Undefined 0 (2019) 0
IOS Press

Constructive approximation of level continuous fuzzy functions [1]

Juan J. Font [a,*], Delia Sanchis [a] and Manuel Sanchis [a]

[a] *Departament de Matemàtiques, and Institut Universitari de Matemàtiques i Aplicacions de Castelló (IMAC), Universitat Jaume I, Campus del Riu Sec. s/n, 12071 Castelló, Spain*
E-mail: font@uji.es, dsanchis@uji.es, sanchis@uji.es

Abstract. We consider the space of continuous functions defined between a locally compact Hausdorff space and the space of fuzzy numbers endowed with the level convergence topology. We obtain a Stone-Weierstrass type theorem for such space of functions equipped with the compact open topology.

As a corollary of the above results, we prove that such functions can be approximated by certain fuzzy-number-valued neural networks and sums of fuzzy-number-valued ridge functions.

Keywords: Stone-Weierstrass theorem, fuzzy-number-valued continuous functions, level convergence, neural networks, ridge functions

1. Introduction

Fuzzy Analysis has developed based on the notion of fuzzy number just as much as classical Real Analysis did based on the concept of real number. Fuzzy-number-valued functions, that is, functions defined on a topological space taking values in the space of fuzzy numbers, play a central role in Fuzzy Analysis as real-valued functions do in the classical setting. Namely, fuzzy-number-valued functions have become the main tool in several fuzzy contexts, such as fuzzy differential equations ([3]), fuzzy integrals ([27]) or fuzzy optimization ([13]). However the primary drawback of dealing with these functions is the fact that the space they form is not a linear space; indeed it is not a group with respect to addition.

The main question in Approximation Theory, a fundamental branch of Mathematical Analysis, is whether a given family of functions from which we plan to approximate is dense in the set of functions we wish to approximate. That is, can we approximate any function in our set, arbitrarily well, using finite linear combinations of functions from our given family? In the fuzzy setting, most results in this topic deal with the approximation capabilities of fuzzy neural networks (see e.g., [20], [14], [15] and [11]) which turn out to be different from the capabilities of classical neural networks (see Section 4). It is known that neural networks are particularly useful in many domains, such as finance, medicine, mechanical engineering, geology, computer science, etc. Generally speaking, neural networks are implemented in all situations where forecasting, decision, classification and control problems arise. Since nature and human brain are inherently fuzzy in characteristic, it is natural to think that fuzzy neural networks have the ability for processing fuzzy information thanks to their learning abilities (which are closely related to their approximation capabilities). Related to these, the so-called ridge functions are also an important tool in Approximation Theory, although they have not been used in a fuzzy context yet. Indeed the term "ridge function" appeared after the seminal paper [21] about an approximation problem in computer tomog-

[1] This research is supported by Spanish Government (grant MTM2016-77143-P, AEI/FEDER, UE) and Generalitat Valenciana (Projecte GV/2018/110).

* Corresponding author. E-mail: font@uji.es.

raphy. They also arise naturally in several fields such as partial differential equations and neural networks ([19]).

In this paper we focus on finding dense (with respect to the compact-open topology) subspaces of the space of continuous (with respect to the level convergence topology) fuzzy-number-valued functions defined on a locally compact Hausdorff. The level convergence topology was introduced in [17] and, thanks to Goetschel-Voxman's characterization of fuzzy numbers (see Section 2 below), has become a natural alternative to the usual metrics (d_∞, d_p, sendograph, ...) used in $\mathbb{E}^1$. Indeed, the space of level continuous fuzzy-number-valued functions is strictly larger than the space of d_∞-continuous fuzzy-number-valued functions. Furthermore, as far as the authors are aware, the use of the compact-open topology is also a novelty in this context ([2]).

As a corollary of the above results, we provide fuzzy-number-valued neural networks and sums of fuzzy-number-valued ridge functions which are dense in $C(\mathbb{R}^n, \mathbb{E}^1)$.

In the final sections, we provide a numerical example to illustrate the technique used in our main result and describe several possible applications in decision making and fuzzy optimization problems.

2. Preliminaries

Let $F(\mathbb{R})$ denote the family of all fuzzy subsets on the real numbers $\mathbb{R}$ (see, e.g., [6]). For $u \in F(\mathbb{R})$ and $\lambda \in [0,1]$, the λ-level set of u is defined by

$$[u]^\lambda := \{x \in \mathbb{R} : u(x) \geq \lambda\}, \quad \lambda \in]0,1],$$

$$[u]^0 := cl_{\mathbb{R}}\{x \in \mathbb{R} : u(x) > 0\}.$$

The fuzzy number space $\mathbb{E}^1$ is the set of elements u of $F(\mathbb{R})$ satisfying the following properties (see, e.g., [6]):

1. u is normal, i.e., there exists an $x_0 \in \mathbb{R}$ with $u(x_0) = 1$;
2. u is convex, i.e., for all $x, y \in \mathbb{R}, \lambda \in [0,1]$,

$$u(\lambda x + (1-\lambda)y) \geq \min\{u(x), u(y)\};$$

3. u is upper-semicontinuous;
4. $[u]^0$ is a compact set in $\mathbb{R}$.

