ISSN 1974-4110 (on line edition)
ISSN 1594-7645 (print edition)

Università
degli Studi di Urbino
"Carlo Bo"

Facoltà
di Economia

www.econ.uniurb.it

WP-EMS
Working Papers Series in
Economics, Mathematics and Statistics

"FUZZIFICATION VIA F-TRANSFORM"

- Luciano Stefanini (University of Urbino, Italy)
- Maria Letizia Guerra (University of Bologna, Italy)

WP-EMS # 2013/10

Fuzzification via F-transform

Luciano Stefanini[*], Maria Letizia Guerra[†]

Abstract

In this paper we show how a fuzzification process can benefit of the F-transform and possibility distributions.

1 Introduction

In order to handle incomplete information, possibility theory is a powerful choice. It is similar to probability theory because it is based on set functions; in possibility theory, however, the additivity does not hold and there are two dual functions and not only one: the possibility and the necessity measures ([6]). The possibility distributions were primarly introduced by Zadeh ([24]) in order to provide a graded semantics to natural language statements. Hereinafter the two measures become a common way to model uncertainty.

Liu in [9] provides a survey of credibility theory that is a new branch of mathematics for studying fuzzy events. The number $Cr\{A\}$ indicates the credibility that an event A will occur and it satisfies certain mathematical properties.

The paper is organized in four sections After the introduction we suggest the correct use of the quantiles in the possibilistic distribution function.

2 Use of quantiles in the (possibilistic) distribution function

Consider a fuzzy number $u \in \mathbb{R}_{\mathcal{F}}$ with membership function (in LR-representation) $u : \mathbb{R} \longrightarrow [0,1]$, with support $[a,b]$ and core $\{c\}$, $c \in]a,b[$

$$u(x) = \begin{cases} 0 & \text{if} \quad x < a \\ u^L(x) & \text{if} \quad a \le x \le c \\ 1 & \text{if} \quad c \le x \le d \\ u^R(x) & \text{if} \quad d \le x \le b \\ 0 & \text{if} \quad x > b \end{cases};$$

[*]Department of Economics, Society and Politics, University of Urbino, Italy; luciano.stefanini@uniurb.it

[†]Department of Mathematics, University of Bologna, Italy; mletizia.guerra@unibo.it

the α-cuts (in LU-representation) are

$$[u]_\alpha = [u_\alpha^-, u_\alpha^+],\ \alpha \in [0,1].$$

The possibility function of u is defined for all $x \in \mathbb{R}$ by

$$Pos_u(x) = \sup\{u(t) \mid t \le x\};$$

and the necessity function of u is

$$Nec_u(x) = 1 - \sup\{u(t) \mid t > x\}$$

then the (possibilistic) distribution function of u is defined by

$$F_u(x) = \frac{Pos(x) + Nec(x)}{2} \tag{1}$$

i.e., if u is continuous,

$$F_u(x) = \begin{cases} 0 & \text{if} \quad x < a \\ \frac{1}{2}u^L(x) & \text{if} \quad a \le x \le c \\ \frac{1}{2} & \text{if} \quad c \le x \le d \\ 1 - \frac{1}{2}u^R(x) & \text{if} \quad d \le x \le b \\ 1 & \text{if} \quad x > b \end{cases}$$

Remark that $F_u(c) = \frac{1}{2}$.

We can prove the following result:

Proposition 1 *Let $u \in \mathbb{R}_{\mathcal{F}}$. For all $\alpha \in]0,1]$, the α-cut $[u_\alpha^-, u_\alpha^+]$ of u is given by*

$$u_\alpha^- = \inf\{x | F_u(x) \ge \frac{\alpha}{2}\} \tag{2}$$

$$u_\alpha^+ = \sup\{x | F_u(x) \le 1 - \frac{\alpha}{2}\} \tag{3}$$

Proof. We fix $\alpha \in]0,1]$ and for (2) we have that $F_u(x) \ge \frac{\alpha}{2}$ means

$$\begin{aligned} 0 &\ge \frac{\alpha}{2} \text{ if } x \le a \\ \frac{1}{2}u^L(x) &\ge \frac{\alpha}{2} \text{ if } x \in [a,c] \\ \frac{1}{2} &\ge \frac{\alpha}{2} \text{ if } x \in [c,d] \\ 1 - \frac{1}{2}u^R(x) &\ge \frac{\alpha}{2} \text{ if } x \in [d,b] \\ 1 &\ge \frac{\alpha}{2} \text{ if } x \ge b \end{aligned}$$

so that

$$\inf\left\{x | F_u(x) \ge \frac{\alpha}{2}\right\} =$$

$$= \inf \left\{ x | u^L(x) \geq \alpha \text{ if } x \in [a, c]\,, \text{ or } u^R(x) \leq 2 - \alpha \text{ if } x \in [c, b] \right\} \tag{4}$$

where we can exclude the case $x \leq a$ because the inequality $0 \geq \frac{\alpha}{2}$ is never true and we can omit $x \geq b$, too, because the inequality $1 \geq \frac{\alpha}{2}$ is always verified.
Finally we observe that $u^L(u_\alpha^-) \geq \alpha$ and that the following implication holds $x < u_\alpha^- \Rightarrow u^L(x) < \alpha$; on this way we obtain that the expression in (4) is equal to:

$$\inf \left\{ x | u^L(x) \geq \alpha \right\} = u_\alpha^-$$

due to the fact that $u^R(x) \leq 2 - \alpha$ is always true because $u^R(x) \in [0, 1]$ and $\alpha \in \,]0, 1]\,$.
In a similar way for (3) we have that $F_u(x) \leq 1 - \frac{\alpha}{2}$ means

$$\begin{aligned} 1 &\geq \frac{\alpha}{2} \text{ if } x \leq a \\ 1 - \frac{1}{2} u^L(x) &\geq \frac{\alpha}{2} \text{ if } x \in [a, c] \\ \frac{1}{2} &\geq \frac{\alpha}{2} \text{ if } x \in [c, d] \\ \frac{1}{2} u^R(x) &\geq \frac{\alpha}{2} \text{ if } x \in [d, b] \\ 0 &\geq \frac{\alpha}{2} \text{ if } x \geq b \end{aligned}$$

