CONSTRUCTIVE APPROXIMATION OF CONTINUOUS INTERVAL-VALUED FUNCTIONS

JUAN J. FONT AND SERGIO MACARIO

Abstract. In this paper we present a Stone-Weierstrass type result in the context of continuous interval-valued functions defined on a compact Hausdorff space. We also provide a Jackson type approximation result involving the modulus of continuity of interval-valued functions based on the gH-difference of intervals. Finally we show how an interval-valued continuous function can be approximated using interval neural networks. All the proofs are constructive.

1. Introduction

When modelling a complex process, some uncertainty in the input and output variables is expected. In such situations, input and output data should mix numbers and intervals. So, often in these modelling processes, it is needed to deal with functions giving an interval as an output rather than real numbers. Interval Analysis is a relatively new branch of research of Numerical Analysis which studies how to handle such interval uncertainty which appears in a lot of computer-mathematical models of certain real-world phenomena; in particular, in control theory, linear programming, optimization problems, etc. The first principles of interval arithmetic were set independently, and almost simultaneously, in the fifties by Paul S. Dwyer ([7]) and Ramon E. Moore ([17], [18]) in the Unites States, Mieczyslaw Warmus ([25]) in Poland, and Teruo Sunaga ([24]) in Japan, although Interval Analysis is often said to have begun with Moore's book [19].

Interval-valued functions, that is, functions defined on a topological space taking values in the space of closed intervals, should play a central role in Interval Analysis, just like real-valued functions do in the Classical Analysis. However some difficulties arise when dealing with these interval-valued functions, mainly because the space they form is not a linear space; indeed it is not a group with respect to addition.

Stimulated by the interaction between Interval Analysis and Optimization Theory, interval-valued functions have received considerable attention

Date: July 28, 2021.

2010 *Mathematics Subject Classification.* 41A30, 65G40.

Key words and phrases. Stone-Weierstrass theorem, interval-valued continuous functions, Jackson-type theorem, interval neural networks.

Research supported by Spanish AEI Project PID2019-106529GB-I00 / AEI / 10.13039/501100011033 and grant UJI-B2019-08.

recently motivated mainly by the necessity of formulating a formal framework for a differential calculus in this context. Namely, an appropriate metric space and a well-behaved subtraction are needed for such development. Thus, for the theoretical framework of calculus of interval-valued functions and among others, Hukuhara ([12]) introduced a concept of difference of two intervals (H-difference) which was used to define the H-derivative of an interval-valued function. This concept, however, turned out to be a very restrictive concept. So, in 2008, Stefanini proposed an improved version of the H-difference (the gH-difference) which seems to be a very useful tool for dealing with interval-valued functions.

The literature on other aspects of the theory of interval-valued functions is reduced compared to the plethora of results dealing with the differential calculus mentioned in the previous paragraph. One of such aspects is the study of the approximation of continuous interval-valued functions.

In this paper, by combining certain techniques from [8], [14]) and [21], we provide some sufficient conditions on a subset of the space of continuous interval-valued functions in order that it be dense, which is to say, a Stone-Weierstrass type result for interval-valued continuous functions which doesn't seem to have made its way into the literature yet. The proofs are constructive and use only straightforward concepts.

In this context, we also provide a Jackson type approximation result involving the modulus of continuity of interval-valued functions based on the gH-difference mentioned above.

Finally, based on the results of the previous sections and taking advantage of an striking result by Guliyev and Ismailov ([10]), we show how an interval-valued continuous function can be approximated using interval neural networks.

2. Preliminaries

Following [6], let $\mathcal{K}_{\mathcal{C}}$ denote the set of all finite closed intervals of the real line $\mathbb{R}$.

$$\mathcal{K}_{\mathcal{C}} = \{[a,b] \ : a, b \in \mathbb{R}, \ a \leq b\}.$$

where $[a,a]$ denotes the singleton $\{a\}$.

In $\mathcal{K}_{\mathcal{C}}$ we shall consider two operations, (Minkowski) addition and scalar multiplication, defined by

$$I + J = \{a + b \ : \ a \in I, \ b \in J\}$$

and, for every $\lambda \in \mathbb{R}$,

$$\lambda I = \{\lambda a \ : a \in I\}.$$

Hereafter 0 will denote the singleton $[0,0] = \{0\}$. and it is the neutral element for the Minkowski addition. Unless J is a singleton , $J + (-J) \neq 0$. So, in general, there is no inverse for the sum operation and the structure of $\mathcal{K}_{\mathcal{C}}$ is that of a cone rather than a linear space.

We consider now the metric space $(\mathcal{K}_{\mathcal{C}}, d_H)$, where d_H is the Pompeiu-Hausdorff metric defined as

$$d_H(I,J) = \max\left\{\max_{a\in I} d(a,J), \max_{b\in J} d(I,b)\right\}$$

where

$$d(a,J) = \min_{b\in J} |a-b| \quad \text{and} \quad d(I,b) = \min_{a\in I} |a-b|.$$

In particular, if $I = [a,b]$ and $J = [c,d]$, then

$$d_H(I,J) = \max\{|a-c|, |b-d|\}.$$

It is well known that $(\mathcal{K}_{\mathcal{C}}, d_H)$ is a complete metric space (see [1]).

Proposition 2.1. *The metric d_H satisfies the following properties:*

(1) $d_H(\sum_{i=1}^m A_i, \sum_{i=1}^m B_i) \le \sum_{i=1}^m d_H(A_i, B_i)$ *where* $A_i, B_i \in \mathcal{K}_{\mathcal{C}}$ *for* $i = 1, ..., m$.
(2) $d_H(\alpha A, \alpha B) = \alpha d_H(A,B)$ *where* $A, B \in \mathcal{K}_{\mathcal{C}}$ *and* $\alpha > 0$.
(3) $d_H(\alpha A, \beta A) = \mid \alpha - \beta \mid d_H(A, 0)$, *where* $A \in \mathcal{K}_{\mathcal{C}}$, $\alpha,\ \beta \ge 0$.
(4) $d_H(\alpha A, \beta B) \le \mid \alpha - \beta \mid d_H(A,0) + \beta d_H(A,B)$, *where* $A, B \in \mathcal{K}_{\mathcal{C}}$, $\alpha,\ \beta \ge 0$.
(5) $d_H(A + C, B + C) = d_H(A,B)$, *where* $A, B, C \in \mathcal{K}_{\mathcal{C}}$.