Notice that if $u \in \mathbb{E}^1$, then the λ-level set $[u]^\lambda$ of u is a compact interval for each $\lambda \in [0,1]$. We denote $[u]^\lambda = [u^-(\lambda), u^+(\lambda)]$. Every real number r can be considered a fuzzy number since r can be identified with the fuzzy number $\tilde{r}$ defined as

$$\tilde{r}(t) := \begin{cases} 1 \text{ if } t = r, \\ 0 \text{ if } t \neq r. \end{cases}$$

We can now state the characterization of fuzzy numbers provided by Goetschel and Voxman ([12]):

Let $u \in \mathbb{E}^1$ and $[u]^\lambda = [u^-(\lambda), u^+(\lambda)]$, $\lambda \in [0,1]$. Then the pair of functions $u^-(\lambda)$ and $u^+(\lambda)$ has the following properties:

1. $u^-(\lambda)$ is a bounded left continuous nondecreasing function on $(0,1]$;
2. $u^+(\lambda)$ is a bounded left continuous nonincreasing function on $(0,1]$;
3. $u^-(\lambda)$ and $u^+(\lambda)$ are right continuous at $\lambda = 0$;
4. $u^-(1) \leq u^+(1)$.

Conversely, if a pair of functions $\alpha(\lambda)$ and $\beta(\lambda)$ satisfy the above conditions (i)-(iv), then there exists a unique $u \in \mathbb{E}^1$ such that $[u]^\lambda = [\alpha(\lambda), \beta(\lambda)]$ for each $\lambda \in [0,1]$.

Example 2.1. Let $u(x)$ be a fuzzy number defined as

$$u(x) = \begin{cases} 0 & x \notin [0,1], \\ 1 & x \in [0, \frac{1}{2}], \\ \frac{1}{2} & x \in (\frac{1}{2}, 1]. \end{cases}$$

Then, its corresponding $u^-(\lambda)$ and $u^+(\lambda)$ functions turn out to be

$$u^-(\lambda) = 0, u^+(\lambda) = \begin{cases} 1 & \lambda \in [0, \frac{1}{2}], \\ \frac{1}{2} & \lambda \in (\frac{1}{2}, 1]. \end{cases}$$

Given $u, v \in \mathbb{E}^1$ and $k \in \mathbb{R}$, we can define $u + v := [u^-(\lambda), u^+(\lambda)] + [v^-(\lambda), v^+(\lambda)]$ and $ku := k[u^-(\lambda), u^+(\lambda)]$. It is well-known that $\mathbb{E}^1$ endowed with these two natural operations is not a vector space. Indeed $(\mathbb{E}^1, +)$ is not a group ([6]).

The space of fuzzy numbers is usually endowed with the topology induced by certain metrics, mainly by the supremum metric d_∞; namely, for $u, v \in \mathbb{E}^1$, $d_\infty(u,v) := \sup_{\lambda \in [0,1]} d_H([u]^\lambda, [v]^\lambda)$, where d_H

stands for the Hausdorff metric. That is, $d_H([u]^\lambda, [v]^\lambda) = \max\{| u^-(\lambda) - v^-(\lambda) |, | u^+(\lambda) - v^+(\lambda) |\}$. It is well-known (see, e.g., [6]) that $(\mathbb{E}^1, d_\infty)$ is a complete metric space. In [17], the authors introduced a new topology on $\mathbb{E}^1$ based on the following convergence:

We say that the net $\{u_k\}_{k\in D} \subset \mathbb{E}^1$ *levelly converges* to $u \in \mathbb{E}^1$ if $\lim_k d_H([u_k]^\lambda, [u]^\lambda) = 0$ for any $\lambda \in [0, 1]$ or, equivalently, if $\lim_k u_k^+(\lambda) = u^+(\lambda)$ and $\lim_k u_k^-(\lambda) = u^-(\lambda)$ for each $\lambda \in [0, 1]$.

Notice that a net $\{u_k\}_{k\in D} \subset \mathbb{E}^1$ d_∞-converges to $u \in \mathbb{E}^1$ if and only if $\lim_k u_k^+$ converges uniformly to u^+ and $\lim_k u_k^-$ converges uniformly to u^-. Thus, the d_∞-convergence implies the level convergence. The converse fails to be true (see Example 2.1 in [26]).

In [8], [9] and [10], the authors studied the topology τ_ℓ associated with this level convergence in $\mathbb{E}^1$. Thus, it is known that $(\mathbb{E}^1, \tau_\ell)$ is a Hausdorff, separable, Baire, first countable topological space. Also a local basis for $u \in \mathbb{E}^1$ in τ_ℓ is of the form

$$U(u, \{\lambda_1, ..., \lambda_n\}, \epsilon) := \{v \in \mathbb{E}^1 :$$

$$\max_{1\le j\le n} \{d_H([v]^{\lambda_j}, [u]^{\lambda_j})\} < \epsilon\},$$

for $\{\lambda_1, ..., \lambda_n\} \subset [0, 1]$ and $\epsilon > 0$.

Let X be a locally compact space. By the above paragraph, it is apparent that the space of continuous functions $C(X, (\mathbb{E}^1, \tau_\ell))$ contains $C(X, (\mathbb{E}^1, d_\infty))$; indeed, by [9, page 429], it is a proper subspace. In the sequel, we shall endow $C(X, (\mathbb{E}^1, \tau_\ell))$ with the compact-open topology. That is, a local basis at $f_0 \in C(X, (\mathbb{E}^1, \tau_\ell))$ is formed by sets of the form

$$V(f_0, K, \{\lambda_1, ..., \lambda_n\}, \epsilon) :=$$

$$= \{f \in C(X, (\mathbb{E}^1, \tau_\ell)) : d_H([f_0(x)]^{\lambda_j}, [f(x)]^{\lambda_j}) < \varepsilon$$

$$\text{for all } x \in K, j = 1, ..., n\}$$

for a compact subset K of X, $\{\lambda_1, ..., \lambda_n\} \subset [0, 1]$ and $\epsilon > 0$.

3. A Stone-Weierstrass type theorem in Fuzzy Analysis.

Let us first recall the following classical result known as Uryshon's Lemma:

Lemma 3.1. Let X be a locally compact Hausdorff space, and let $K, F \subset X$ be two disjoint sets, with K compact, and F closed. Then there exists a continuous function $f : X \longrightarrow [0, 1]$ such that $f \equiv 1$ on K and f vanishes on F.