Now, for $\alpha \in]0, 1]$ it is impossible to have $0 \geq \frac{\alpha}{2}$ and then all values $x \geq b$ can be omitted; in addition the inequality $1 \geq \frac{\alpha}{2}$ is always true and we can remove the condition $x \leq a$. These consideration imply that

$$\sup\{x | F_u(x) \leq 1 - \frac{\alpha}{2}\} =$$

$$= \sup\{x | u^L(x) \leq 2 - \alpha \text{ if } x \in [a, c]\,, \text{ or } u^R(x) \geq \alpha \text{ if } x \in [c, b]\}$$

Observing, finally, that the inequality $u^L(x) \leq 2 - \alpha$ is always true we can conclude as follows

$$= \sup\{x | u^R(x) \geq \alpha\} = u_\alpha^+$$

because $u^R(u_\alpha^+) \geq \alpha$ and that the following implication holds $x < u_\alpha^+ \Rightarrow u^R(x) < \alpha$. ■

A consequence is that, if $F_u(x)$ is continuous and strictly increasing, u_α^- is such that $F_u(u_\alpha^-) = \frac{\alpha}{2}$ and u_α^+ is such that $F_u(u_\alpha^+) = 1 - \frac{\alpha}{2}$; furthermore, considering $\alpha = 1$, we obtain $c = \inf\{x | F_u(x) \geq \frac{1}{2}\} = \sup\{x | F_u(x) \leq \frac{1}{2}\}$ i.e. $c = \{x | F_u(x) = \frac{1}{2}\}$; the core value c (assumed to be unique) has the same property as the *median* of a probability distribution.

Consider the following function (the integral is Riemann-Stieltjes assuming $F_u(x)$ absolutely continuous)

$$S_u(m) = \int\limits_{-\infty}^{+\infty} |m - x| dF_u(x).$$

The minimization of $S_u(m)$ with respect to m gives a value of m such that $\int\limits_{-\infty}^{m} dF_u(x) \geq \frac{1}{2}$ and $\int\limits_{m}^{+\infty} dF_u(x) \geq \frac{1}{2}$; and if $F_u(x)$ is absolutely continuous this gives $F_u(m) \geq \frac{1}{2}$ and $1 - F_u(m) \geq \frac{1}{2}$, i.e. $F_u(m) = \frac{1}{2}$; it follows that $m = c$.

Proposition 2 *Let $F_u(x)$ be the (possibilistic) distribution function of $u \in \mathbb{R}_{\mathcal{F}}$ and $u(x) = 1 \iff x = c$. Then $S_u(c) \leq S_u(\xi)$ for all $\xi \in \mathbb{R}$.*

Proof. We can write

$$S_u(\xi) = \int\limits_{-\infty}^{\xi} (\xi - x) dF_u(x) + \int\limits_{\xi}^{+\infty} (x - \xi) dF_u(x).$$

Then, considering that $\xi - x = (\xi - c) + (c - x)$ or $x - \xi = (x - c) + (c - \xi)$ and using the properties $\int\limits_{-\infty}^{\xi} = \int\limits_{-\infty}^{c} + \int\limits_{c}^{\xi}$, $\int\limits_{\xi}^{+\infty} = \int\limits_{\xi}^{c} + \int\limits_{c}^{+\infty}$, we have

$$\begin{aligned} S_u(\xi) \quad = \quad & \int\limits_{-\infty}^{c} (\xi - c) dF_u(x) + \int\limits_{-\infty}^{c} (c - x) dF_u(x) + \\ & \int\limits_{c}^{\xi} (\xi - c) dF_u(x) + \int\limits_{c}^{\xi} (c - x) dF_u(x) + \\ & \int\limits_{\xi}^{c} (x - c) dF_u(x) + \int\limits_{\xi}^{c} (c - \xi) dF_u(x) + \\ & \int\limits_{c}^{+\infty} (x - c) dF_u(x) + \int\limits_{c}^{+\infty} (c - \xi) dF_u(x) \end{aligned}$$

But

$$\int\limits_{-\infty}^{c} (\xi - c) dF_u(x) = \frac{1}{2}(\xi - c),$$

$$\int\limits_{c}^{+\infty} (c - \xi) dF_u(x) = \frac{1}{2}(c - \xi),$$

$$\int\limits_{-\infty}^{c} (c - x) dF_u(x) + \int\limits_{c}^{+\infty} (x - c) dF_u(x) = \int\limits_{c}^{+\infty} |x - c| dF_u(x)$$

$$2 \int\limits_{c}^{\xi} (\xi - c) dF_u(x) = \int\limits_{c}^{\xi} (\xi - c) dF_u(x) + \int\limits_{c}^{\xi} (c - x) dF_u(x) + \int\limits_{\xi}^{c} (x - c) dF_u(x) + \int\limits_{\xi}^{c} (c - \xi) dF_u(x)$$

and it follows that

$$S_u(\xi) = S_u(c) + 2 \int\limits_{c}^{\xi} (\xi - x) dF_u(x).$$

Now, from the basic properties of the Riemann- Stijelties integrals, we have $\int\limits_{c}^{\xi} (\xi - x) dF_u(x) \geq 0$; as a consequence, $S_u(\xi) \geq S_u(c)$ and the proof is complete.
■

Definition 3 (possibilistic) *r-quantile*

Analogously, we can prove the following property.

Proposition 4 *If u is a fuzzy number with possibilistic distribution $F_u(x)$, $x \in \mathbb{R}$, then for all $\alpha \in]0,1]$, the α-cuts $[u_\alpha^-, u_\alpha^+]$ of u are such that u_α^- is the $\frac{\alpha}{2}$-quantile and u_α^+ is the $(1-\frac{\alpha}{2})$-quantile of $F_u(x)$, i.e. they are obtained, respectively, by minimizing the functionals*

$$S_\alpha^-(m) = (1-\frac{\alpha}{2}) \int_{-\infty}^{m} (m-x)dF_u(x) + \frac{\alpha}{2} \int_{m}^{+\infty} (x-m)dF_u(x)$$

and

$$S_\alpha^+(m) = \frac{\alpha}{2} \int_{-\infty}^{m} (m-x)dF_u(x) + (1-\frac{\alpha}{2}) \int_{m}^{+\infty} (x-m)dF_u(x).$$

The proposition above is useful to estimate the level cuts (and the membership function) of a normal fuzzy number $u \in \mathbb{R}_\mathcal{F}$. This application mimics very well known results in statistics and we apply it without proof.