Proof. The proofs of (1) and (2) can be found, for example, in [6]. In order to prove (3), let us assume that $\beta < \alpha$ and rewrite αA and βA as $\alpha A = (\beta + (\alpha - \beta))A = \beta A + (\alpha - \beta)A$ and $\beta A = \beta A + (\alpha - \beta)0$. By (1), we know that

$$d_H(\alpha A, \beta A) \le d_H(\beta A, \beta A) + d_H((\alpha - \beta)A, (\alpha - \beta)0) = \mid \alpha - \beta \mid d_H(A, 0).$$

Consequently, by (2) and (3),

$$d_H(\alpha A, \beta B) \le d_H(\alpha A, \beta A) + d_H(\beta A, \beta B) \le \mid \alpha - \beta \mid d_H(A,0) + \beta d_H(A,B).$$

For (5), take $A = [a^-, a^+]$, $B = [b^-, b^+]$ and $C = [c^-, c^+]$. Then

$$\begin{aligned} d_H(A+C, B+C) &= \max\{|a^- + c^- - (b^- + c^-)|,\ |a^+ + c^+ - (b^+ + c^+|\} \\ &= \max\{|a^- - b^-|,\ |a^+ - b^+|\} = d_H(A,B) \end{aligned}$$

□

In the sequel, let K be a compact Hausdorff space and let $C(K, \mathcal{K}_{\mathcal{C}})$ denote the space of continuous functions from K to the metric space $(\mathcal{K}_{\mathcal{C}}, d_H)$, that is, the space of continuous interval-valued functions defined on K. We shall consider $C(K, \mathcal{K}_{\mathcal{C}})$ endowed with the supremum metric:

$$D_\infty(f,g) = \sup_{t\in K} d_H(f(t), g(t)),$$

which induces, as usual, the uniform convergence topology on $C(K, \mathcal{K}_{\mathcal{C}})$.

Proposition 2.2. *Let $f \in C(K, \mathcal{K}_{\mathcal{C}})$ and $\varphi \in C(K, \mathbb{R}^+)$. Then the mapping from K to $\mathcal{K}_{\mathcal{C}}$ defined by $k \mapsto \varphi(k)f(k)$, belongs to $C(K, \mathcal{K}_{\mathcal{C}})$ as well.*

Proof. Since $\varphi(k) > 0$ for all $k \in K$, then the pointwise product $\varphi(k)f(k)$ is well defined. Let us see that the mapping is continuous. Choose $k_1, k_2 \in K$. By Proposition 2.1 (4), we know that

$$\begin{aligned} d_H(\varphi(k_1)f(k_1), \varphi(k_2)f(k_2)) \leq\, & |\varphi(k_1) - \varphi(k_2)| d_H(f(k_1), 0) \\ & + \varphi(k_2) d_H(f(k_1), f(k_2)). \end{aligned}$$

The proof follows since both φ and f are continuous and bearing in mind that every continuous function from a compact space to a metric space is bounded. $\square$

3. A version of the Stone-Weierstrass theorem in Interval Valued Analysis.

The Stone-Weierstrass Theorem is a main tool in the General Approximation Theory. Our version states conditions for a subset H of continuous interval-valued functions from a compact space K being dense in the whole space of continuous interval-valued functions, endowed with the supremum metric. That is, to find, for every continuous interval-valued function f, another one close enough to f and belonging to this smaller subset of continuous interval-valued functions. Following [8], we introduce a useful tool to get our main theorem (Theorem 3.8).

Definition 3.1. Let H be a nonempty subset of $C(K, \mathcal{K}_{\mathcal{C}})$. We define

$$Conv(H) = \{\varphi \in C(K, [0,1]) : \varphi f + (1 - \varphi)g \in H \ \ for \ \ all \ \ f, g \in H\}.$$

This set, $Conv(H)$, have some nice properties that will be helpful in order to prove our results.

Proposition 3.2. *Let H be a nonempty subset of $C(K, \mathcal{K}_{\mathcal{C}})$. Then we have:*

1. *$\phi \in Conv(H)$ implies that $1 - \phi \in Conv(H)$.*
2. *If $\phi, \varphi \in Conv(H)$, then $\phi \cdot \varphi \in Conv(H)$.*
3. *If ϕ belongs to the uniform closure of $Conv(H)$, then so does $1 - \phi$.*
4. *If ϕ, φ belong to the uniform closure of $Conv(H)$, then so does $\phi \cdot \varphi$.*

Proof. (1) is clear. To see (2), let us assume that $\phi, \varphi \in Conv(H)$. To get that $\phi \cdot \varphi \in Conv(H)$, we use the identity

$$1 - \phi \cdot \varphi = (1 - \phi) + \phi(1 - \varphi)$$

which implies, for every pair $f, g \in H$, that

$$(\phi \cdot \varphi)f + (1 - \phi \cdot \varphi)g = \phi[\varphi f + (1 - \varphi)g] + (1 - \phi)g \in H.$$

To show (3), let us suppose that there exists a sequence $\{\phi_n\} \subset Conv(H)$ uniformly converging to $\phi \in C(K, [0,1])$. Hence, $\{1 - \phi_n\}$, which is contained in $Conv(H)$ by (1), converges uniformly to $1 - \phi$.

Finally, (4) can be proved in the same way.

$\square$

We still need to state another property of $Conv(H)$ whose proof follows from two technical lemmas that can be found in [14].

Lemma 3.3. [14, Lemma 2] *Let $0 \leq a < b \leq 1$ and $0 < \delta < \frac{1}{2}$. There exists a polynomial $p(x) = (1 - x^m)^n$ such that*

(1) $p(x) > 1 - \delta$ *for all* $0 \leq x \leq a$,
(2) $p(x) < \delta$ *for all* $b \leq x \leq 1$.

