Conditional expectation and fuzzy regression

Andreas Wünsche
Faculty of Mathematics and Computer Science, Agricolastraße 1,
Freiberg University of Mining and Technology
09596 Freiberg, Germany
wuensche@math.tu-freiberg.de

Abstract

We show that analogously to classical probability theory the conditional expectation $\mathbf{E}(\widetilde{Y}|\widetilde{X})$ of a fuzzy random variable $\widetilde{Y}$ w.r.t. a fuzzy random variable $\widetilde{X}$ is w.r.t. a suitable metric the best approximation of $\widetilde{Y}$ by measurable functions of $\widetilde{X}$. Furthermore, several linear regression functions, i.e. best approximation of $\widetilde{Y}$ by linear functions of $\widetilde{X}$ and examples for random LR-fuzzy numbers and Gaussian fuzzy random variables are presented.

Keywords: fuzzy random variables, conditional expectation, fuzzy regression

1.Preliminaries. Let $(\Omega, \mathcal{A}, \mathbf{P})$be a probability space and $X_0, Y_0 : \Omega \longrightarrow \mathbb{R}^n$ random variables with $\mathbf{E}(X_0^2)$, $\mathbf{E}(Y_0^2) < \infty$, i.e. $X_0, Y_0 \in \mathrm{L}^2$,where L^2 denotes the set of all measurable squared integrable functions. There is the well known result:

$$\mathbf{E}(Y_0 - \mathbf{E}(Y_0|X_0))^2 \leq \inf_{f(X_0) \in \mathrm{L}^2} \mathbf{E}(Y_0 - f(X_0))^2 \tag{1}$$

Hence , w.r.t. quadratic loss the conditional expectation $\mathbf{E}(Y_0|X_0)$ is the best approximation of Y_0 by measurable functions of X_0. For example let X_0 and Y_0 be Gaussian as follow:

$$\begin{pmatrix} Y_0 \\ X_0 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}; \begin{pmatrix} \sigma_y^2 & \rho\sigma_y\sigma_x \\ \rho\sigma_x\sigma_y & \sigma_x^2 \end{pmatrix}\right)$$

$$\Rightarrow \mathbf{E}(Y_0|X_0) = \mu_y + \rho\frac{\sigma_y}{\sigma_x}(X_0 - \mu_x)$$
$$= \frac{Cov(X_0,Y_0)}{Var(X_0)}X_0 + \mathbf{E}(Y_0) - \frac{Cov(X_0,Y_0)}{Var(X_0)}\mathbf{E}(X_0) \tag{2}$$

If we restrict ourselves to find the best approximation of Y_0 by linear functions of X_0, i.e. if we consider $\min_{a,b\in\mathbb{R}} \mathbf{E}(Y_0 - (aX_0 + b))^2$, the solution is the so-called linear regression of Y_0 w.r.t. X_0 which coincides with (2), i.e.

$$a^* = \frac{Cov(X_0,Y_0)}{Var(X_0)} \text{ and } b^* = \mathbf{E}Y_0 - a^*\mathbf{E}X_0. \tag{3}$$

The aim of this paper is to show similar results for fuzzy random variables.

Let $\mathcal{P}_c(\mathbb{R}^n)$ denote the family of all nonempty compact, convex subsets of $\mathbb{R}^n$. A fuzzy set on $\mathbb{R}^n$ is identified by its membership function $v : \mathbb{R}^n \longrightarrow [0,1]$. Let be $v_\alpha := \{t \in \mathbb{R}^n | v(t) \geq \alpha\}$ the α-cut of v, $\alpha \in (0,1]$. Let be $\mathcal{F}_c(\mathbb{R}^n)$ the family of convex fuzzy sets with the following properties :

(i) v is normal, i.e. there exists a $t \in \mathbb{R}^n$ such that $v(t) = 1$;
(ii) v is upper semicontinuous ;
(iii) $v(\alpha t_1 + (1-\alpha)t_2) \geq min(v(t_1), v(t_2))$, $t_1, t_2 \in \mathbb{R}^n$, $\alpha \in [0,1]$;
(iv) $v_o = cl\{t \in \mathbb{R}^n | v(t) > 0\} \in \mathcal{P}_c(\mathbb{R}^n)$.