Suppose that a fuzzy number $u \in \mathbb{R}_\mathcal{F}$ is known at N observations $(x_i, u(x_i))$, $i = 1, ..., N$ and that the known membership values $u(x_i)$ are uniform on $[0,1]$.

For a value of $\alpha \in]0,1]$ determine the (empirical) $\frac{\alpha}{2}$-quantile $u_\alpha^-(N)$ and the $(1-\frac{\alpha}{2})$-quantile $u_\alpha^+(N)$ to be the minimizers of the functionals

$$\begin{aligned} S_\alpha^-(m) &= (1-\frac{\alpha}{2}) \sum_{\substack{i=1 \\ x_i<m}}^{N} (m-x_i) + \frac{\alpha}{2} \sum_{\substack{i=1 \\ x_i>m}}^{N} (x_i-m) \\ S_\alpha^+(m) &= \frac{\alpha}{2} \sum_{\substack{i=1 \\ x_i<m}}^{N} (m-x_i) + (1-\frac{\alpha}{2}) \sum_{\substack{i=1 \\ x_i>m}}^{N} (x_i-m). \end{aligned}$$

The obtained values

$$m_\alpha^-(N) = \arg\min_m S_\alpha^-(m) \tag{5}$$

$$m_\alpha^+(N) = \arg\min_m S_\alpha^+(m) \tag{6}$$

are an estimate $[m_\alpha^-(N), m_\alpha^+(N)]$ of the α-cut $[u_\alpha^-, u_\alpha^+]$ of u.

The minimization can be performed by linear programming.in the following way.

Inserire e spiegare

For $k = 1, 2, ..., n$ let j_k and n_k be the integers such that $t_j \in [x_{k-1}, x_{k+1}]$, $k = 1, 2, ..., n$ for $j = j_k + 1, ..., j_k + n_k$. We have $n_k \geq 1$ because the points t_j are sufficiently dense with respect to the partition $(\mathbb{P}, \mathbb{A})$. Then the two linear programming formulations can be obtained as follows: define $2n_k$ nonnegative variables y_i^-, y_i^+, $i = 1, 2, ..., n_k$ such that

$$\begin{aligned} y_i^- &= \begin{cases} 0 & \text{if } f(t_{j_k+i}) \geq F \\ F - f(t_{j_k+i}) & \text{if } f(t_{j_k+i}) < F \end{cases} \\ y_i^+ &= \begin{cases} 0 & \text{if } f(t_{j_k+i}) \leq F \\ f(t_{j_k+i}) - F & \text{if } f(t_{j_k+i}) > F \end{cases} . \end{aligned}$$

Then

$$\begin{aligned} y_i^- + y_i^+ &= |F - f(t_{j_k+i})|,\ i = 1, 2, ..., n_k, \\ y_i^- - y_i^+ &= F - f(t_{j_k+i}),\ i = 1, 2, ..., n_k \end{aligned}$$

and

$$\begin{aligned} S_\alpha^-(F) &= (1 - \frac{\alpha}{2}) \sum_{i=1}^{n_k} y_i^- A_k(t_{j_k+i}) + \frac{\alpha}{2} \sum_{i=1}^{n_k} y_i^+ A_k(t_{j_k+i}) \\ S_\alpha^+(F) &= \frac{\alpha}{2} \sum_{i=1}^{n_k} y_i^- A_k(t_{j_k+i}) + (1 - \frac{\alpha}{2}) \sum_{i=1}^{n_k} y_i^+ A_k(t_{j_k+i}). \end{aligned}$$

Defining $c_i^{(k)} = A_k(t_{j_k+i})$ and $b_i^{(k)} = f(t_{j_k+i})$, $i = 1, ..., n_k$, $k = 1, ..., n$, the minimization of $S_\alpha^-(F)$ and $S_\alpha^+(F)$ requires to minimize the two linear objective functions (with nonnegative cost coefficients)

$$S_\alpha^-(F) = \sum\nolimits_{i=1}^{n_k} (1 - \frac{\alpha}{2}) c_i^{(k)} y_i^- + \frac{\alpha}{2} c_i^{(k)} y_i^+ \tag{7}$$

$$S_\alpha^+(F) = \sum\nolimits_{i=1}^{n_k} \frac{\alpha}{2} c_i^{(k)} y_i^- + (1 - \frac{\alpha}{2}) c_i^{(k)} y_i^+ \tag{8}$$

s.t.

subject to the contraints

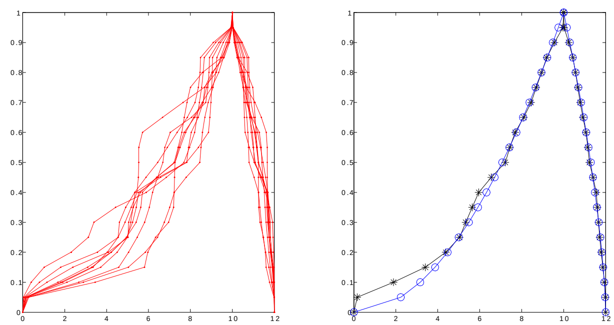
$$\begin{aligned} &F - y_i^- + y_i^+ = b_i^{(k)},\ i = 1, ..., n_k \\ &y_i^- \geq 0,\ i = 1, ..., n_k \\ &y_i^+ \geq 0,\ i = 1, ..., n_k \\ &F \text{ unconstrained.} \end{aligned}$$

Example.

