.02, -111.68, -93.334)$
Case 4 (21) (without (-))	$(165.28, 192.67, 220.11)$ $(100.24, 108.22, 116.26)$ $(-124.05, -111.68, -91.63)$	$(165.28, 192.67, 220.11)$ $(100.24, 108.22, 116.26)$ $(-124.05, -111.68, -91.62)$

[1]Table displaying the fuzzy regression coefficients obtained via a given fuzzy regression model shown in the first column.
[2]Solutions of a given constrained fuzzy least squares regression model.
[3]Solutions of a given unconstrained fuzzy least squares regression model.
[4]The orthogonal projection has not been multiplied by a minus sign.
[5]The orthogonal projection has been multiplied by a minus sign.

Table 3: Sum of squared residuals[1]

Models	Constraints	Left value [2]	Center value[3]	Right value [4]	Total value[5]
(14)	constrained	489.8294	256.9274	555.5718	1302.329
(14)	unconstrained	426.2246	256.9274	491.9669	1175.119
(15)	constrained	555.5714	303.3096	1189.367	2048.248
(15)	unconstrained	599.9359	256.9274	630.1572	1487.021
(16) (with (-) [6])	constrained	489.8294	256.9274	555.5718	1302.329
(16)(w/o (-) [7])	constrained	3306.798	256.9274	3372.54	6936.265
(16)	unconstrained	426.2246	256.9274	491.9669	1175.119
(21) (w/o (-))	constrained & unconstrained	599.9359	256.9274	630.1572	1487.021

[1]Table displaying the squared residuals of the different models studied throughout this work.
[2]Squared residuals of the left fuzzy parts.
[3]Squared residuals of the central fuzzy parts.
[4]Squared residuals of the right fuzzy parts.
[5]Sum of the left, center and right squared residuals.
[6]The orthogonal projection has been multiplied by a minus sign.
[7]The orthogonal projection has not been multiplied by a minus sign.