Lemma 3.4. [14, Theorem 1] *For every $0 < \varepsilon < \frac{1}{32}$, there exists a polynomial in two variables $q(x, y) = (1 - p_1(x))p_2(x)(1 - p_3(y))p_4(y)$, with $p_k(z) = (1 - z^{m_k})^{n_k}$, m_k, $n_k \in \mathbb{N}$, $k = 1, 2, 3, 4$; such that*

$$|x \wedge y - q(x, y)| < \varepsilon, \qquad \textit{for every } x, y \in [0, 1]$$

where $x \wedge y$ stands for $min\{x, y\}$.

Proposition 3.5. *Let $H \subseteq C(K, \mathcal{K}_\mathcal{C})$. Given two elements of $Conv(H)$, its maximum lies in the uniform closure of $Conv(H)$.*

Proof. Let ϕ and ψ be two elements in $Conv(H)$. Let $\phi \vee \psi$ stand for the maximum of ϕ and ψ. By property (3) of Proposition 3.2, since

$$\phi \vee \psi = 1 - ((1 - \phi) \wedge (1 - \psi)),$$

it suffices to prove that $(1 - \phi) \wedge (1 - \psi)$ belongs to the uniform closure of $Conv(H)$.

Take $0 < \varepsilon < 1/32$ and the corresponding polynomial $q(x, y)$ given by Lemma 3.4. Then,

$$|(1 - \phi(s) \wedge (1 - \psi(s) - q(1 - \phi(s), 1 - \psi(s))| < \varepsilon$$

for all $s \in K$. By (1) and (2) of Proposition 3.2 and the given form of $q(x, y)$ we can claim that $\varphi := q(1 - \phi, 1 - \psi)$ belongs to $Conv(H)$. So we have just shown that there exists $\varphi \in Conv(H)$ with

$$|(1 - \phi(s)) \wedge (1 - \psi(s)) - \varphi(s)| < \varepsilon$$

for all $s \in K$ and we get the conclusion.

□

It is convenient to remark that not only the maximum of two elements of $Conv(H)$ belongs to the uniform closure of $Conv(H)$, but also its minimum.

Definition 3.6. Let H be a subset $C(K, [0, 1])$. It is said that H separates the points of K if given $s, t \in K$, there exists $\phi \in H$ such that $\phi(s) \neq \phi(t)$.

The next lemma provides the selection of the elements $\phi \in Conv(H)$ needed for the construction of the approximation function g in our main result. The proof is similar to the one from [8] but we include it here for the sake of completeness.

Lemma 3.7. *Let H be a subset of $C(K, \mathcal{K}_\mathcal{C})$ such that $Conv(H)$ separates the points of K. Given $x_0 \in K$ and an open neighborhood V of x_0, there exists a neighborhood U of x_0, with $U \subseteq V$, such that, for all $0 < \delta < \frac{1}{2}$, there is $\phi \in Conv(H)$ satisfying*

(1) $\phi(t) > 1-\delta$, *for all* $t \in U$;
(2) $\phi(t) < \delta$, *for all* $t \notin V$.

Proof. Let $W = K \setminus V$. Since $Conv(H)$ separates the points of K, we can assume, with no loss of generality, that for each $t \in W$, there is a $\varphi_t \in Conv(H)$ such that $\varphi_t(t) < \varphi_t(x_0)$.

Pick two real numbers a_t and b_t such that $\varphi_t(t) < a_t < b_t < \varphi_t(x_0)$. Taking $\delta = \frac{1}{4}$ in Lemma 3.3, we can find a polynomial $p_t(x) = (1-x^m)^n$ such that $p_t(x) < \frac{1}{4}$ for $b_t \leq x \leq 1$, and $p_t(x) > \frac{3}{4}$ for $0 \leq x \leq a_t$. Hence, $p_t(\varphi_t(x_0)) < \frac{1}{4}$ and $p_t(\varphi_t(t)) > \frac{3}{4}$.

Then, for every $t \in W$, we can define

$$U(t) := \{s \in K : p_t(\varphi_t(s)) > \frac{3}{4}\},$$

which is an open neighborhood of t. Since W is compact, there exist $t_1, ..., t_m \in W$ such that $W \subset U(t_1) \cup U(t_2) \cup ... \cup U(t_m)$. For each $i = 1, ..., m$ and all $s \in K$ we can define

$$\varphi_i(s) = p_{t_i}(\varphi_{t_i}(s)).$$

We have $p_{t_i}(\varphi_{t_i}(s)) = (1-[\varphi_{t_i}(s)]^m)^n$ and since $\varphi_{t_i}(s) \in Conv(H)$, we infer, by Proposition 3.2, that so is $\varphi_i = p_{t_i}(\varphi_{t_i})$, for all $i = 1, ..., m$.

Let us define $\psi(s) = \varphi_1(s) \vee ... \vee \varphi_m(s)$, $s \in K$ and, by Proposition 3.5, we know that ψ lies in the uniform closure of $Conv(H)$. We remark that $\psi(x_0) < \frac{1}{4}$ and $\psi(t) > \frac{3}{4}$, for all $t \in W$ due to the properties of the polynomials $p_t(x)$. Now, let us define

$$U = \{s \in K ; \psi(s) < \frac{1}{4}\}.$$

Clearly, U is an open neighborhood of x_0 in K. We claim that U is contained in V. Indeed, if $s \in U$ and $s \notin V$, then $s \in W$ and, consequently, $\psi(s) > \frac{3}{4}$ which means it cannot be in U.

Take $0 < \delta < \frac{1}{2}$ and let p be the polynomial defined by Lemma 3.3, applied to $a = \frac{1}{4}$, $b = \frac{3}{4}$ and $\delta/2$. Define $\mu(s) = p(\psi(s))$, for $s \in K$. By Proposition 3.2, (3) and (4), the function μ also belongs to the uniform closure of $Conv(H)$.

If $s \in U$, then $\mu(s) > 1 - \delta/2$ by construction. If $s \notin V$, then $s \in W$ and $\psi(s) > \frac{3}{4}$ gives $\mu(s) < \delta/2$.