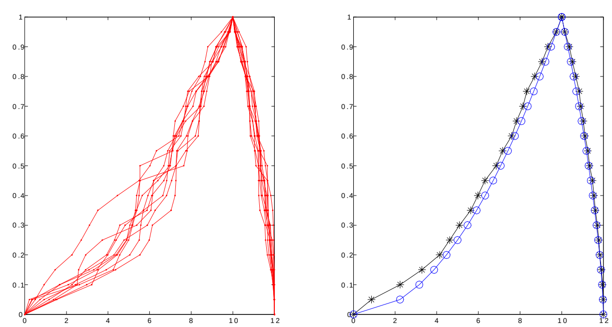
We consider the fuzzy number $u \in \mathbb{R}_\mathcal{F}$ having α-cuts $[u_\alpha^-, u_\alpha^+] = [10\alpha^{0.5}, 12 - 2\alpha^{1.5}]$ and we sample u_α^- and u_α^+ at $\frac{N}{2}$ points α_i generated on $[0, 1]$ as follows: $\alpha_1 = 0$, $\alpha_i = rand()$ and $\alpha_{\frac{N}{2}} = 1$ (N is even and $rand()$ is a uniform random number between 0 and 1). The N observations of u are given by $x_i = 10\alpha_i^{0.5}$ and $x_{\frac{N}{2}+i} = 12 - 2\alpha_i^{1.5}$, $i = 1, ..., \frac{N}{2}$. Then, using the N observations $x_1, ..., x_N$, the α-cuts $[u_k^-, u_k^+]$ of u are estimated for $\alpha_k = \frac{k-1}{20}$ ($k = 1, ..., 21$) by averaging the results of $P = \{10, 20, 50\}$ independent runs of (5)-(6).

Cite the convergence theorem of big numbers

The figures[1] below report the results for different values of N; when N increases, the estimated α-cuts become better.

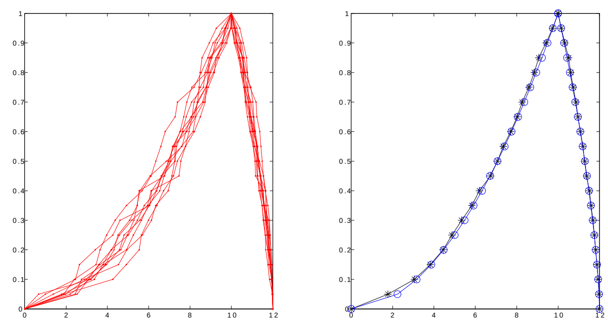


(1): $N/2 = 11$. 10 instances (left) and estimated (right,stars) ws exact (right,circles) α-cuts.

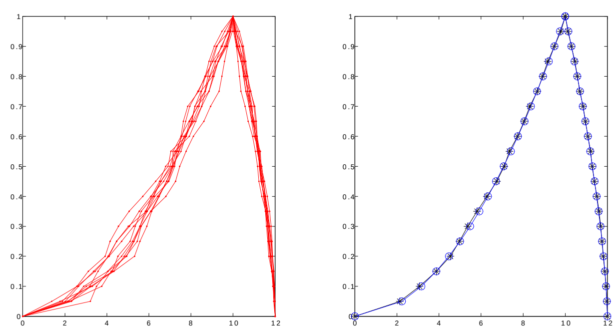


(2): $N/2 = 21$. 10 instances (left) and estimated (right,stars) ws exact (right,circles) α-cuts.

[1] All figures are performed through Matlab elaborations.



(3): $N/2 = 51$. 10 instances (left) and estimated (right,stars) ws exact (right,circles) α-cuts.



(4): $N/2 = 101$. 10 instances (left) and estimated (right,stars) ws exact (right,circles) α-cuts.

3 L_1-norm based F-transform

Insert definition of F-transform in the L1 norm case

We have seen that the direct F-transform, in combination with the $\frac{\alpha}{2}$-quantiles and $(1-\frac{\alpha}{2})$-quantiles, produces the α-cuts of a fuzzy-valued iF-transform of $f(x)$ on $[a, b]$.

An interesting problem is to relate the obtained membership function of the fuzzified $f(x)$ to the basic functions of the fuzzy partition $(\mathbb{P}, \mathbb{A})$.

We first consider the problem for the value $\alpha = 1$, i.e. the L_1-norm based values F_k obtained by minimizing

$$\Phi_k(F) = \int_a^b |f(x) - F| A_k(x)dx = \int_{x_{k-1}}^{x_{k+1}} |f(x) - F| A_k(x)dx$$

or, for a generalized fuzzy partition $(\mathbb{P}, \mathbb{A}^{(r)})$, the functional $(r \geq 1)$

$$\Phi_k(F) = \int_{x_{k-r}}^{x_{k+r}} |f(x) - F| A_k^{(r)}(x)dx.$$

The first step is to determine the possibility distribution functions $G_k(x)$, for each $k = 1, ..., n$, such that the minimization of $\Phi_k(F)$ produces the same result as the minimization of $\int_{x_{k-r}}^{x_{k+r}} |f(x) - F| dG_k(x)$; correspondingly, we can obtain the fuzzy numbers B_k with support $x \in [x_{k-r}, x_{k+r}]$ and with membership functions ($\widetilde{x}_k$ correspond to membership value 1)

$$B_k(x) = \begin{cases} 0 & \text{if} & x < x_{k-r} \\ B^L(x) & \text{if} & x_{k-r} \leq x \leq \widetilde{x}_k \\ B^R(x) & \text{if} & \widetilde{x}_k \leq x \leq x_{k+r} \\ 0 & \text{if} & x > x_{k+r} \end{cases}$$

such that

$$G_k(x) = \begin{cases} \frac{1}{2}B^L(x) & \text{if} & x_{k-r} \leq x \leq \widetilde{x}_k \\ 1 - \frac{1}{2}B^R(x) & \text{if} & \widetilde{x}_k \leq x \leq x_{k+r} \\ 0 & & \text{otherwise} \end{cases}.$$

For simplicity, we denote the basic functions by A_k instead of the more complete notation $A_k^{(r)}$; and assume that

$$A_k(x) = \begin{cases} 0 & \text{if} & x < x_{k-r} \\ A^L(x) & \text{if} & x_{k-r} \leq x \leq x_k \\ A^R(x) & \text{if} & x_k \leq x \leq x_{k+r} \\ 0 & \text{if} & x > x_{k+r} \end{cases}.$$

We will use the two integrals

$$\begin{aligned} I_k^L &= \int_{x_{k-r}}^{x_k} A_k^L(x)dx,\ I_k^R = \int_{x_k}^{x_{k+r}} A_k^R(x)dx \text{ and} \\ I_k &= I_k^L + I_k^R. \end{aligned}$$

Defining the integral functions

$$G_k(x) = \begin{cases} 0 & \text{if} & x \leq x_{k-r} \\ \frac{1}{I_k}\int_{x_{k-r}}^{x} A_k(x)dx & \text{if} & x \in [x_{k-r}, x_{k+r}] \\ 1 & \text{if} & x \geq x_{k+r} \end{cases},\ k = 1, ..., n$$

and assuming that each $A_k(x)$ is a continuous function, we have immediately that

$$\Phi_k(F) = I_k \int_{x_{k-r}}^{x_{k+r}} |f(x) - F| dG_k(x)$$

and the minimization (with respect to F) of $\int_{x_{k-r}}^{x_{k+r}} |f(x) - F| dG_k(x)$ is equivalent to the minimization of $\Phi_k(F)$.

