Degree of Approximation by Kantorovich-Choquet quasi-interpolation neural network operators revisited

George A. Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

Abstract

In this article we exhibit univariate and multivariate quantitative approximation by Kantorovich-Choquet type quasi-interpolation neural network operators with respect to supremum norm. This is done with rates using the first univariate and multivariate moduli of continuity. We approximate continuous and bounded functions on $\mathbb{R}^N$, $N \in \mathbb{N}$. When they are also uniformly continuous we have pointwise and uniform convergences. Our activation functions are induced by the arctangent, algebraic, Gudermannian and generalized symmetrical sigmoid functions.

2020 AMS Mathematics Subject Classification: 41A17, 41A25, 41A30, 41A35.

Keywords and Phrases: arctangent-algebraic-Gudermannian-generalized symmetrical activation functions, univariate and multivariate quasi-interpolation neural network approximation, Kantorovich-Choquet type operators.

1 Introduction

The author in [1] and [2], see Chapters 2-5, was the first to establish neural network approximations to continuous functions with rates by very specifically defined neural network operators of Cardaliaguet-Euvrard and "Squashing" types, by employing the modulus of continuity of the engaged function or its high order derivative, and producing very tight Jackson type inequalities. He treats there both the univariate and multivariate cases. The defining these operators "bell-shaped" and "squashing" functions are assumed to be compact support. Also

RGMIA Res. Rep. Coll. **25** (2022), Art. 45, 20 pp. Received 09/04/22

in [2] he gives the Nth order asymptotic expansion for the error of weak approximation of these two operators to a special natural class of smooth functions, see Chaptes 4-5 there.

The author inspired by [20], continued his studies on neural networks approximation by introducing and using the proper quasi-interpolation operators of sigmoidal and hyperbolic tangent type which resulted into [3] - [7], by treating both the univariate and multivariate cases. He did also the corresponding fractional case [8]. For recent works see [9] - [19].

The author here performs univariate and multivariate arctangent-algebraic-Gudermannian-generalized symmetrical sigmoid activation functions based neural network approximations to continuous functions over the whole $\mathbb{R}^N$, $N \in \mathbb{N}$, then he extends his results to complex valued functions. All convergences here are with rates expressed via the modulus of continuity of the involved function and given by very tight Jackson type inequalities. This is a continuation of [12], Chapter 1.

The author comes up with the "right" precisely defined flexible quasi-interpolation, Kantorovich-Choquet type integral coefficient neural networks operators associated with: arctangent-algebraic-Gudermannian-generalized symmetrical sigmoid activation functions. In preparation to prove our results we establish important properties of the basic density functions defining our operators.

Feed-forward neural networks (FNNs) with one hidden layer, the only type of networks we deal with in this article, are mathematically expressed as

$$N_n(x) = \sum_{j=0}^{n} c_j \sigma\left(\langle a_j \cdot x \rangle + b_j\right), \quad x \in \mathbb{R}^s,\ s \in \mathbb{N},$$

where for $0 \leq j \leq n$, $b_j \in \mathbb{R}$ are the thresholds, $a_j \in \mathbb{R}^s$ are the connection weights, $c_j \in \mathbb{R}$ are the coefficients, $\langle a_j \cdot x \rangle$ is the inner product of a_j and x, and σ is the activation function of the network. In many fundamental neural network models, the activation functions are the arctangent-algebraic-Gudermannian-generalized symmetrical sigmoid activation functions. About neural networks in general read [25], [26], [27].

2 Background

Next we present briefly about the Choquet integral.

We make

Definition 1 *Consider $\Omega \neq \varnothing$ and let $\mathcal{C}$ be a σ-algebra of subsets in Ω.*

(i) (see, e.g., [28], p. 63) The set function $\mu : \mathcal{C} \to [0, +\infty]$ is called a monotone set function (or capacity) if $\mu(\varnothing) = 0$ and $\mu(A) \leq \mu(B)$ for all

$A, B \in \mathcal{C}$, with $A \subset B$. Also, μ is called submodular if

$$\mu(A \cup B) + \mu(A \cap B) \leq \mu(A) + \mu(B), \text{ for all } A, B \in \mathcal{C}.$$

μ is called bounded if $\mu(\Omega) < +\infty$ and normalized if $\mu(\Omega) = 1$.