Given $u \in \mathbb{E}^1$, we shall write $\widehat{u}$ to denote the function in $C(X, \mathbb{E}^1)$ which takes the constant value u.

We can now state and prove a version of the Stone-Weierstrass theorem for $C(X, (\mathbb{E}^1, \tau_l))$ endowed with the compact-open topology:

Theorem 3.2. Let $\mathcal{H}$ be a subspace of $C(X, (\mathbb{E}^1, \tau_l))$ which contains the finite combinations of the form $\psi_1\widehat{u_1} + ... + \psi_m\widehat{u_m}$, where $\psi_i \in C(X, [0, 1])$ and $u_i \in \mathbb{E}^1$, $i = 1, ..., m$. Then $\mathcal{H}$ is dense in $C(X, (\mathbb{E}^1, \tau_l))$.

Proof. Fix $\varepsilon > 0$, $\{\lambda_1, ..., \lambda_n\} \subset [0, 1]$, a compact subset $K \subset X$ and $f_0 \in C(X, (\mathbb{E}^1, \tau_l))$. For each $x \in X$ and $0 < \varepsilon(x) < \varepsilon$, let us define

$$N(x) = \{t \in X : d_H([f_0(t)]^{\lambda_j}, [f_0(x)]^{\lambda_j}) < \varepsilon(x)$$

$$< \varepsilon, j = 1, ..., n\},$$

which is an open neighborhood of x since $f_0 \in C(X, (\mathbb{E}^1, \tau_l))$. As X is locally compact, we can find a relatively compact neighborhood of x, $V(x)$, such that $clV(x) \subset N(x)$. By Uryshon's Lemma (Lemma 3.1), there exists a continuous function $f_x : X \longrightarrow [0, 1]$ such that $f_x \equiv 1$ on $clV(x)$ and f_x vanishes on $X \setminus N(x)$.

Choose $x_1 \in K$. By the compacity of $X \setminus N(x_1) \cap K$, there is a finite set $\{x_2, \ldots, x_m\} \subset X \setminus N(x_1) \cap K$ such that $X \setminus N(x_1) \cap K \subset V(x_2) \cup \ldots \cup V(x_m)$. Define $\varepsilon' = \max\{\varepsilon(x_i) : 1 \le i \le m\}$.

Let us define the functions
$\psi_2 := f_{x_2}$,
$\psi_3 := (1 - f_{x_2})f_{x_3}$,

$\vdots$
$\psi_m := (1-f_{x_2})(1-f_{x_3})\cdots(1-f_{x_{m-1}})f_{x_m}$.

Next we claim that

$$\psi_2+\ldots+\psi_j = 1-(1-f_{x_2})(1-f_{x_3})\cdots(1-f_{x_j}),$$

$j=2,\ldots,m$. Indeed, it is clear that

$$\psi_2+\psi_3 = f_{x_2}+(1-f_{x_2})f_{x_3} = 1-(1-f_{x_2})\cdot(1-f_{x_3}).$$

We proceed by induction. Assume that the result is true for a certain $j\in\{4,...,m-1\}$ and let us check

$$\psi_2+\ldots+\psi_j+\psi_{j+1} = 1-(1-f_{x_2})(1-f_{x_3})\cdots$$

$$(1-f_{x_j})(1-f_{x_{j+1}}).$$

Namely,

$$\psi_2+\ldots+\psi_j+\psi_{j+1} = 1-(1-f_{x_2})(1-f_{x_3})\cdots$$

$$(1-f_{x_j})+(1-f_{x_2})(1-f_{x_3})\cdots(1-f_{x_j})f_{x_{j+1}} =$$

$$= 1-(1-f_{x_2})(1-f_{x_3})\cdots(1-f_{x_j})(1-f_{x_{j+1}}),$$

as was to be checked. Finally, we can define $\psi_1 := (1-f_{x_2})\cdots(1-f_{x_m})$. Consequently, we have

$$\psi_1+\psi_2+\ldots+\psi_m \equiv 1.$$

On the other hand, we claim that

$$\psi_i(t)=0 \quad \text{for all } t\notin N(x_i),\ i=1,\ldots,m. \quad (1)$$

Indeed, if $i\geq 2$, then the claim is clear by construction. If $t\notin N(x_1)$, then $t\in V(x_j)$ for some $j=2,\ldots,m$. Hence $f_{x_j}(t)=1$ and then

$$\psi_1(t) = (1-f_{x_j}(t))\prod_{i\neq j}(1-f_{x_i}(t)) = 0.$$

Let us define

$$g := \psi_1\widehat{f_0(x_1)}+\psi_2\widehat{f_0(x_2)}+\ldots+\psi_m\widehat{f_0(x_m)} \in \mathcal{H}. \quad (2)$$

Next, given $x_0\in K$ and by the properties of the Hausdorff metric (see, e.g., [6]), we infer

$$d_H([f_0(x_0)]^{\lambda_j}, [(\psi_1\widehat{f_0(x_1)}+...+\psi_m\widehat{f_0(x_m)})(x_0)]^{\lambda_j}) =$$

$$d_H(\left[\sum_{i=1}^m \psi_i(x_0)f_0(x_0)\right]^{\lambda_j},$$

$$\left[(\psi_1\widehat{f_0(x_1)}+...+\psi_m\widehat{f_0(x_m)})(x_0)\right]^{\lambda_j}) \leq$$

$$\sum_{i=1}^m \psi_i(x_0)d_H([f_0(x_0)]^{\lambda_j},[f_0(x_i)]^{\lambda_j})$$

for $j=1,...,n$.