4 A variant to F-Transform

A *fuzzy partition* of a given real compact interval $[a,b]$ is constructed by a decomposition $\mathbb{P} = \{a = x_1 < x_2 < ... < x_n = b\}$ of $[a,b]$ into $n-1$ subintervals $[x_{k-1}, x_k]$, $k = 2, ..., n$ and by a family $\mathbb{A} = \{A_1, A_2, ..., A_n\}$ of n fuzzy numbers (the *basic functions*), identified by their membership functions $A_1(x), A_2(x), ..., A_n(x)$ on $[a,b]$ and with the properties (to complete this notation, we set $x_0 = a$ and $x_{n+1} = b$):

1. $A_k : [a,b] \longrightarrow [0,1]$ is continuous with $A_k(x_k) = 1$, $A_k(x) = 0$ for $x \notin [x_{k-1}, x_{k+1}]$;
2. for $k = 2, 3, ..., n$, A_k is increasing on $[x_{k-1}, x_k]$ and decreasing on $[x_k, x_{k+1}]$; A_1 is decreasing on $[a, x_2]$; A_n is increasing on $[x_{n-1}, b]$;
3. for all $x \in [a,b]$ the partition of unity condition holds $\sum_{k=1}^{n} A_k(x) = 1$.

We denote a fuzzy partition by the pair $(\mathbb{P}, \mathbb{A})$.

The basic functions can be defined by monotonic branches $A_1(x) = R_1(x)$ for $x \in [a, x_2]$, $A_n(x) = L_n(x)$ for $x \in [x_{n-1}, b]$ and

$$A_k(x) = \begin{cases} L_k(x) & \text{if} \quad x \in [x_{k-1}, x_k] \\ R_k(x) & \text{if} \quad x \in [x_k, x_{k+1}] \end{cases} \quad \text{for } k = 2, ..., n-1 \tag{9}$$

where each function $L_k(x)$ is increasing with $L_k(x_{k-1}) = 0$, $L_k(x_k) = 1$ and each $R_k(x)$ is decreasing with $R_k(x_k) = 1$, $R_k(x_{k+1}) = 0$.

The support of each basic function $A_k(x)$ is the compact interval $[x_{k-1}, x_{k+1}]$ so that, on each subinterval $[x_{k-1}, x_k]$ of the decomposition $\mathbb{P}$ only two basic functions $A_{k-1}(x)$ and $A_k(x)$ are non zero and we can introduce the conditions

$$L_k(x) + R_{k-1}(x) = 1 \text{ for all } x \in [x_{k-1}, x_k] \text{ and } k = 2, ..., n$$

and use only $n-1$ functions $L_k(x)$, $k = 2, ..., n$ (or $R_k(x)$, $k = 1, 2, ..., n-1$) to define a fuzzy partition $(\mathbb{P}, \mathbb{A})$:

$$\begin{aligned} A_1(x) &= 1 - L_2(x) \text{ for } x \in [a, x_2], \\ A_k(x) &= \begin{cases} L_k(x) & \text{if} \quad x \in [x_{k-1}, x_k] \\ 1 - L_{k+1}(x) & \text{if} \quad x \in [x_k, x_{k+1}] \end{cases} \quad \text{for } k = 2, ..., n-1, \\ A_n(x) &= L_n(x) \text{ for } x \in [x_{n-1}, b]. \end{aligned} \tag{10}$$

So, each fuzzy partition is essentially defined by decomposition $\mathbb{P}$ and by $n-1$ increasing functions $L_k(x)$, $k = 2, 3, ..., n$ such that

$$L_k(x_{k-1}) = 0 \text{ and } L_k(x_k) = 1.$$

Given a continuous function $f : [a, b] \longrightarrow \mathbb{R}$ and a fuzzy partition $(\mathbb{P}, \mathbb{A})$, the *direct fuzzy transform* (F-transform) of f with respect to $(\mathbb{P}, \mathbb{A})$ is the n-tuple of real numbers $F = (F_1, F_2, ..., F_n)^T$ (notation $(\cdot)^T$ means transposition) such that each F_k, given by

$$F_k = \frac{\int_a^b f(x)A_k(x)dx}{\int_a^b A_k(x)dx}, \; k = 1, 2, ..., n, \tag{11}$$

is the unique minimizer of the convex functional

$$\Phi_k(F) = \int_a^b [f(x) - F]^2 A_k(x)dx \tag{12}$$

i.e.

$$F_k = \arg\min_F \Phi_k(F) \Longleftrightarrow F_k = \frac{\int_a^b f(x)A_k(x)dx}{\int_a^b A_k(x)dx}.$$

Each F_k is a local wighted least squares approximation of function $f(x)$ with weighting function $A_k(x)$ so that F_k is a wighted average of $f(x)$ on $[x_{k-1}, x_{k+1}]$.

The setting (12) can be generalized by considering a real vlue $p > 0$ and defining the L_p-norm based approximation of function $f(x)$ on $[x_{k-1}, x_{k+1}]$, by minimizing the functional

$$\Phi_k^{(p)}(F) = \int_a^b |f(x) - F|^p A_k(x)dx$$

and by setting

$$F_k^{(p)} = \arg\min_F \Phi_k^{(p)}(F). \tag{13}$$

Of particular interest is the case $p = 1$,

$$\Phi_k^{(1)}(F) = \int_a^b |f(x) - F| A_k(x)dx.$$

L_1-norm based approximation defines the direct F-transform component $F_k^{(1)}$ as the minimizer of the absolute deviations $|f(x) - F|$ on $[x_{k-1}, x_{k+1}]$ (weighted by $A_k(x)$). In statistical terms, the minimizer of $\Phi_k^{(1)}(F)$ produces an analogous of the weighted median of function $f(x)$ on $[x_{k-1}, x_{k+1}]$.