Let $X, Y \in \mathcal{P}_c(\mathbb{R}^n)$. As distance between X and Y we use

$$d_2(X,Y) := \sqrt{\int_{S^{n-1}} (s_X(t) - s_Y(t))^2 \mu(dt)},$$

where μ is the Lebesgue measure normalized on S^{n-1}, the $(n-1)$-dimensional unit sphere on $\mathbb{R}^n$. Notice that this metric and the Hausdorff metric are topological equivalent ([2]). From this metric we can define a metric in $\mathcal{F}_c(\mathbb{R}^n)$ by:

$$\delta_2(\widetilde{Y},\widetilde{X}) = \sqrt{n\int_0^1 \int_{S^{n1}} (s_{\widetilde{Y}_\alpha}(t) - s_{\widetilde{X}_\alpha}(t))^2 \mu(dt) d\alpha}.$$

Now, a compact convex random set (crs) X can be seen as a Borel measurable function

$$X: \quad \Omega \to (\mathcal{P}_c(\mathbb{R}^n), d_2),$$

and a fuzzy random variable (frv) $\widetilde{X}$ can be defined as a Borel measurable function

$$\widetilde{X}: \quad \Omega \to (\mathcal{F}_c(\mathbb{R}^n), \delta_2).$$

Then all α−cuts of $\widetilde{X}$ are crs .

Following [1] the expected value $\mathbf{E}X$ of a random set X is defined by:
$\mathbf{E}X =$

$$\{\mathbf{E}x \,|\, x : \Omega \to \mathbb{R}^n, \mathbf{E}||x|| < \infty, x(\omega) \in X(\omega)\ P-a.e.\}.$$

The expected value of a frv $\widetilde{X}$ was introduced by Puri and Ralescu ([7]) as the unique fuzzy set $\mathbf{E}\widetilde{X}$ with:

$$(\mathbf{E}\widetilde{X})_\alpha = \mathbf{E}\widetilde{X}_\alpha, \quad 0 \le \alpha \le 1.$$

Following [4] the variance of a frv $\widetilde{X}$ is defined as $Var\widetilde{X} = \mathbf{E}\delta_2^2(\widetilde{X},\mathbf{E}\widetilde{X})$. Using $(\mathbf{E}\widetilde{X})_\alpha = \mathbf{E}\widetilde{X}_\alpha$ and $s_{\mathbf{E}\widetilde{X}_\alpha} = \mathbf{E}s_{\widetilde{X}_\alpha}$ this can be written as

$$Var\widetilde{X} = n\int_0^1 \int_{S^{n-1}} Var s_{\widetilde{X}_\alpha}(t)\mu(dt)d\alpha.$$

Analogously, the covariance between two frv's $\widetilde{X}$ and $\widetilde{Y}$ is defined as

$$Cov(\widetilde{X},\widetilde{Y}) = n\int_0^1 \int_{S^{n-1}} Cov(s_{\widetilde{X}_\alpha}(t), s_{\widetilde{Y}_\alpha}(t))\mu(dt)d\alpha.$$

For more details see [5].

Consider now a sub-σ-algebra $\mathcal{B} \subseteq \mathcal{A}$. The conditional expectation of $\widetilde{Y}$ with respect to $\mathcal{B}$ is the frv $\mathbf{E}(\widetilde{Y}|\mathcal{B})$ which satisfies the following properties:

(a) $\mathbf{E}(\widetilde{Y}|\mathcal{B})$ is $\mathcal{B}$ measurable.

(b) $\int_B \mathbf{E}(\widetilde{Y}|\mathcal{B})dP = \int_B \widetilde{Y}dP \quad \forall B \in \mathcal{B}.$

Then it holds (see [8] Proposition 4.1) $(\mathbf{E}(\widetilde{X}|\mathcal{B}))_\alpha = \mathbf{E}(\widetilde{X}_\alpha|\mathcal{B})$. Clearly, if $\mathcal{B}$ is induced by a frv $\widetilde{X}$ we write $\mathbf{E}(\widetilde{X}|\mathcal{B}) = \mathbf{E}(\widetilde{Y}|\widetilde{X})$.

Let us introduce the Hukuhara-difference between fuzzy sets: Let be given two fuzzy sets $\widetilde{A}$ and $\widetilde{B}$. If there exists a fuzzy set $\widetilde{C}$ with $\widetilde{A} = \widetilde{B} \oplus \widetilde{C}$ then $\widetilde{C}$ is called Hukuhara-difference between $\widetilde{A}$ and $\widetilde{B}$, denoted by

$$\widetilde{C} := \widetilde{A} \ominus_H \widetilde{B}.$$

Note that $\oplus$ describes the (extended) sum between fuzzy sets $\widetilde{B}$ and $\widetilde{C}$.