Let $I=\{1\leq i\leq m : x_0\in N(x_i)\}$ and $J=\{1\leq i\leq m : x_0\notin N(x_i)\}$. Then, for all $i\in I$, we have

$$\psi_i(x_0)d_H([f_0(x_0)]^{\lambda_j},[f_0(x_i)]^{\lambda_j}) \leq \psi_i(x_0)\varepsilon'$$

for $j=1,...,n$ and, for all $i\in J$, (1) yields

$$\psi_i(x_0)d_H([f_0(x_0)]^{\lambda_j},[f_0(x_i)]^{\lambda_j}) = 0$$

for $j=1,...,n$. Hence, we deduce

$$\sum_{i=1}^m \psi_i(x_0)d_H([f_0(x_0)]^{\lambda_j},[f_0(x_i)]^{\lambda_j}) \leq \sum_{i\in I}\psi_i(x_0)\varepsilon' < \varepsilon$$

for $j=1,...,n$. Since x_0 is arbitrary in K, we infer $g\in V(f_0,K,\{\lambda_1,...,\lambda_n\},\epsilon)$.

Remark 3.3. The following example shows that a subspace of $C(X,(\mathbb{E}^1,\tau_l))$ which does not contain all the finite combinations of the form $\psi_1\widehat{u_1}+...+\psi_m\widehat{u_m}$, where $\psi_i\in C(X,[0,1])$ and $u_i\in\mathbb{E}^1$, $i=1,...,m$, may fail to be dense in $C(X,(\mathbb{E}^1,\tau_l))$:

Let $X=[0,1]$ and let $H=span\{x^{n_j} : 0=n_0<n_1<n_2<...\}$ with $\sum_j \frac{1}{n_j}\neq\infty$, which is a subset of $C([0,1],[0,1])$. Let u be the fuzzy number introduced in Example 2.1 and consider the subspace $\mathcal{H}=\{f\hat{u} : f\in H\}$ of $C([0,1],(\mathbb{E}^1,\tau_l))$. Let us suppose that $\mathcal{H}$ is dense in $C([0,1],(\mathbb{E}^1,\tau_l))$. In that case, given any $f_0\in C([0,1],[0,1])$ and $\varepsilon>0$, we could find $f \in H$ such that $d_H([f_0\hat{u}(x)]^{0.5}, [f\hat{u}(x)]^{0.5}) = |f_0(x) - f(x)| max\{|u^+(0.5)|, |u^-(0.5)|\} < \varepsilon$ for all $x \in [0,1]$. Since $u^-(0.5) = 0$ and $u^+(0.5) = 1$, we infer that $|f_0(x) - f(x)| < \varepsilon$ for all $x \in [0,1]$. This would imply that H is dense in $C([0,1],[0,1])$, which is not true by [22, Section 5].

4. Density of fuzzy-number-valued regular neural networks and sums of fuzzy-number-valued ridge functions.

A fuzzy-number-valued four-layer regular neural network (four-layer RFNN) is defined by

$$H(\mathbf{x}) = \sum_{i=1}^{m} u_i \left(\sum_{l=1}^{s} w_{il} \cdot \sigma(\mathbf{x} \cdot \mathbf{a_l} + \theta_l) \right)$$

for each $\mathbf{x} \in \mathbb{R}^n$, where $u_i \in \mathbb{E}^1$, the coeficients w_{ij} and the thresholds θ_j are real numbers, the weights $\mathbf{a_j} \in \mathbb{R}^n$ and $\sigma : \mathbb{R} \to \mathbb{R}$ stands for the activation function in the hidden layer. That is, $H(\mathbf{x})$ is a $\mathbb{R}^n$-based fuzzy-number-valued function. In [5] (see also [20]), the authors proved that, although three-layer RFNNs cannot (unlike the classical real case) approximate the set of all d_∞-continuous $\mathbb{R}$-based fuzzy-number-valued functions, four-layer RFNNs can. They used sigmoidal or bounded continuous nonconstant activation functions to achieve such approximation property of four-layer RFNNs.

For a fixed activation function σ, we denote the set of all possible $\mathbb{R}^n$-based fuzzy-number-valued four-layer RFNNs by $\mathcal{H}(\sigma)$.

We will show, based on the results in the previous section, that $\mathcal{H}(\sigma)$ is dense in $C(\mathbb{R}^n, (\mathbb{E}^1, \tau_l))$ endowed with the compact-open topology provided the activation function σ be a non-polynomial continuous function.

Theorem 4.1. Assume that $\sigma : \mathbb{R} \to \mathbb{R}$ is either a non-polynomial continuous function or a bounded (not necessarily continuous) sigmoidal function. Then $\mathcal{H}(\sigma)$ is dense in $C(\mathbb{R}^n, (\mathbb{E}^1, \tau_l))$.

Proof. Let $f_0 \in C(\mathbb{R}^n, (\mathbb{E}^1, \tau_l))$ and take a neighborhood of f_0,

$$V(f_0, K, \{\lambda_1, ..., \lambda_q\}, \epsilon).$$

By Theorem 3.2, there exist finitely many functions $\psi_i \in C(K,[0,1])$ and $u_i \in \mathbb{E}^1$, $i = 1, ..., m$, such that

$$d_H([f_0(\mathbf{x})]^{\lambda_j}, [(\psi_1\widehat{u_1} + ... + \psi_m\widehat{u_m})(\mathbf{x})]^{\lambda_j}) < \frac{\varepsilon}{2}$$

for all $\mathbf{x} \in K$ and for $j = 1, ..., q$.