In the discrete case, assume that f is given at a sufficiently dense set of m points $t_1 < t_2 < ... < t_m$; the functional to be minimized is

$$\Psi_k(F) = \sum_{j=1}^{m} |f(t_j) - F| A_k(t_j).$$

Function $\Psi_k(F)$ is nondifferentiable and an explicit closed-form solution is not available. The minimization of $\Psi_k(F)$ can be obtained by linear programming.

For $k = 1, 2, ..., n$ let j_k and n_k be the integers such that $t_j \in [x_{k-1}, x_{k+1}]$, $k = 1, 2, ..., n$ for $j = j_k + 1, ..., j_k + n_k$. We have $n_k \geq 1$ because the points t_j are sufficiently dense with respect to the partition $(\mathbb{P}, \mathbb{A})$. Then

$$\Psi_k(F) = \sum_{i=1}^{n_k} |f(t_{j_k+i}) - F| A_k(t_{j_k+i})$$

and a linear programming formulation can be obtained as follows: define $2n_k$ nonnegative variables y_i^-, y_i^+, $i = 1, 2, ..., n_k$ such that

$$y_i^+ - y_i^- + F = f(t_{j_k+i}),\ i = 1, 2, ..., n_k$$

and

$$y_i^- + y_i^+ = |F - f(t_{j_k+i})|,\ i = 1, 2, ..., n_k.$$

Then

$$\Psi_k = \sum_{i=1}^{n_k} (y_i^- + y_i^+) A_k(t_{j_k+i})$$

and, defining $c_i^{(k)} = A_k(t_{j_k+i})$ and $b_i^{(k)} = f(t_{j_k+i})$, $i = 1, ..., n_k$, $k = 1, ..., n$, the minimization of Ψ_k becomes

$$\begin{aligned} &\min\ \sum_{i=1}^{n_k} c_i^{(k)} y_i^- + c_i^{(k)} y_i^+ \\ &\text{s.t.} \\ &\qquad F - y_i^- + y_i^+ = b_i^{(k)},\ i = 1, ..., n_k \\ &\qquad y_i^- \geq 0 \\ &\qquad y_i^+ \geq 0 \\ &\qquad F \text{ unconstrained} \end{aligned} \tag{14}$$

The dual of (14) is

$$\begin{aligned} &\max\ \sum_{j=1}^{n_k} b_j^{(k)} w_j \\ &\text{s.t.} \\ &\qquad \textstyle\sum_{j=1}^{n_k} w_j = 0 \\ &\qquad -c_j^{(k)} \leq w_j \leq c_j^{(k)},\ j = 1, ..., n_k \end{aligned} \tag{15}$$

We see immediatly that problem (15) is feasible ($w_j = 0\ \forall j = 1, ..., n_k$ is feasible) and is also bounded ($\sum_{j=1}^{n_k} b_j^{(k)} w_j \leq \sum_{j=1}^{n_k} |b_j^{(k)}| c_j^{(k)}$, because $c_j^{(k)} \geq 0$).

It follows that

a) problem (14) always has an optimal bounded solution,

b) the optimal value F_k corresponding to the optimal solution of (14) is obtained by

$$F_k = \frac{1}{n_k} \sum\nolimits_{i=1}^{n_k} (y_i^- - y_i^+ + b_i^{(k)})$$

Given a continuous function $f : [a,b] \longrightarrow \mathbb{R}$, a fuzzy partition $(\mathbb{P}, \mathbb{A})$ and the L_1-norm based direct F-transform $(F_1, F_2, ..., F_n)^T$ of f with respect to $(\mathbb{P}, \mathbb{A})$, the L_1-norm based *inverse F-transform* is the continuous function $\widehat{f}_F : [a,b] \longrightarrow \mathbb{R}$ given by

$$\widehat{f}_F(x) = \sum_{k=1}^{n} F_k A_k(x) \text{ for } x \in [a,b]. \tag{16}$$

REMARK

The same construction can be obtained in terms of an extended fuzzy partition $(\mathbb{P}, \mathbb{A}^{(r)})$ of bandwidth $r \geq 1$, where each basic function $A_k^{(r)}(x)$ is defined to "cover" the subinterval $[x_{k-r}, x_{k+r}]$, $k = 2-r, ..., 1, 2, ..., n, ..., n+r-1$.

Now, it is not difficult to determine the fuzzy number $B_k(x)$, corresponding to the possibility distribution $G_k(x)$, in terms of A_k^L and A_k^R.

Omitting the details, define $\widetilde{x}_k \in [x_{k-r}, x_{k+r}]$ such that

$$\int_{x_{k-r}}^{\widetilde{x}_k} A_k(x)dx = \frac{I_k}{2} = \frac{I_k^L + I_k^R}{2}.$$

We have three cases:

(1) if $\widetilde{x}_k = x_k$, i.e. if $I_k^L = I_k^R$, then

$$\begin{aligned} B_k^L(x) &= \frac{2}{I_k} \int_{x_{k-r}}^{x} A_k^L(x)dx \\ B_k^R(x) &= 1 - \frac{2}{I_k} \int_{x_k}^{x} A_k^R(x)dx. \end{aligned}$$

(2) if $\widetilde{x}_k < x_k$, i.e. if $I_k^L > I_k^R$, then

$$\begin{aligned} B_k^L(x) &= \frac{2}{I_k} \int_{x_{k-r}}^{x} A_k^L(x)dx \\ B_k^R(x) &= \begin{cases} 1 - \frac{2}{I_k} \int_{\widetilde{x}_k}^{x} A_k^R(x)dx & \text{if} \quad \widetilde{x}_k \leq x \leq x_k \\ \frac{2I_k^R}{I_k} - \frac{2}{I_k} \int_{x_k}^{x} A_k^R(x)dx & \text{if} \quad x_k \leq x \leq x_{k+r} \end{cases}. \end{aligned}$$