Since μ belongs to the uniform closure of $Conv(H)$, there exists $\phi \in Conv(H)$ such that $\| \phi - \eta \|_\infty = \sup_{s \in K} | \phi(s) - \mu(s) | < \delta/2$. Consequently, ϕ satisfies the desired properties. $\square$

We can now state and prove a version of the Stone-Weierstrass theorem for continuous interval-valued functions, gathering the information obtained above.

For every $x \in K$ let $\overline{f_x} \in C(K, \mathcal{K}_{\mathcal{C}})$ be the constant function which takes the constant value $f(x)$.

Theorem 3.8. *Let H be a subset of $C(K,\mathcal{K}_\mathcal{C})$ that contains the constant functions. Assuming that $Conv(H)$ separates points, then H is D_∞-dense in $C(K,\mathcal{K}_\mathcal{C})$.*

Proof. Let f be in $C(K,\mathcal{K}_\mathcal{C})$ and fix $\varepsilon > 0$. We need to find $g \in H$ such that $D_\infty(f,g) < \varepsilon$.

Take $x \in K$ and $0 < \varepsilon(x) < \varepsilon$ and let us define the following open neighborhood of x:

$$V(x) := \{t \in K : d_H(f(t), f(x)) < \varepsilon(x) < \varepsilon\}.$$

Apply, then, Lemma 3.7 to get $U(x)$, an open neighborhood of x, satisfying the properties there.

Fix any point $x_1 \in K$ and take $W = K \setminus V(x_1)$ which turns out to be a compact set; so we can find a finite number of points, namely $x_2, \ldots, x_m$, in W such that

$$W \subset U(x_2) \cup \ldots \cup U(x_m).$$

Take $M = \max_{1\le i\le m}\{D_\infty(f, \overline{f_{x_i}})\}$ and $\varepsilon' = \max_{1\le i\le m}\{\varepsilon(x_i)\}$ and let us choose $0 < \delta < \frac{1}{2}$ such that $\delta M m < \varepsilon - \varepsilon'$.

On the other hand, Lemma 3.7 also gives $\phi_2, \cdots, \phi_m \in Conv(H)$ such that, for all $i = 2, \ldots, m$,

(i) $\phi_i(t) > 1 - \delta$, for all $t \in U(x_i)$;
(ii) $0 \le \phi_i(t) < \delta$, $\quad if \;\; t \notin V(x_i)$.

Let us define the following functions which belong to $Conv(H)$ as well:

$\psi_2 := \phi_2$,
$\psi_3 := (1-\phi_2)\phi_3$,
$\vdots$
$\psi_m := (1-\phi_2)(1-\phi_3)\cdots(1-\phi_{m-1})\phi_m$.

Next we can compute the sum $\psi_2 + \ldots + \psi_j$, $j = 2, \ldots, m$ and we are going to show, by induction, that,

$$\psi_2 + \ldots + \psi_j = 1 - (1-\phi_2)(1-\phi_3)\cdots(1-\phi_j), \quad j = 2, \ldots, m.$$

It is clear that $\psi_2 + \psi_3$ follows the rule since

$$\psi_2 + \psi_3 = \phi_2 + (1-\phi_2)\phi_3 = 1 - (1-\phi_2)\cdot(1-\phi_3).$$

Assume that it is also true for a certain $j \in \{4, ..., m-1\}$ and let us check

$$\psi_2 + \ldots + \psi_j + \psi_{j+1} = 1 - (1-\phi_2)(1-\phi_3)\cdots(1-\phi_j)(1-\phi_{j+1}).$$

Then, it is easily checked that

$$\begin{aligned}\psi_2 + \ldots + \psi_j + \psi_{j+1} &= 1 - (1-\phi_2)(1-\phi_3)\cdots(1-\phi_j) \\ &\quad + (1-\phi_2)(1-\phi_3)\cdots(1-\phi_j)\phi_{j+1} \\ &= 1 - (1-\phi_2)(1-\phi_3)\cdots(1-\phi_j)(1-\phi_{j+1})\end{aligned}$$

as was to be checked and letting us define $\psi_1 := (1-\phi_2)\cdots(1-\phi_m)$, which also belongs to $Conv(H)$ by Proposition 3.2, allows us to get m functions in $Conv(H)$ satisfying

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

Next we claim that

$$\psi_i(t) < \delta \quad \text{for all } t \notin V(x_i),\ i = 1, \ldots, m. \tag{3.1}$$

It is apparent when $i \geq 2$, since $\psi_i(t) \leq \phi_i(t) < \delta$ for all $t \notin V(x_i)$ from (ii) above. So we only need to show it for $i = 1$. If $t \notin V(x_1)$, we have $t \in W$. Hence, $t \in U(x_j)$ for some $j = 2, \ldots, m$. From (i), we have $1 - \phi_j(t) < \delta$ and then

$$\psi_1(t) = (1-\phi_j(t)) \prod_{i \neq j} (1 - \phi_i(t)) < \delta.$$

Finally, we can define

$$g := \psi_1 \overline{f_{x_1}} + \psi_2 \overline{f_{x_2}} + \ldots + \psi_m \overline{f_{x_m}}, \tag{3.2}$$

Since $\psi_i \in Conv(H)$ for $i = 1, ..., m$ (see Definition 3.1), we have $g \in H$ and it remains to show that this function g satisfies the requested property.