On the other hand, by [18] and [16], we know that for each ψ_i, $i = 1, ..., m$, there exist $w_{il}, \theta_{il} \in \mathbb{R}$ and $\mathbf{a_{il}} \in \mathbb{R}^n$ such that

$$\left| \psi_i(\mathbf{x}) - \sum_{l=1}^{s_i} w_{il} \cdot \sigma(\mathbf{x} \cdot \mathbf{a_{il}} + \theta_{il}) \right| < \frac{\varepsilon}{2m \cdot d_\infty(u_i, 0)},$$

for all $\mathbf{x} \in K$. Hence

$$d_H(\left[\left(\sum_{l=1}^{s_i} w_{il} \cdot \sigma(\mathbf{x} \cdot \mathbf{a_{il}} + \theta_{il}) \right) \widehat{u_i}(\mathbf{x}) \right]^{\lambda_j},$$

$$[\phi_i(\mathbf{x}) \cdot \widehat{u_i}(\mathbf{x})]^{\lambda_j}) =$$

$$= \left| \phi_i(\mathbf{x}) - \left(\sum_{l=1}^{s_i} w_{il} \cdot \sigma(\mathbf{x} \cdot \mathbf{a_{il}} + \theta_{il}) \right) \right| d_H([u_i]^{\lambda_j}, 0) \leq$$

$$\leq \left| \phi_i(\mathbf{x}) - \left(\sum_{l=1}^{s_i} w_{il} \cdot \sigma(\mathbf{x} \cdot \mathbf{a_{il}} + \theta_{il}) \right) \right| d_\infty(u_i, 0)$$

$$< \frac{\varepsilon}{2m},$$

for all $\mathbf{x} \in K$, $i = 1, ..., m$ and $j = 1, ..., q$. As a consequence,

$$d_H([f_0(\mathbf{x})]^{\lambda_j},$$

$$\left[\sum_{i=1}^{m} u_i \left(\sum_{l=1}^{s_i} w_{il} \cdot \sigma(\mathbf{x} \cdot \mathbf{a_{il}} + \theta_{il}) \right) (\mathbf{x}) \right]^{\lambda_j}) < \varepsilon$$

for all $\mathbf{x} \in K$ and for $j = 1, ..., q$ and we are done.

Let us recall that *ridge* functions are multivariate functions of the form $g(a_1x_1 + ... + a_nx_n)$ where

$g : \mathbb{R} \to \mathbb{R}$ and $(a_1, ..., a_n) \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ is a fixed direction. For a subset A of $\mathbb{R}^n \setminus \{\mathbf{0}\}$, we can define

$$\mathcal{R}(A) = span\{g(a_1x_1 + ... + a_nx_n) :$$

$$g \in C(\mathbb{R}, \mathbb{R}), (a_1, ..., a_n) \in A\}.$$

We can adapt these functions to the fuzzy setting as follows: Let $\mathcal{R}(\mathbb{R}^n, \mathbb{E}^1, A)$ stand for the sums of fuzzy-number-valued ridge functions of the form

$$H(\mathbf{x}) = \sum_{i=1}^{m} u_i \left(\sum_{l=1}^{s_i} g_{il}(\mathbf{a_{il}} \cdot \mathbf{x}) \right)$$

for each $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n$ and $\mathbf{a_{il}} = (a^1_{il}, ..., a^n_{il}) \in A$, where $u_i \in \mathbb{E}^1$.

Theorem 4.2. Let A be a subset of $\mathbb{R}^n \setminus \{\mathbf{0}\}$. Then $\mathcal{R}(A)$ and $\mathcal{R}(\mathbb{R}^n, \mathbb{E}^1, A)$ are dense in $C(\mathbb{R}^n, \mathbb{R})$ and in $C(\mathbb{R}^n, (\mathbb{E}^1, \tau_l))$, respectively, if and only if $\mathcal{R}(A)$ contains all polynomials.

Proof. Assume $\mathcal{R}(A)$ contains all polynomials. As in the proof of Theorem 4.1, given $f_0 \in C(\mathbb{R}^n, (\mathbb{E}^1, \tau_l))$ and a neighborhood of f_0, $V(f_0, K, \{\lambda_1, ..., \lambda_q\}, \epsilon)$, there exist finitely many functions $\psi_i \in C(K, [0, 1])$ and $u_i \in \mathbb{E}^1$, $i = 1, ..., m$, such that

$$d_H([f_0(\mathbf{x})]^{\lambda_j}, [(\psi_1\widehat{u_1} + ... + \psi_m\widehat{u_m})(\mathbf{x})]^{\lambda_j}) < \frac{\varepsilon}{2}$$

for all $\mathbf{x} \in K$ and for $j = 1, ..., q$. By [19, Theorem 2.1 and Remark 2.2], we can find, for each ψ_i, $i = 1, ..., m$, $\mathbf{a_{il}} \in A$ and $g_{il} \in C(\mathbb{R}, \mathbb{R})$ such that

$$\left| \psi_i(\mathbf{x}) - \sum_{l=1}^{s_i} g_{il}(\mathbf{a_{il}} \cdot \mathbf{x}) \right| < \frac{\varepsilon}{2m \cdot d_\infty(u_i, 0)},$$

for all $\mathbf{x} \in K$. Hence

$$d_H \left(\left[\left(\sum_{l=1}^{s_i} g_{il}(\mathbf{a_{il}} \cdot \mathbf{x}) \right) \widehat{u_i}(\mathbf{x}) \right]^{\lambda_j}, \ [\phi_i(\mathbf{x}) \cdot \widehat{u_i}(\mathbf{x})]^{\lambda_j} \right) =$$

$$= \left| \phi_i(\mathbf{x}) - \left(\sum_{l=1}^{s_i} g_{il}(\mathbf{a_{il}} \cdot \mathbf{x}) \right) \right| d_H([u_i]^{\lambda_j}, 0) \leq$$

$$\leq \left| \phi_i(\mathbf{x}) - \left(\sum_{l=1}^{s_i} g_{il}(\mathbf{a_{il}} \cdot \mathbf{x}) \right) \right| d_\infty(u_i, 0) < \frac{\varepsilon}{2m},$$

for all $\mathbf{x} \in K$, $i = 1, ..., m$ and $j = 1, ..., q$. As a consequence,

$$d_H \left([f_0(\mathbf{x})]^{\lambda_j}, \left[\sum_{i=1}^{m} u_i \left(\sum_{l=1}^{s_i} g_{il}(\mathbf{a_{il}} \cdot \mathbf{x}) \right) (\mathbf{x}) \right]^{\lambda_j} \right) < \varepsilon$$

for all $\mathbf{x} \in K$ and for $j = 1, ..., q$.