2. Properties of the conditional expectation. For all the following let be

$$\begin{aligned} X_0 &: \Omega \to \mathbb{R} \quad \text{a random variable,} \\ X,Y &: \Omega \to \mathcal{P}_c(\mathbb{R}^n) \quad \text{a random sets and} \\ \widetilde{X},\widetilde{Y} &: \Omega \to \mathcal{F}_c(\mathbb{R}^n) \quad \text{fuzzy random variables.} \end{aligned}$$

Analogously to the known result $\mathbf{E}(s_X) = s_{\mathbf{E}X}$ it holds:

Theorem 1 *For any sub-σ-algebra $\mathcal{F} \subseteq \mathcal{A}$:*

$$\mathbf{E}(s_X(t)|\mathcal{F}) = s_{\mathbf{E}(X|\mathcal{F})}(t) \quad \text{for all } t \in S^{n-1}.$$

A direct consequence of the classical result (1) is:

Theorem 2 *For all $t \in S^{n-1}$ and measurable $f : \mathcal{F}_c(\mathbb{R}^n) \to \mathcal{P}_c(\mathbb{R}^n)$ it holds:*

$$\mathbf{E}(s_Y(t) - \mathbf{E}(s_Y(t)|\widetilde{X}))^2 \le \mathbf{E}(s_Y(t) - s_{f(\widetilde{X})}(t))^2.$$

Now Theorem1 and Theorem 2 leads to:

Theorem 3 *For all $f : \mathcal{F}_c(\mathbb{R}^n) \to \mathcal{F}_c(\mathbb{R}^n)$, f_α measurable for all α it holds:*

$$\mathbf{E}(\delta_2^2(\widetilde{Y},\mathbf{E}(\widetilde{Y}|\widetilde{X}))) \le \mathbf{E}(\delta_2^2(\widetilde{Y},f(\widetilde{X}))).$$

Example 1 Let $\widetilde{Y}$ be a LR-f.r.v. and X_0 a random variable, i.e.

$$\widetilde{Y}(\omega) = \langle \mu_Y, l_Y, r_Y \rangle_{LR}$$

The conditional expectation in this case is:

$\mathbf{E}(\widetilde{Y}|X_0)(\omega) =$

$$\langle \mathbf{E}(\mu_Y|X_0), \mathbf{E}(l_Y|X_0), \mathbf{E}(r_Y|X_0) \rangle_{LR}$$

Example 2 Let $\widetilde{X}$ and $\widetilde{Y}$ be Gaussian fuzzy random variables (see [6] and [3]), which can be characterized as:

$$\widetilde{Y} = \mathbf{E}(\widetilde{Y}) + \{\underline{\xi}_1\} \quad \text{and} \quad \widetilde{X} = \mathbf{E}(\widetilde{X}) + \{\underline{\xi}_2\}$$

where $\mathbf{E}(\widetilde{Y})$ and $\mathbf{E}(\widetilde{X})$ are the Aumann expectations of $\widetilde{Y}$ and $\widetilde{X}$ and

$$\begin{pmatrix} \underline{\xi}_1 \\ \underline{\xi}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \underline{0} \\ \underline{0} \end{pmatrix}; \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right).$$

The conditional expectation results in

$$\mathbf{E}(\widetilde{Y}|\widetilde{X}) = \mathbf{E}(\widetilde{Y}) + \{\Sigma_{12}\Sigma_{22}^{-1}\underline{\xi}_2\}$$

Hence, $\mathbf{E}(\widetilde{Y}|\widetilde{X})$ is also a Gaussian fuzzy random variable and has a structure similar to (2).

3. Fuzzy Regression.

In this section we are interested in best approximation of a given frv $\widetilde{Y}$ by a linear function, i.e. we look for results similar to (3). There are at least two possible generalizations of $\min_{a,b} \mathbf{E}(Y_0 - (aX_0 + b))^2$. Firstly, we will approximate $\widetilde{Y}$ by a linear (regression) function of a further frv $\widetilde{X}$, i.e.

$$\inf_{a \in \mathbb{R}, \widetilde{B} \in \mathcal{F}_c(\mathbb{R}^n)} \mathbf{E}\delta_2^2(\widetilde{Y}, a\widetilde{X} \oplus \widetilde{B}) \qquad (4)$$

and secondly we will approximate $\widetilde{Y}$ by a linear (regression) function of a real random variable X_0 but with fuzzy coefficients, i.e.

$$\inf_{\widetilde{A}, \widetilde{B} \in \mathcal{F}_c(\mathbb{R}^n)} \mathbf{E}\delta_2^2(\widetilde{Y}, \widetilde{A}X_0 \oplus \widetilde{B}). \qquad (5)$$

Clearly, for writing down (4), (5) it is necessary that $\widetilde{Y}, \widetilde{X}, X_0$ are quadratic integrable random variables. In some cases solution of (4), (5) can be expressed by the Hukuhara difference $\ominus_H$ introduced in section 1.