(3) if $x_k < \widetilde{x}_k$, i.e. if $I_k^L < I_k^R$, then

$$B_k^L(x) = \begin{cases} \frac{2}{I_k} \int\limits_{x_{k-r}}^{x} A_k^L(x)dx & \text{if} \quad x \le x_k \\ \frac{2I_k^L}{I_k} + \frac{2}{I_k} \int\limits_{x_k}^{x} A_k^R(x)dx & \text{if} \quad x_k \le x \le \widetilde{x}_k \end{cases}$$

$$B_k^R(x) = 1 - \frac{2}{I_k} \int\limits_{\widetilde{x}_k}^{x} A_k^R(x)dx.$$

The discussion above shows a relevant fact: the α-cuts $[F_{k,\alpha}^-, F_{k,\alpha}^+]$, estimated by computing the $\frac{\alpha}{2}$-quantiles and $(1 - \frac{\alpha}{2})$-quantiles for $\alpha \in]0, 1]$, by L_1-norm based F-transform, define a fuzzy number F_k such that, for $x \in [x_{k-r}, x_{k+r}]$, the membership value of $f(x)$ with respect to F_k is $B_k(x)$.

We can prove some additional theoretical results about this results.

Crouzet in [2] introduces the fuzzy projection that is closely related to fuzzy and inverse fuzzy transforms.

The fuzzy projection, as the inverse fuzzy transform, is another way to address the sampling/interpolation problem in a fuzzy context. The main advantage is that they are numerically very simple to implement and they require little computation. We have also shown that interpolation kernels are easy to compute, numerically stable, address the requirement of being very well localized.

The fuzzy projection is defined as

$$Pf = Arg \min_{g \in \Pi} \|f - g\|_2$$

and it minimizes classical global least square criteria, while the inverse fuzzy transform

$$Qf = F * Ff$$

is computed using the components of the fuzzy transform $(Ff)_k$ that minimize the $p + 2$ local functionals

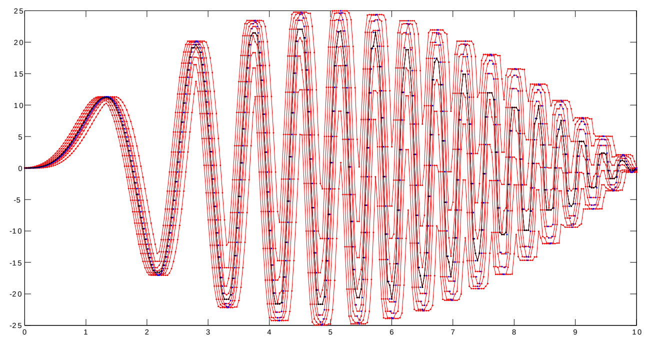
$$\Phi_k(y) = \int_0^d (y - f(x))^2 \mu_k(x) \, dx$$

For the moment, we illustrate a computational example.

The function (as in Crouzet) is $f(x) = x(10 - x)sin(x^2)$ on $[0, 10]$. Two runs are executed with $m = 601$ and $m = 1201$ points; the number n of basic functions is 301 and $r = 5$.

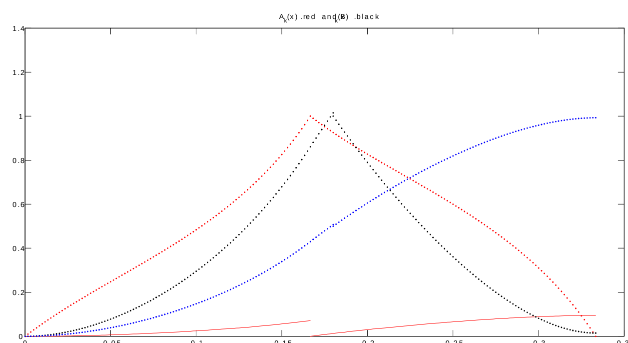
The quantile F-transform to determine the α-cuts for $\alpha \in \{0.05, 0.25, 0.5, 0.75, 1.0\}$ is illustrated in the figure below[2]

[2] The routines for quantile L1-based F-transform are available.



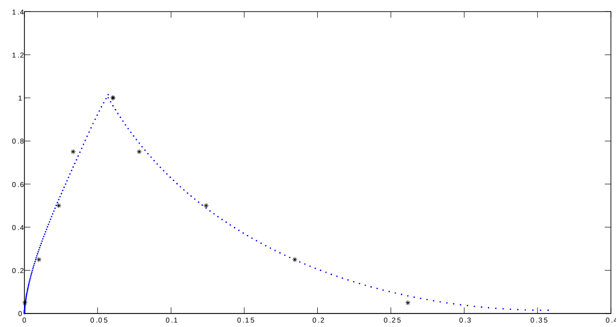
Quantile (L1-norm based) F-transform for example

and the membership functions $A_k(x)$, $B_k(x)$ and the corresponding possibility distribution are illustrated in the next figure;

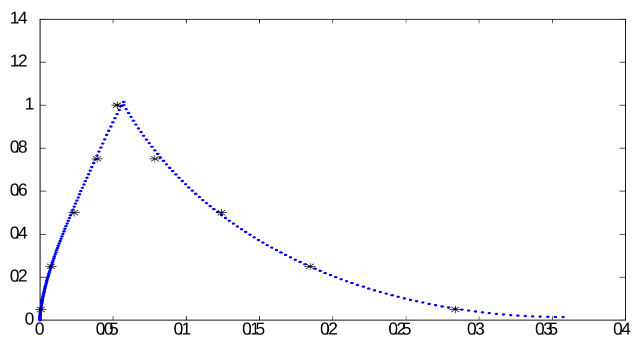


Membership functions $A_k(x)$ (red-dots), $B_k(x)$ (black-dots) and corresponding possibility distribution $G_k(x)$ (blue-dots)

The membership function of $f(x)$ derived from $B_k(x)$ is then computed and compared with the estimated α-cuts for the two runs with $m = 601$ and $m = 1201$ data points; with $m = 601$ data, the computed $[F^{-}_{k,\alpha}, F^{+}_{k,\alpha}]$ by L1-based F-transform are illustrated in the two final figures; observe that $m = 1201$ observations give α-cuts quite near to the membership function of $f(x)$ corresponding to $B_k(x)$ (here, $k = r + 1$)



Theoretical (dot-blue) and empirical (star-black) values of membership function of $f(x)$ from L1-norm based F-transform; case of $m = 601$ data points and $n = 301$, $r = 5$.