Fix now $x_0 \in K$ and, from Proposition 2.1, we get that

$$\begin{aligned} d_H(f(x_0), g(x_0)) &= d_H\left(\sum_{i=1}^m \psi_i(x_0) f(x_0), \sum_{i=1}^m \psi_i(x_0) f(x_i))\right) \\ &\leq \sum_{i=1}^m \psi_i(x_0) d_H(f(x_0), f(x_i)). \end{aligned}$$

Let us split the set $\{1, 2 \ldots, m\}$ into two disjoint sets: $\mathcal{I} = \{1 \leq i \leq m : x_0 \in V(x_i)\}$ and $\mathcal{J} = \{1 \leq i \leq m : x_0 \notin V(x_i)\}$. Then, for all $i \in \mathcal{I}$, we have

$$\psi_i(x_0) d_H(f(x_0), f(x_i)) \leq \psi_i(x_0)\varepsilon'$$

and, for all $i \in \mathcal{J}$, the inequality (3.1) yields

$$\psi_i(x_0) d_H(f(x_0), f(x_i)) \leq \delta M.$$

From these two inequalities, we deduce

$$\sum_{i=1}^m \psi_i(x_0) d_H(f(x_0), f(x_i)) \leq \sum_{i \in \mathcal{I}} \psi_i(x_0)\varepsilon' + \sum_{i \in \mathcal{J}} \delta Q \leq \varepsilon' + \delta M m < \varepsilon.$$

Finally, gathering all the information above, we infer $d_H(f(x_0), g(x_0)) < \varepsilon$ for all $x_0 \in K$, which yields $D_\infty(f, g) \leq \varepsilon$ as desired.

□

Given $J \in \mathcal{K}_\mathcal{C}$, we shall keep writing $\overline{J}$ to denote the function in $C(K, \mathcal{K}_\mathcal{C})$ which takes the constant value J.

Corollary 3.9. *Given $f \in C(K, \mathcal{K_C})$, there exist finitely many functions $\psi_i \in C(K,[0,1])$ and $J_i \in \mathcal{K_C}$, $i = 1, ..., m$, such that*

$$D_\infty(f, \psi_1 \overline{J_1} + ... + \psi_m \overline{J_m}) < \varepsilon.$$

Proof. Let us first remark that $Conv(C(K, \mathcal{K_C})) = C(K,[0,1])$, which clearly separates the points of K. Then, it is enough to take $J_i := f(x_i)$, $i = 1, ..., m$, in the definition of the function g in the proof of Theorem 3.8 (see formula (3.2)).

□

4. A Jackson type approximation result for interval-valued functions.

As pointed out in the corollary above, the family of all finite sums of the form $\sum_{i=0}^m \psi_i(x) A_i$, where $\psi_i(x)$ are continuous functions from K into $[0,1]$ and A_i are closed intervals of the real line, is dense in $C(K, \mathcal{K_C})$. When we consider the particular case $K = [a, b]$, we will be able to provide an upper bound of the approximation error between a continuous function and a member of such family, obtaining a Jackson-type result. Similar results can be in found in, e.g., [5], [11] and [2] for classical neural networks and in [9] in the fuzzy setting. We follow the techniques in [5], but some difficulties arise because we deal with interval arithmetic, whose properties differ considerably from those of the arithmetic of real numbers. Without loss of generality, we will consider the unit interval instead of $[a, b]$.

Let $f \in C([0,1], \mathcal{K_C})$. The modulus of continuity of f is defined to be

$$\omega(f, \delta) := \sup\{d_H(f(x), f(y)) \ : \ x, y \in [0,1]; |x - y| < \delta\}.$$

Let $\mathfrak{T}_n$ denote the family

$$\left\{ \sum_{i=0}^{n} \psi_i(x) A_i \ : \ \psi_i \in C([0,1],[0,1]); \ A_i \in \mathcal{K_C}, = 0, 1, 2, \dots n \right\}.$$

Define the approximation error between a member of $\mathfrak{T}_n$ and a continuous function $f \in C([0,1], \mathcal{K_C})$ by

$$E_{n,f} := \inf_{g \in \mathfrak{T}_n} D_\infty(f, g)$$

As usual, the main concern when dealing with interval-valued functions is to find a well-behaved substraction for intervals with respect to the Hausdorff metric and with some sort of cancelation law. Namely, Minkowski difference has not the desired properties but there have been other approaches to provide suitable interval differences (see, for instance, Hukuhara [12], Markov [16], Lodwick [15], Chalco-Cano et al [3]). Maybe the most used, due to its simplicity, is the Hukuhara difference which was generalized in 2008 by L. Stefanini ([22],[23]). This generalization has the requested properties we need.

Let us denote by $A \ominus B$ the generalized Hukuhara difference, gH-difference for short, defined in [22] by

$$A \ominus B = C \Leftrightarrow \begin{cases} A = B + C \\ \text{or} \\ B = A + (-1)C \end{cases}$$

It is worth remarking that the gH-difference always exists in $\mathcal{K}_{\mathcal{C}}$ and in fact,

$$[a^-, a^+] \ominus [b^-, b^+] = [\min\{a^- - b^-, a^+ - b^+\}, \max\{a^- - b^-, a^+ - b^+\}]$$

If $A = \{a\}$ and $B = \{b\}$ are two singletons, then $A \ominus B = a - b$.

We denote by $L(A)$ the diameter or length of the interval A; that is, $L(A) = a^+ - a^-$, when $A = [a^-, a^+]$.

The following properties will be useful in the sequel and can be found in [23]:

Proposition 4.1. *Let $A = [a^-, a^+]$, $B = [b^-, b^+]$ be in $\mathcal{K}_{\mathcal{C}}$. Then*

(1) $A \ominus A = 0$.
(2) *Either* $A + (B \ominus A) = B$ *or* $B - (B \ominus A) = A$.
(3) *If* $L(A) \leq L(B)$, *then* $A + (B \ominus A) = B$.
(4) $d_H(A \ominus B, 0) = d_H(A, B)$

The cancelation law (item (3) above) allows us to provide the following property:

Proposition 4.2. *Let $\{A_j\}_{j=0}^n$ be a family of subsets of $\mathcal{K}_{\mathcal{C}}$.*

(1) *If $L(A_j) \geq L(A_{j+1})$, for all $j = 0, 1, \dots, n-1$ then, for any $0 \leq k \leq n-1$,*

$$A_n + \sum_{j=k}^{n-1} (A_j \ominus A_{j+1}) = A_k$$

(2) *If $L(A_j) \leq L(A_{j+1})$, for all $j = 0, 1, \dots, n-1$ then, for any $0 \leq k \leq n-1$,*

$$A_0 + \sum_{j=0}^{k} (A_{j+1} \ominus A_j) = A_{k+1}$$

Proof. For (1) we use that, for every j, $A_{j+1} + (A_j \ominus A_{j+1}) = A_j$.