Theorem 4 *Let $\widetilde{X}, \widetilde{Y}$ be two quadratic integrable frv's and consider the optimization problem (4).*
(a) Let be $a \geq 0$: If $Cov(\widetilde{X}, \widetilde{Y}) \geq 0$ and $\mathbf{E}\widetilde{Y} \ominus_H \frac{Cov(\widetilde{X}\widetilde{Y})}{Var\widetilde{X}} \mathbf{E}\widetilde{X}$ exists then

$$a^* = \frac{Cov(\widetilde{X}, \widetilde{Y})}{Var\widetilde{X}} \quad , \quad \widetilde{B}^* = \mathbf{E}\widetilde{Y} \ominus_H a^* \mathbf{E}\widetilde{X}$$

are solutions of (4).
(b) Let be $a \leq 0$: If $Cov(-\widetilde{X}, \widetilde{Y}) \geq 0$ and $\mathbf{E}\widetilde{Y} \ominus_H -\frac{Cov(-\widetilde{X}\widetilde{Y})}{Var\widetilde{X}} \mathbf{E}\widetilde{X}$ exists then

$$a^* = -\frac{Cov(-\widetilde{X}, \widetilde{Y})}{Var\widetilde{X}} \quad , \quad \widetilde{B}^* = \mathbf{E}\widetilde{Y} \ominus_H a^* \mathbf{E}\widetilde{X}$$

are solutions of (4).

Note that the results in Theorem 4 formally coincide with (3).

Let us apply Theorem 4 to random LR-fuzzy numbers. For two LR-fuzzy numbers $A = \langle \mu_A, l_A, r_A \rangle_{LR}$ and $B = \langle \mu_B, l_B, r_B \rangle_{LR}$ the Hukuhara difference $\widetilde{A} \ominus_H \widetilde{B}$ exists if $l_A \geq l_B$ and $r_A \geq r_B$ and is given by

$$\widetilde{A} \ominus_H \widetilde{B} = \langle \mu_A - \mu_B, l_A - l_B, r_A - r_B \rangle_{LR}. \qquad (6)$$

Assume now that two random LR-fuzzy numbers $\widetilde{X} = \langle \mu_X, l_X, r_X \rangle_{LR}$, $\widetilde{Y} = \langle \mu_Y, l_Y, r_Y \rangle_{LR}$ are given. Assume

$$a^* = \frac{Cov(\widetilde{X}, \widetilde{Y})}{Var\widetilde{X}} \geq 0 \quad , \quad \mathbf{E}l_Y \geq a^* \mathbf{E}l_X, \ \mathbf{E}r_Y \geq a^* \mathbf{E}r_X$$

then $\widetilde{B}^* = \mathbf{E}\widetilde{Y} \ominus_H a^* \mathbf{E}\widetilde{X}$ exists and can be written as

$$\begin{aligned} \widetilde{B}^* &= \langle \mathbf{E}\mu_Y, \mathbf{E}l_Y, \mathbf{E}r_Y \rangle_{LR} \\ &\overset{(6)}{=} \left\langle \mathbf{E}\mu_Y - \frac{Cov(\widetilde{X}, \widetilde{Y})}{Var\widetilde{X}} \mathbf{E}\mu_X, \mathbf{E}l_Y - \frac{Cov(\widetilde{X}, \widetilde{Y})}{Var\widetilde{X}} \mathbf{E}l_X, \right. \\ &\quad \left. \mathbf{E}r_Y - \frac{Cov(\widetilde{X}, \widetilde{Y})}{Var\widetilde{X}} \mathbf{E}r_X \right\rangle_{LR} \end{aligned}$$

Now, let us turn towards problem (5). Here we also present an explicit solution in that case where certain Hukuhara differences exist.