Theoretical (dot-blue) and empirical (star-black) values of membership function of $f(x)$ from L1-norm based F-transform; case of $m = 1201$ data points and $n = 301$, $r = 5$.

It is immediate to see that more data produce, as expected, more precise estimated α-cuts.

5 Use of the L_1-norm based F-transform as a fuzzification tool

Possibly important consequence

We have seen that fixing the basic functions $A_k(x)$ induces a family of membership functions $B_k(x)$ or, equivalently, a family of possibility distribution functions $G_k(x)$.

So, instead of chosing the basic functions $A_k(x)$ we can chose directly the family $\{G_k | k = 1, ..., n\}$ of possibility distributions and determine the (direct) L1-norm based F-transform by minimizing

$$\Phi_k(F) = \int\limits_{x_{k-r}}^{x_{k+r}} |f(x) - F| dG_k(x)$$

and similarly, for the $\frac{\alpha}{2}$-quantiles and $(1 - \frac{\alpha}{2})$-quantiles, by minimizing (the integrals are Stieltjes)

$$\begin{aligned}
\Phi^{(-)}_{k,\alpha}(F) &= (1 - \frac{\alpha}{2}) \int\limits_a^b \langle f(x) - F \rangle^- \, dG_k(x) + \frac{\alpha}{2} \int\limits_a^b \langle f(x) - F \rangle^+ \, dG_k(x) \\
\Phi^{(+)}_{k,\alpha}(F) &= \frac{\alpha}{2} \int\limits_a^b \langle f(x) - F \rangle^- \, dG_k(x) + (1 - \frac{\alpha}{2}) \int\limits_a^b \langle f(x) - F \rangle^+ \, dG_k(x).
\end{aligned}$$

Final comment

Any family of distributions $\mathbb{G} = \{G_k | k = 1, ..., n\}$ *defined on a decomposition* P *of* $[a, b]$ *will produce a fuzzy (or possibilistic) partition* $(\mathbb{P}, \mathbb{G})$ *and a corresponding fuzzy-valued F-transform.*

6 Conclusion

We have proposed

References

[1] I. Couso, D. Dubois, Statistical reasoning with set-valued information: Ontic vs. epistemic views, *International Journal of Approximate Reasoning* (2013), http://dx.doi.org/10.1016/j.ijar.2013.07.002.

[2] J. F. Crouzet, Fuzzy projection versus inverse fuzzy transform as sampling/interpolation schemes, *Fuzzy Sets and Systems*, 193 (2012) 108-121.

[3] P. Craven, G. Wahba, Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 31 (1979) 377–403.

[4] F. Di Martino, V. Loia, S. Sessa, Fuzzy transforms method and attribute dependency in data analysis, *Information Sciences*, 180 (2010) 493-505.

[5] F. Di Martino, V. Loia, S. Sessa, Fuzzy transforms for compression and decompression of color videos, *Information Sciences*, 180 (2010) 3914-3931.

[6] D. Dubois, H. Prade, Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York, 1980.

[7] R. Eubank, *Spline Smoothing and Nonparametric Regression*, M. Dekker, 1988.

[8] A. Kovac, Smooth functions and local extreme values, *Computational Statistics & Data Analysis*, 51 (2007) 5155-5171.

[9] B. Liu, A survey of credibility theory, *Fuzzy Optimization Decision Making*, 5 (2006) 387–408.

[10] K.M. Mullen, M. Vengris, I.H.M. van Stokkum, Algorithms for separable nonlinear least squares with application to modelling time-resolved spectra, *Journal of Global Optimization*, 38 (2007) 201-213.

[11] I. Perfilieva, Fuzzy Transforms: Theory and Applications, *Fuzzy Sets and Systems*, 157 (2006) 993-1023.

[12] I. Perfilieva, Fuzzy Transforms: A Challenge to conventional transforms. In P.W. Hawkes (Ed.), *Advances in Images and Electron Physics*, 147, Elsevier Academic Press, (2007) 137-196.

[13] I. Perfilieva, Fuzzy Transforms and their applications to data compression. In I. Bloch et al. (Eds), *Fuzzy Logic and Applications, LNAI 3849, Springer,* (2006) 19-31.

[14] I. Perfilieva, M. Dancova, Image fusion on the basis of fuzzy transform, in D. Ruan et Al. (Eds) "Computational Intelligence in Decision and Control", World Scientific, (2008) 471-476.

[15] I. Perfilieva, B. De Baets, Fuzzy transforms of monotone functions with applications to image compression, *Information Sciences*, 180 (2010) 3304-3315.

[16] I. Perfilieva, V. Novak, A. Dvorak, Fuzzy transform in the analysis of data, *International Journal of Approximate Reasoning*, 48 (2008) 36-46.

[17] S.K. Schnabel, P.H.C. Eilers, Optimal expectile smoothing, *Compuational Statistics and Data Analysis*, 53 (2009) 4168-4177.

[18] J.S. Simonoff, *Smoothing Methods in Statistics*, Springer Series in Statistics, Springer Verlag, 1996.

[19] L. Stefanini, Fuzzy Transform with Parametric LU-Fuzzy Partitions, in D. Ruan et Al. (Eds), *Computational Intelligence in Decision and Control*, World Scientific, (2008) 399-404.

[20] L. Stefanini, Fuzzy Transform and Smooth Functions, *Proceeding of the IFSA-EUSFLAT 2009 Conference, Lisbon, July 2009*, edited by J.P. Carvalho, D. Dubois, U. Kaymak, J.M.C. Sousa, 579-584.

[21] L. Stefanini, L. Sorini, M.L. Guerra, Parametric representations of fuzzy numbers and application to fuzzy calculus, *Fuzzy Sets and Systems*, 157 (2006) 2423–2455.

[22] L. Stefanini, L. Sorini, M.L. Guerra, Fuzzy Numbers and Fuzzy Arithmetic. In W. Pedrycz, A. Skowron, V. Kreynovich (Eds), *Handbook of Granular Computing*, Chapter 12, J. Wiley & Sons, 2009.

[23] L.Stefanini, F-transform with parametric generalized fuzzy partitions, *Fuzzy Sets and Systems*, 180 1 (2011) 98-120.

[24] L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems,* 1 (1978) 3–28.