(ii) (see, e.g., [28], p. 233, or [21]) If μ is a monotone set function on $\mathcal{C}$ and if $f : \Omega \to \mathbb{R}$ is $\mathcal{C}$-measurable (that is, for any Borel subset $B \subset \mathbb{R}$ it follows $f^{-1}(B) \in \mathcal{C}$), the for any $A \in \mathcal{C}$, the Choquet integral is defined by

$$(C)\int_A f d\mu = \int_0^{+\infty} \mu(F_\beta(f) \cap A)\, d\beta + \int_{-\infty}^0 [\mu(F_\beta(f) \cap A) - \mu(A)]\, d\beta,$$

where we used the notation $F_\beta(f) = \{\omega \in \Omega : f(\omega) \geq \beta\}$. Notice that if $f \geq 0$ on A, then in the above formula we get $\int_{-\infty}^0 = 0$.

The integrals on the right-hand side are the usual Riemann integral.

The function f will be called Choquet integrable on A if $(C)\int_A f d\mu \in \mathbb{R}$.

Next we list some well known properties of the Choquet integral.

Remark 2 *If $\mu : \mathcal{C} \to [0, +\infty]$ is a monotone set function, then the following properties hold:*

(i) For all $a \geq 0$ we have $(C)\int_A af d\mu = a \cdot (C)\int_A f d\mu$ (if $f \geq 0$ then see, e.g., [28], Theorem 11.2, (5), p. 228 and if f is arbitrary sign, then see, e.g., [22], p. 64, Proposition 5.1, (ii)).

(ii) For all $c \in \mathbb{R}$ and f of arbitrary sign, we have (see, e.g., [28], pp. 232-233, or [22], p. 65) $(C)\int_A (f + c) d\mu = (C)\int_A f d\mu + c \cdot \mu(A)$.

If μ is submodular too, then for all f, g of arbitrary sign and lower bounded, we have (see, e.g., [22], p. 75, Theorem 6.3)

$$(C)\int_A (f + g)\, d\mu \leq (C)\int_A f d\mu + (C)\int_A g d\mu.$$

(iii) If $f \leq g$ on A then $(C)\int_A f d\mu \leq (C)\int_A g d\mu$ (see, e.g., [28], p. 228, Theorem 11.2, (3) if $f, g \geq 0$ and p. 232 if f, g are of arbitrary sign).

(iv) Let $f \geq 0$. If $A \subset B$ then $(C)\int_A f d\mu \leq (C)\int_B f d\mu$. In addition, if μ is finitely aubadditive, then

$$(C)\int_{A \cup B} f d\mu \leq (C)\int_A f d\mu + (C)\int_B f d\mu.$$

(v) It is immediate that $(C)\int_A 1 \cdot d\mu(t) = \mu(A)$.

(vi) The formula $\mu(A) = \gamma(M(A))$, where $\gamma : [0,1] \to [0,1]$ is an increasing and concave function, with $\gamma(0) = 0$, $\gamma(1) = 1$ and M is a probability measure (or only finitely additive) on a σ-algebra on Ω (that is, $M(\emptyset) = 0$, $M(\Omega) = 1$ and M is countably additive), gives simple examples of normalized, monotone

and submodular set functions (see, e.g., [22], pp. 16-17, Example 2.1). Such of set functions μ are also called distorsions of countably normalized, additive measures (or distorted measures). For a simple example, we can take $\gamma(t) = \frac{2t}{1+t}$, $\gamma(t) = \sqrt{t}$.

If the above γ function is increasing, concave and satisfies only $\gamma(0) = 0$, then for any bounded Borel measure m, $\mu(A) = \gamma(m(A))$ gives a simple example of bounded, monotone and submodular set function.

(vii) If μ is a countably additive bounded measure, then the Choquet integral $(C)\int_A f d\mu$ reduces to the usual Lebesgue type integral (see, e.g., [22], p. 62, or [28], p. 226).

(viii) If $f \geq 0$, then $(C)\int_A f d\mu \geq 0$.

(ix) Let $\mu = \sqrt{M}$, where M is the Lebesgue measure on $[0, +\infty)$, then μ is a monotone and submodular set function, furthermore μ is strictly positive, see [24].