$$
\begin{aligned}
A_n + \sum_{j=k+1}^{n-1} (A_j \ominus A_{j+1}) &= A_n + (A_{n-1} \ominus A_n) \\
&\quad + (A_{n-2} \ominus A_{n-1}) + \dots \\
&= A_{n-1} + (A_{n-2} \ominus A_{n-1}) \\
&\quad + (A_{n-3} \ominus A_{n-2}) + \dots \\
&\ \ \vdots \\
&= A_{k+2} + (A_{k+1} \ominus A_{k+2}) + (A_k \ominus A_{k+1}) \\
&= A_{k+1} + (A_k \ominus A_{k+1}) = A_k
\end{aligned}
$$

The proof of (2) is similar but using $A_j + (A_{j+1} \ominus A_j) = A_{j+1}$, for every j. □

Next set $A_j := f(j/n) \in \mathcal{K_C}$ for $j = 0, 1, \dots, n-1$, where $f \in C([0,1], \mathcal{K_C})$. Then, Proposition 4.2 forces us to demand some kind of monotonicity as follows: for a continuous function $f \in C([0,1], \mathcal{K_C})$ we define the length function $len(f) : [0,1] \to \mathbb{R}$ as

$$len(f)(x) = f(x)^+ - f(x)^-, \qquad \text{for all } x \in [0,1]$$

where $f(x) = [f(x)^-, f(x)^+]$. As usual, the length function is said non-increasing when

$$len(f)(x) \geq len(f)(y), \quad \text{when } x \leq y. \tag{4.1}$$

and non-decreasing when

$$len(f)(x) \leq len(f)(y), \quad \text{when } x \leq y. \tag{4.2}$$

For instance, type (i)-gH-differentiable functions satisfy (4.2) and type (ii)-gH-differentiable functions satisfy (4.1) (see [4, Proposition 3.8]).

So, for every $j = 0, 1, \dots n-1$, we have:

$$A_{j+1} + (A_j \ominus A_{j+1}) = A_j, \quad \text{when (4.1) is satisfied.}$$

and

$$A_j + (A_{j+1} \ominus A_j) = A_{j+1}, \quad \text{when (4.2) is satisfied.}$$

Now, we are ready to show the following Jackson-type theorem:

Theorem 4.3. *Let $f \in C([0,1], \mathcal{K_C})$ satisfying (4.1). Then, for every $n \in \mathbb{N}$,*

$$E_{n,f} \leq 2\omega\left(f, \frac{1}{n}\right)$$

Proof. Fix $\varepsilon > 0$ and $n \in \mathbb{N}$. Consider the points $a_j = \frac{j}{n}$, $j = 0, 1, \ldots, n$. Take $0 < \delta < \frac{1}{2n}$. Being f a continuous function on a compact set, it is bounded, so we can find a constant $M > 0$ such that $d_H(f(x), f(y)) \leq 2M$, for all $x, y \in [0, 1]$. Take $\epsilon' = \frac{\varepsilon}{2nM}$ and apply Lemma 3.3 to $a_j - \delta < a_j + \delta$. So we can find $n+1$ continuous functions ψ_j, $j = 0, 1, \ldots n$, with

(1) $\psi_j(t) < \epsilon'$, for $t \geq a_j + \delta$, for all $j = 0, 1, \ldots, n-1$;
(2) $\psi_j(t) > 1 - \epsilon'$, for $t \leq a_j - \delta$, for all $j = 1, \ldots, n$.

For every $j = 0, 1, \ldots, n$, set $A_j = f(a_j)$. Define now

$$g(x) := A_n + \sum_{j=0}^{n-1} \psi_j(x) \Big(A_j \ominus A_{j+1} \Big) \in \mathfrak{T}_n$$

Now, take any $x \in [0, 1]$. We are going to study the value of $d_H(f(x), g(x))$ and two cases arise depending on where x is located.

Case 1. There exists k with $x \in [a_k - \delta, a_k + \delta]$ (or $0 \leq x \leq \delta$ if $k = 0$ or $1 - \delta \leq x \leq 1$ if $k = n$).

In this case, $x > a_j + \delta$, for $j < k$ and $x < a_j - \delta$, for $j > k$. Then $\psi_j(x) < \epsilon'$ for $j < k$ and $1 - \psi_j(x) < \epsilon'$, for $j > k$.

Then we can rewrite $g(x)$ as

$$\begin{aligned} g(x) = &\sum_{j=0}^{k-1} \psi_j(x) \Big(A_j \ominus A_{j+1} \Big) \\ &+ \psi_k(x) \Big(A_k \ominus A_{k+1} \Big) \\ &+ \sum_{j=k+1}^{n-1} \psi_j(x) \Big(A_j \ominus A_{j+1} \Big) + A_n \end{aligned} \tag{4.3}$$

Using (5) from Proposition 2.1 and (1) from Proposition 4.2, we have

$$\begin{aligned} d_H(g(x), f(x)) &= d_H(g(x) + A_k,\ f(x) + A_k) \\ &= d_H\Big(g(x) + A_k,\ f(x) + A_n + (A_k \ominus A_{k+1}) + \sum_{j=k}^{n-1} (A_j \ominus A_{j+1}) \Big) \end{aligned}$$

Finally, by (1) and (2) in Proposition 2.1, we infer

$$
\begin{aligned}
d_H(g(x), f(x)) &\leq \sum_{j=0}^{k-1} \psi_j(x) d_H(A_j \ominus A_{j+1},\ \mathbf{0}) \\
&+ d_H\Big(\psi_k(x)(A_k \ominus A_{k+1}),\ A_k \ominus A_{k+1}\Big) \\
&+ \sum_{j=k+1}^{n-1} d_H\Big(\psi_j(x)(A_j \ominus A_{j+1}),\ A_j \ominus A_{j+1}\Big) \\
&+ d_H(A_k,\ f(x)) \\
&\leq \sum_{j=0}^{k-1} \psi_j(x) d_H(A_j,\ A_{j+1}) \\
&+ (1-\psi_k(x)) d_H(A_k,\ A_{k+1}) \\
&+ \sum_{j=k+1}^{n-1} (1-\psi_j(x)) d_H(A_j,\ A_{j+1}) \\
&+ d_H(A_k,\ f(x)) \\
&\leq k\epsilon'\, 2M + \omega(f, 1/n) + (n-1-k)\epsilon'\, 2M + \omega(f, 1/n) \\
&< 2\omega(f, 1/n) + \varepsilon
\end{aligned}
$$