Theorem 5 *(a) If $X_0 \geq 0$ and $VarX_0 > 0$ then*

$$\begin{aligned} \widetilde{A}^* &= \frac{1}{VarX_0} \left(\mathbf{E}(\widetilde{Y}X_0) \ominus_H \mathbf{E}\widetilde{Y}\mathbf{E}X_0 \right) \text{ and} \\ \widetilde{B}^* &= \mathbf{E}\widetilde{Y} \ominus_H \widetilde{A}^* \mathbf{E}X_0 \end{aligned}$$

are solutions of (5) if the corresponding Hukuhara-differences exist.

(b) If $X_0 \leq 0$ and $VarX_0 > 0$ then

$$\begin{aligned} \widetilde{A}^* &= -\frac{1}{VarX_0} \left(-\mathbf{E}(\widetilde{Y}X_0) \ominus_H -\mathbf{E}\widetilde{Y}\mathbf{E}X_0 \right) \text{ and} \\ \widetilde{B}^* &= \mathbf{E}\widetilde{Y} \ominus_H \widetilde{A}^* \mathbf{E}X_0 \end{aligned}$$

are solutions of (5) if the corresponding Hukuhara-differences exist.

As an example, in the LR-fuzzy number case $\widetilde{Y} = <\mu_Y, l_Y, r_Y>_{LR}$ and $X_0 \geq 0$ it can be seen that the Hukuhara difference exists if $\mathbf{E}l_Y \geq \frac{Cov(X_0,l_Y)}{VarX_0}\mathbf{E}X_0$, $\mathbf{E}r_Y \geq \frac{Cov(X_0,r_Y)}{VarX_0}\mathbf{E}X_0$, $Cov(X_0,l_Y) \geq 0$ and $Cov(X_0,r_Y) \geq 0$. Then the result is

$$\begin{aligned}
\widetilde{A}^* &= \frac{1}{VarX_0}(\mathbf{E}(\widetilde{Y}X_0)) \ominus_H \mathbf{E}X_0\mathbf{E}\widetilde{Y}) \\
&= \frac{1}{VarX_0}(\langle \mathbf{E}(X_0\mu_Y), \mathbf{E}(X_0 l_Y), \mathbf{E}(X_0 r_Y)\rangle_{LR} \\
&\quad \ominus_H \langle \mathbf{E}X_0\mathbf{E}\mu_Y, \mathbf{E}X_0\mathbf{E}l_Y, \mathbf{E}X_0\mathbf{E}r_Y\rangle_{LR}) \\
&= \left\langle \frac{Cov(X_0,\mu_Y)}{VarX_0}, \frac{Cov(X_0,l_Y)}{VarX_0}, \frac{Cov(X_0,r_Y)}{VarX_0} \right\rangle_{LR}
\end{aligned}$$

$$\begin{aligned}
&\widetilde{B}^* = \mathbf{E}\widetilde{Y} \ominus_H \widetilde{A}^*\mathbf{E}X_0 \\
&= \left\langle \mathbf{E}\mu_Y - \frac{Cov(X_0,\mu_Y)}{VarX_0}\mathbf{E}X_0, \right. \\
&\left. \mathbf{E}l_Y - \frac{Cov(X_0,l_Y)}{VarX_0}\mathbf{E}X_0, \mathbf{E}r_Y - \frac{Cov(X_0,r_Y)}{VarX_0}\mathbf{E}X_0 \right\rangle_{LR} .
\end{aligned}$$

Remark 1 *Proofs of the Theorems and further information you can find in [9].*

Literatur

[1] R. J. Aumann. Integrals of set-valued functions. *J. Math. Anal. Appl.*, 12:1–12, 1965.

[2] P. Diamond and P. Kloeden. *Metric Space of fuzzy sets: Theory and Applications*. World Scientific, Singapore, New Jessey, London, Hongkong, 1994.

[3] Y. Feng. Gaussian fuzzy random variable. *Fuzzy Sets and Systems*, 111(3):325–330, 2000.

[4] R. Körner and W. Näther. Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. *Information Sciences*, 109:95–118, 1998.

[5] W. Näther. On random fuzzy variables of second order and their application to linear statistical inference with fuzzy data. *Metrika*, 51:201–221, 2000.

[6] M. Puri and D. Ralescu. The concept of normality for fuzzy random variables. *The Annals of Probability*, 13(4):1373–1379, 1985.

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

[8] M. Puri and D. Ralescu. Convergence theorem for fuzzy martingals. *J. Math. Anal. Appl.*, 160:107–122, 1991.

[9] A. Wünsche and W. Näther. Conditional expectation of fuzzy random variables and fuzzy regression. *submitted to Fuzzy Sets and Systems*, 2001.