On the other hand, if we assume that $\mathcal{R}(A)$ is dense in $C(\mathbb{R}^n, \mathbb{R})$, then $\mathcal{R}(A)$ contains all polynomials by [19, Theorem 2.1 and Remark 2.2]

Remark 4.3. According to [19, Theorem 2.1 and Remark 2.2], the expression "$\mathcal{R}(A)$ contains all polynomials" can be replaced by "no nonzero homogeneous polynomial vanishes on A".

5. Example

In this section, we provide a numerical example to illustrate the method which we have used in our main result (Theorem 3.2).

Let $X = [0, 1[$ and let us define a level-continuous function $f_0 : [0, 1[\longrightarrow \mathbb{E}^1$ as follows:

$$f_0(0)(x) = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1 \end{cases}$$

and $f_0(t)(x) = x^{\frac{1}{t}}$ for $0 \leq x \leq 1$. Hence, we deduce that

$$[f_0(0)]^\lambda = \begin{cases} \{1\} & 0 < \lambda \leq 1 \\ [0,1] & \lambda = 0 \end{cases}$$

and for $t \in]0, 1[$,

$$[f_0(t)]^\lambda = \begin{cases} [\lambda^t, 1] & 0 < \lambda \leq 1 \\ [0,1] & \lambda = 0. \end{cases}$$

Fix $\epsilon = 0.1$ and $\lambda_1 = 0.5$. Hence, if take $t \in]0, 1[$, then

$$N(t) = \{s \in X : |0.5^t - 0.5^s| < 0.1\}.$$

If we denote $t_i = 0.1i - 0.05$ for $i = 1, 2, ..., 10$, then it is apparent that $[0, 1[\subset \cup_{i=1}^{10} N(t_i)$.

Let us define the following trapezoidal functions for $i = 2, ..., 10$:

$$f_i(t) = \begin{cases} 20t - (2i-3) & 0.1i - 0.15 < t < 0.1i - 0.1 \\ 1 & 0.1i - 0.1 \leq t \leq 0.1i \\ -20t + (2i+1) & 0.1i < t < 0.1i + 0.05 \\ 0 & otherwise \end{cases}$$

From these functions we can next construct the following ones, which turn out to be trapezoidal too:
$\psi_2 := f_{x_2}$,
$\psi_3 := (1 - f_{x_2})f_{x_3}$,
$\vdots$
$\psi_{10} := (1 - f_{x_2})(1 - f_{x_3}) \cdots (1 - f_{x_9})f_{x_{10}}$. Namely,

$$\psi_3(t) = \begin{cases} 20t - 4 & 0.2 < t < 0.25 \\ 1 & 0.25 \leq t \leq 0.3 \\ -20t + 7 & 0.3 < t < 0.35 \\ 0 & otherwise \end{cases}$$

and a routine manipulation shows

$$\psi_2(t) + \psi_3(t) = \begin{cases} 20t - 1 & 0.05 < t < 0.1 \\ 1 & 0.1 \leq t \leq 0.3 \\ -20t + 7 & 0.3 < t < 0.35 \\ 0 & otherwise \end{cases}$$

Similarly, we would obtain

$$\psi_2(t) + ... + \psi_{10}(t) = \begin{cases} 20t - 1 & 0.05 < t < 0.1 \\ 1 & 0.1 \leq t < 1 \\ 0 & otherwise \end{cases}$$

Finally, let us define $\psi_1(t) := 1 - (\psi_2(t) + ... + \psi_{10}(t))$, which yields

$$\psi_1 = \begin{cases} 1 & 0 \leq t < 0.05 \\ -20t + 2 & 0.05 \leq t \leq 0.1 \\ 0 & otherwise \end{cases}$$

It is also apparent that the support of each ψ_i lies in $N(t_i)$, $i = 1, 2, ..., 10$. Finally, we define

$$g := \psi_1 \widehat{f_0(0.05)} + \psi_2 \widehat{f_0(0.15)} + \ldots + \psi_{10} \widehat{f_0(0.95)}. \quad (3)$$

Fix $t_0 \in [0, 1[$. Then, since $\sum_{i=1}^{10} \psi_i(t_0) = 1$, we infer

$$d_H([f_0(t_0)]^{0.5}, [(\psi_1 \widehat{f_0(0.05)} + ... + \psi_{10} \widehat{f_0(0.95)})(t_0)]^{0.5}) =$$

$$d_H(\left[\sum_{i=1}^{10} \psi_i(t_0) f_0(t_0)\right]^{0.5},$$

$$\left[(\psi_1 \widehat{f_0(0.05)} + ... + \psi_{10} \widehat{f_0(0.95)})(t_0)\right]^{0.5}) \leq$$

$$\psi_1(t_0) d_H([f_0(t_0)]^{0.5}, [f_0(0.05)]^{0.5}) + ...$$

$$+ \psi_{10}(t_0) d_H([f_0(t_0)]^{0.5}, [f_0(0.95)]^{0.5}) =$$

$$= \psi_1(t_0)|0.5^{t_0} - 0.5^{0.05}| + ...$$

$$+ \psi_{10}(t_0)|0.5^{t_0} - 0.5^{0.95}|.$$

Since t_0 belongs to the support of, at most, two functions ψ_i and we know that such supports lie in their respective $N(t_i)$, we infer that

$$d_H([f_0(t_0)]^{0.5},$$

$$[(\psi_1 \widehat{f_0(0.05)} + ... + \psi_{10} \widehat{f_0(0.95)})(t_0)]^{0.5}) \leq 0.1.$$

As a consequence, since t_0 is arbitrary, we deduce $g \in V(f_0, \lambda_1, \epsilon)$ with $\epsilon = 0.1$ and $\lambda_1 = 0.5$.