Case 2. There exists $0 \leq k \leq n-1$ with $x \in (a_k + \delta, a_{k+1} - \delta)$. In this case, $x > a_j + \delta$, for $j \leq k$ and $x < a_j - \delta$, for $j > k$. Then $\psi_j(x) < \epsilon'$ for $j \leq k$ and $1 - \psi_j(x) < \epsilon'$, for $j > k$.

Now, we can rewrite $g(x)$ as

$$
\begin{aligned}
g(x) = &\sum_{j=0}^{k} \psi_j(x)\Big(A_j \ominus A_{j+1}\Big) \\
&+ \sum_{j=k+1}^{n-1} \psi_j(x)\Big(A_j \ominus A_{j+1}\Big) + A_n
\end{aligned}
\tag{4.4}
$$

And, as in the previous case, but bearing in mind that now $d_H(A_{k+1}, f(x)) \leq \omega(f, 1/n)$ as well,

$$\begin{aligned}
d_H(g(x), f(x)) &= d_H(g(x) + A_{k+1},\ f(x) + A_{k+1}) \\
&= d_H\Big(g(x) + A_{k+1},\ f(x) + A_n + \sum_{j=k+1}^{n-1} (A_j \ominus A_{j+1})\Big) \\
&\leq \sum_{j=0}^{k} \psi_j(x) d_H(A_j, A_{j+1}) \\
&+ \sum_{j=k+1}^{n-1} (1 - \psi_j(x))\, d_H(A_j, A_{j+1}) \\
&+ d_H(A_{k+1}, f(x)) \\
&\leq (k+1)\epsilon'\, 2M + (n-1-k)\epsilon'\, 2M + \omega(f, 1/n) \\
&< 2\omega(f, 1/n) + \varepsilon
\end{aligned}$$

So we get, for all $x \in [0, 1]$,

$$d_H(g(x), f(x)) \leq 2\omega\left(f, \frac{1}{n}\right) + \varepsilon$$

which means,

$$D_\infty(f, g) = \sup_{x \in [0,1]} d_H(f(x), g(x)) \leq 2\omega\left(f, \frac{1}{n}\right) + \varepsilon$$

Then,

$$\inf_{h \in \mathfrak{T}_n} D_\infty(f, h) \leq D_\infty(f, g) \leq 2\omega\left(f, \frac{1}{n}\right) + \varepsilon$$

and being true for all $\varepsilon > 0$, we finally get

$$E_{n,f} \leq 2\omega\left(f, \frac{1}{n}\right).$$

$\square$

Theorem 4.4. *Let $f \in C([0, 1], \mathcal{K}_{\mathcal{C}})$ satisfying 4.2. Then, for every $n \in \mathbb{N}$,*

$$E_{n,f} \leq 2\omega\left(f, \frac{1}{n}\right)$$

Proof. The proof goes parallel to the previous theorem but taking $\phi_i = 1 - \psi_i$ and considering the function in $\mathfrak{T}_n$ defined by

$$g(x) := A_0 + \sum_{j=0}^{n-1} \phi_j(x)(A_{j+1} \ominus A_j) \in \mathfrak{T}_n$$

$\square$

5. Approximation by means of interval-valued neural networks

In [13] (see also [20]), an architecture of neural networks with interval weights and interval biases was proposed. Such networks maps an input of real numbers to an output interval and they are called interval neural networks.

Based on the results of the previous section and taking advantage of an striking result by Guliyev and Ismailov ([10]), we show how an interval-valued continuous functions can be approximated using interval neural networks. Namely, we shall deal with interval neural networks of the following form:

$$H(x) = \sum_{i=1}^{m} W_i \left(c_{i1} \cdot \sigma(x - \theta_{i1}) + c_{i2} \cdot \sigma(x - \theta_{i2}) \right)$$

for each $x \in \mathbb{R}$, where the weights $W_i \in \mathcal{K}_{\mathcal{C}}$ and the weights c_{ij} and the thresholds θ_{ij} are real numbers for $i = 1, ...m$ and $j = 1, 2$. Here $\sigma : \mathbb{R} \to \mathbb{R}$ stands for the activation function in the hidden layer.

Theorem 5.1. *Let K be a compact subset of $\mathbb{R}$ and let $f \in C(K, \mathcal{K}_{\mathcal{C}})$ and $\varepsilon > 0$. Then there exists a sigmoidal function $\sigma : \mathbb{R} \to \mathbb{R}$ and an interval neural network $H(x)$ as above such that $D_\infty(f, H) < \varepsilon$.*

Proof. By Corollary 3.9, there exist finitely many functions $\psi_i \in C(K, [0, 1])$ and $J_i \in \mathcal{K}_{\mathcal{C}}$, $i = 1, ..., m$, such that

$$D_\infty(f, \psi_1 \widehat{J_1} + ... + \psi_m \widehat{J_m}) < \frac{\varepsilon}{2}.$$

On the other hand, by [10, Theorem 4.2], we know that, for each ψ_i, $i = 1, ..., m$, there exist weights c_{i1}, c_{i2}, thresholds θ_{i1}, θ_{i2}, in $\mathbb{R}$ and a sigmoidal function $\sigma : \mathbb{R} \to \mathbb{R}$ such that

$$\left| \psi_i(x) - \sum_{i=1}^{m} \left(c_{i1} \cdot \sigma(x - \theta_{i1}) + c_{i2} \cdot \sigma(x - \theta_{i2}) \right) \right| < \frac{\varepsilon}{2m \cdot d_H(J_i, 0)},$$

for all $x \in K$. Hence

$$\begin{aligned} & D_\infty \left(\left(c_{i1} \cdot \sigma(x - \theta_{i1}) + c_{i2} \cdot \sigma(x - \theta_{i2}) \right) \widehat{J_i}, \psi_i(x) \cdot \widehat{J_i} \right) \\ &= \sup_{x \in K} d_H \left(\left(c_{i1} \cdot \sigma(x - \theta_{i1}) + c_{i2} \cdot \sigma(x - \theta_{i2}) \right) J_i, \psi_i(x) \cdot J_i \right) \\ &= \sup_{x \in K} \left| \psi_i(x) - \left(c_{i1} \cdot \sigma(x - \theta_{i1}) + c_{i2} \cdot \sigma(x - \theta_{i2}) \right) \right| d_H(J_i, 0) < \frac{\varepsilon}{2m} \end{aligned}$$

which clearly yields

$$D_\infty \left(f, \ \sum_{i=1}^{m} \widehat{J_i} \left(c_{i1} \cdot \sigma(x - \theta_{i1}) + c_{i2} \cdot \sigma(x - \theta_{i2}) \right) \right) < \varepsilon.$$

□

References

[1] J. P. Aubin, H. Franskowska, *Set Valued Analysis*, Birkhauser, London (1990). 2

[2] F. Cao; T. Xie; Z. Xu, *The estimate for approximation error of neural networks: A constructive approach*, Neurocomputing, **71** (2008), 626–630. 4

[3] Y. Chalco-Cano, W. A. Lodwick, B. Bede, *Single level constraint interval arithmetic*, Fuzzy Sets and Systems, **257**, (2014), 146–168. 4

[4] Y. Chalco-Cano, T. M. Costa, H. Román-Flores, A. Rufián-Lizana, *New properties of the switching points for the generalized Hukuhara differentiability and some results on calculus*, Fuzzy Sets and Systems, **404**, (2021), 62–74. 4

[5] D. Chen, *Degree of approximation by superpositions of a sigmoidal function*, Approximation Theory and its Applications, **9** (1993), no. 3, 17–28. 4

[6] P. Diamond, P. Kloeden, *Metric Spaces of Fuzzy Sets: Theory and Applications*, World Scientific, Singapore, (1994). 2, 2

[7] P. S. Dwyer, *Linear Computations*, John Wiley and Sons, New York, (1951). 1

[8] 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. 1, 3, 3

[9] S.G. Gal, *Degree of approximation of fuzzy mappings by fuzzy polynomials*, J. Fuzzy Math. **2** (1994) (4), 847–853. 4

[10] N.J. Guliyev, V.E. Ismailov, *On the approximation by single hidden layer feedforward neural networks with fixed weights*, Neural Networks, **98** (2018), 296–304. 1, 5, 5

[11] B. I. Hong; N. Hahm, *Approximation Order to a Function in $C(\mathbb{R})$ by Superposition of a Sigmoidal Function* , Applied Mathematical Letters, **15** (2002), 591–597. 4

[12] M. Hukuhara, *Integration des applications mesurables dont la valeur est un compact convexe*, Funkc. Ekvacioj, **10** (1967), 205–223. 1, 4

[13] H. Ishibuchi, H. Tanaka, H. Okada, *An architecture of neural networks with interval weights and its application to fuzzy regresion analysis*, Fuzzy Sets and Systems, **57** (1993), 27-39. 5

[14] R.I. Jewett, *A variation on the Stone-Weierstrass theorem*, Proc. Amer. Math. Soc., **14** (1963), 690–693. 1, 3, 3.3, 3.4

[15] W. A. Lodwick, *Constrained interval arithmetic*, Technical Report, Center for Computational Mathematics, University of Colorado at Denver, Denver, USA, 1999. 4

[16] S. Markov, *Extended interval arithmetic*, C. R. Acad. Bulgare Sci., **30** (1977), 1239–1242. 4

[17] R. E. Moore, *Automatic error analysis in digital computation*, LMSD Technical report 48421, Lockheed Aircraft Corporation, Sunnyvale, California, January 1959. 1

[18] R. E. Moore, C. T. Yang, *Interval Analysis I*, LMSD Technical report 285875, Lockheed Aircraft Corporation, Sunnyvale, California, September 1959. 1

[19] R. E. Moore, Interval Analysis. Prentice-Hall, Englewood Cliffs N. J., 1966. 1

[20] R. B. Patil, *Interval neural networks* in *Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, TX, 1995*, Reliable Computing, (1995). Supplement. 5

[21] J.B. Prolla, *On the Weierstrass-Stone Theorem*, J. Approx. Theory, **78** (1994), 299–313. 1

[22] L. Stefanini; *A generalization of Hukuhara difference for interval and fuzzy arithmetic*, D. Dubois, M.A. Lubiano, H. Prade, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), Soft Methods for Handling Variability and Imprecision, Series on Advances in Soft Computing, vol. 48, Springer (2008) An extended version is available online at the RePEc service: http://econpapers.repec.org/RAS/pst233.htm 4

[23] L. Stefanini, *A generalization of Hukuhara difference and division for interval and fuzzy arithmetic*, Fuzzy Sets Syst., **161** (2010), 1564–1584. 4

[24] Sunaga, T., *Theory of interval algebra and its application to numerical analysis*, RAAG Memoirs, Ggujutsu Bunken Fukuy-kai, Tokyo, 2 (1958), 29-48. 1

[25] M. Warmus, *Calculus of approximations*, Bulletin de l'Academie Polonaise de Sciences, **4**, 5 (1956), 253-257. 1

Institut Universitari de Matemàtiques i Aplicacions de Castelló (IMAC), Universitat Jaume I, Avda. Sos Baynat s/n, 12071, Castelló (Spain).

Email address: `font@uji.es`

Institut Universitari de Matemàtiques i Aplicacions de Castelló (IMAC), Universitat Jaume I, Avda. Sos Baynat s/n, 12071, Castelló (Spain).

Email address: `macario@uji.es`