6. Application

The study on the theory of fuzzy optimization has been active since the concept of fuzzy decision was proposed by Bellman and Zadeh ([4]) in 1970. This is a useful methodology, since it allows us to represent the underlying uncertainty of the optimization problem. As indicated by Dubois and Prade ([7]), constrained fuzzy optimization (also known as fuzzy mathematical programming) refers to the search for extrema of a fuzzy-valued utility (or objective) function define on a bounded domain and, among a large number of forms, could be described as follows: consider a (level-continuous) fuzzy-valued objective function $f : \mathbb{R} \longrightarrow \mathbb{E}^1$ subject to a constraint set $A \subset \mathbb{R}$. Then

$$\begin{cases} \text{Maximize (or minimize) } f(x) \\ \\ \text{subject to } x \in A \end{cases}$$

It is known (see [9, Theorem 5.1]) that if $A = [a, b]$, then there exists the supremum and the infimum of $f(x)$ on A. Indeed, if u denotes such maximum (similarly for the infimum), then, for each $\lambda \in [0, 1]$,

$$[u]^\lambda = [\sup_{t\in[a,b]} [f(t)]^-(\lambda), \sup_{t\in[a,b]} [f(t)]^+(\lambda)].$$

However, such $f(x)$ might not attain either its supremum or its infimum (see [9, Remark 5.3]). Consequently, our unique option is to approximate the supremum (resp. the infimum) and our Theorem 3.2 can help us by realizing the objective function based on a finite subset of fuzzy numbers rather than the whole $\mathbb{E}^1$, which will reduce the problem to a classical (crisp) real-valued optimization problem. In order to illustrate this technique, we can consider the problem of maximizing the function provided in the previous section subject to $x \in [0, 1]$. Then we can approximate the solution of such problem as follows: fix, for instance, $\varepsilon = 0.1$ and $\lambda_1 = 0.5$, and let $\psi_i(t)$, $i = 1, ..., 10$ be as in the previous section. Then we can define the function

$$g(t) = \psi_1(t)\widehat{f_0(0.05)} + ... + \psi_{10}(t)\widehat{f_0(0.95)}$$

for $t \in [0, 1]$ and, consequently,

$$[g(t)]^{0.5} = \psi_1(t)[0.5^{0.05}, 1]+...+\psi_{10}(t)[0.5^{0.95}, 1] =$$

$$= [\psi_1(t)0.5^{0.05} + ... + \psi_{10}(t)0.5^{0.95}, 1].$$

If we denote $v = sup\{g(t) : t \in [0, 1]\}$, then we get

$$[v]^{0.5} = [\sup_{t\in[0,1]} (\psi_1(t)0.5^{0.05}+...+\psi_{10}(t)0.5^{0.95}), 1] =$$

$$= [0.5^{0.95}, 1]$$

On the other hand, if $u = sup\{f(t) : t \in [0, 1]\}$, then it is apparent that $[u]^{0.5} = [0.5, 1]$, which yields

$$d_H([u]^{0.5}, [v]^{0.5}) = 0.017632461 < \varepsilon = 0.1.$$

Similarly we can proceed with any $\lambda = [0, 1]$.

7. Conclusion

In this paper we have studied density problems in the space of level continuous fuzzy-number-valued functions defined on a locally compact Hausdorff space endowed with the compact-open topology, whose use does not seem to have made its way in the fuzzy literature so far. We provide a numerical example to illustrate the techniques we have based on. As a corollary of the above results we find that many fuzzy-number-valued neural networks (with two hidden layers) and sums of fuzzy-number-valued ridge functions are dense in $C(\mathbb{R}^n, \mathbb{E}^1)$. We hope that our techniques, together with the introduction of ridge functions, will be helpful for obtaining stronger results and further applications in environments related to fuzzy approximation.

One of such fields is decision theory, which comprises a broad diversity of approaches to modeling behavior of a human decision maker under various information frameworks in management science, economics and other areas. Among the multiple approaches to decision making in a fuzzy context, the use of fuzzy-valued utility (or objective) functions as a quantitative representation of preferences of a decision maker was proposed in 1970 by Bellman and Zadeh in their seminal paper, [4], and was concretized in [23] by showing that their approach reduces to a fuzzy optimization problem of fuzzy-valued utility functions based on λ-level sets. In general, an optimization problem deals with two elements: a goal or utility function and a set of feasible domains and, in a fuzzy context, it consists of finding an x "belonging" to the domain X of a fuzzy-valued function $f : X \longrightarrow \mathbb{E}^1$ such that $f(x)$ can reach a possible "extremum" in a fuzzy sense. However, how to interpret the terms "belonging" and "extremum" in this fuzzy environment is not apparent since $\mathbb{E}^1$ is not a linearly ordered space and that is why we can find a plethora of approaches in the literature (see [24]). For example, in [1], the authors use fuzzy-number-valued utility functions which represent linguistic preferences based on the Hausdorff distance of their images.

In fuzzy optimization it is desirable that all fuzzy solutions under consideration be attainable, but very often one may find that maximum covering problems are computationally complex and not easy to solve. In these cases the decision maker must usually accept approximate solutions instead of optimum ones. Thus, in [5] (see also [25]), Buckley and Hayashi were the first authors to introduce a technique to solve fuzzy optimization problems approximately. Their technique was based on maximizing the centroids of the fuzzy number in the range of the utility function.

As shown in the previous section, our results can contribute to find approximate solutions of constrained fuzzy optimization problems by realizing utility fuzzy-valued functions based on a finite subset of fuzzy numbers rather than the whole $\mathbb{E}^1$, which boils down the fuzzy problem to a classical (crisp) real-valued optimization problem.

Acknowledgement. The authors wish to thank the referees and the Associate Editor, Prof. X. Ma, for their remarks which contributed to improve the quality of this manuscript.

References

[1] R. Aliev, W. Pedrycz, B. Fazlollahi, O.H. Huseynov, A.V. Alizadeh and B.G. Guirimov, *Fuzzy logic-based generalized decision theory with imperfect information*, Information Sciences, **189** (2012), 243–246.

[2] A. Aygunoglu, V. Cetkin and H. Aygun, *On embedding problem of fuzzy number valued continuous functions*, J. Intell. Fuzzy Systems, **25** (2013), 243–246.

[3] B. Bede and S. Gal, *Generalizations of the differentiability of fuzzy-number-valued functions with applications to fuzzy differential equations*, Fuzzy Sets and Systems, **151** (2005), 581–599.

[4] R. E. Bellman and L. A. Zadeh, *Decision making in a fuzzy environment*, Management Science, **17** (1970), 141–164.

[5] J. Buckley and Y. Hayashi, *Can fuzzy neural nets approximate continuous fuzzy functions?* Fuzzy Sets and Systems, **61** (1994), 43–51.

[6] P. Diamond and P. Kloeden, *Metric Spaces of Fuzzy Sets: Theory and Applications*, World Scientific, Singapore, (1994).

[7] D. Dubois and H. Prade, *Decision making under fuzzy constraints and fuzzy criteria mathematical programming vs rule-based system approach*, in M. Delgado, et al. Ed. Fuzzy Optimization-Recent Advances, Physica-Verlag, Heidelberg, (1994), 21–32.

[8] J-X. Fang and H. Huang, *Some properties of the level convergence topology on fuzzy number space* $\mathbb{E}^n$, Fuzzy Sets and Systems **140** (2003), 509–517.

[9] J-X. Fang and H. Huang, *On the level convergence of a sequence of fuzzy numbers*, Fuzzy Sets and Systems **147** (2004), 417–435.

[10] J.J. Font, A. Miralles and M. Sanchis, *On the Fuzzy Number Space with the Level Convergence Topology*, J. Funct. Spaces Appl., Volume 2012, Article ID 326417, 11 pages.

[11] J.J. Font, D. Sanchis, M. Sanchis, *A version of the Stone-Weierstrass theorem in fuzzy analysis*. J. Nonlinear Sci. Appl. **10** (2017), 4275–4283.

[12] R. Goetschel and W. Voxman, *Elementary fuzzy calculus*, Fuzzy Sets and Systems, **18** (1986), 31–42.

[13] S. Hai, Z. Gong and H. Li, *Generalized differentiability for n-dimensional fuzzy-number-valued functions and fuzzy optimization*, Information Sciences, **374** (2016), 151–163.

[14] H. Huang and C. Wu, *Approximation of level fuzzy-valued functions by multilayer regular fuzzy neural networks*, Math. Comp. Model., **49** (2009), 1311–1318.

[15] H. Huang and C. Wu, *Approximation of fuzzy-valued functions by regular fuzzy neural networks and the accuracy analysis*, Soft Comput., **18** (2014), 2525–2540.

[16] L.K. Jones, *Constructive Approximations for Neural Networks by Sigmoidal Functions*, Proc. IEEE, **78** (1990), 1586–1589.

[17] O. Kaleva, S. Seikkla, *On fuzzy metric space*, Fuzzy Sets and Systems, **12** (1984), 215-Ű229.

[18] M. Leshno, V.Y. Lin, A. Pinkus and S. Schocken, *Multilayer feedforward networks with a nonpolynomial activation function can approximate any function*, Neural Networks, **6** (1993), 861–867.

[19] V.Y. Lin and A. Pinkus, *Fundamentality of ridge functions*, J. Approx. Th., **75** (1993), 295–311.

[20] P.Y. Liu, *Universal approximations of continuous fuzzy-valued function by multi-layer regular fuzzy neural networks*, Fuzzy Sets and Systems, **119** (2001), 313–320.

[21] B.F. Logan and L.A. Shepp, *Optimal reconstruction of a function from its projections*, Duke Math. J., **42** (1975), 645–659.

[22] A. Pinkus, *Density in approximation theory*, Surv. Approx. Theory, **1** (2005), 1–45.

[23] H. Tanaka, T. Okuda and K. Asai, *On fuzzy-mathematical programming*, J. Cybern., **3** (1974), 37–46.

[24] J.F. Tang, D.W. Wang, R. Fung and K.L. Yung, *Understanding of fuzzy optimization: Theories and methods*, J. Syst. Sci. Complexity, **17** (2004), 117–136.

[25] D. Tikk, L.T. Kóczy and T.D. Gedeon, *A survey on universal approximation and its limits in soft computing techniques*, International Journal of Approximate Reasoning, **33** (2003), 185–202.

[26] C.-x Wu and G.-x Wang, Convergence of sequence of fuzzy numbers and fixed point theorems for increasing fuzzy mappings and applications, Fuzzy Sets and Systems **130** (2002) 383–390.

[27] C. Wu and Z. Gong, *On Henstock integral of fuzzy-number-valued functions (I)*, Fuzzy Sets and Systems, **120** (2001), 523–532.