(x) If $\Omega = \mathbb{R}^N$, $N \in \mathbb{N}$, we call μ strictly positive if $\mu(A) > 0$, for any open subset $A \subseteq \mathbb{R}^N$.

2.1 About the arctangent activation function

We consider the

$$\arctan x = \int_0^x \frac{dz}{1+z^2}, \quad x \in \mathbb{R}. \tag{1}$$

We will be using

$$h(x) := \frac{2}{\pi} \arctan\left(\frac{\pi}{2}x\right) = \frac{2}{\pi} \int_0^{\frac{\pi x}{2}} \frac{dz}{1+z^2}, \quad x \in \mathbb{R}, \tag{2}$$

which is a sigmoid type function and it is strictly increasing. We have that

$$h(0) = 0, \ \ h(-x) = -h(x), \ \ h(+\infty) = 1, \ \ h(-\infty) = -1,$$

and

$$h'(x) = \frac{4}{4+\pi^2 x^2} > 0, \ \text{ all } x \in \mathbb{R}. \tag{3}$$

We consider the activation function

$$\psi_1(x) := \frac{1}{4}(h(x+1) - h(x-1)), \ \ x \in \mathbb{R}, \tag{4}$$

and we notice that

$$\psi_1(-x) = \psi_1(x), \tag{5}$$

it is an even function.

Since $x+1 > x-1$, then $h(x+1) > h(x-1)$, and $\psi_1(x) > 0$, all $x \in \mathbb{R}$.

We see that

$$\psi_1(0) = \frac{1}{\pi} \arctan\frac{\pi}{2} \cong 18.31. \tag{6}$$

Let $x > 0$, we have that

$$\psi_1'(x) = \frac{1}{4}\left(h'(x+1) - h'(x-1)\right) =$$

$$\frac{-4\pi^2 x}{\left(4+\pi^2(x+1)^2\right)\left(4+\pi^2(x-1)^2\right)} < 0. \tag{7}$$

That is

$$\psi_1'(x) < 0, \text{ for } x > 0. \tag{8}$$

That is ψ_1 is strictly decreasing on $[0, \infty)$ and clearly is strictly increasing on $(-\infty, 0]$, and $\psi_1'(0) = 0$.

Observe that

$$\begin{array}{l} \lim_{x\to+\infty} \psi_1(x) = \frac{1}{4}\left(h(+\infty) - h(+\infty)\right) = 0, \\ \text{and} \\ \lim_{x\to-\infty} \psi_1(x) = \frac{1}{4}\left(h(-\infty) - h(-\infty)\right) = 0. \end{array} \tag{9}$$

That is the x-axis is the horizontal asymptote on ψ_1.

All in all, ψ_1 is a bell symmetric function with maximum $\psi_1(0) \cong 18.31$.

We need

Theorem 3 *([11], p. 286) We have that*

$$\sum_{i=-\infty}^{\infty} \psi_1(x-i) = 1, \ \ \forall\, x \in \mathbb{R}. \tag{10}$$

Theorem 4 *([11], p. 287) It holds*

$$\int_{-\infty}^{\infty} \psi_1(x)\, dx = 1. \tag{11}$$

So that $\psi_1(x)$ is a density function on $\mathbb{R}$.

We mention

Theorem 5 *([11], p. 288) Let* $0 < \alpha < 1$, *and* $n \in \mathbb{N}$ *with* $n^{1-\alpha} > 2$. *It holds*

$$\sum_{\begin{cases} k=-\infty \\ : |nx-k| \geq n^{1-\alpha} \end{cases}}^{\infty} \psi_1(nx-k) < \frac{2}{\pi^2\left(n^{1-\alpha}-2\right)}. \tag{12}$$

We introduce (see [17])

$$Z_1(x_1, ..., x_N) := Z_1(x) := \prod_{i=1}^{N} \psi_1(x_i), \quad x = (x_1, ..., x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{13}$$

It has the properties:

(i) $Z_1(x) > 0, \ \forall x \in \mathbb{R}^N$,

(ii)

$$\sum_{k=-\infty}^{\infty} Z_1(x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z_1(x_1-k_1, ..., x_N-k_N) = 1, \tag{14}$$

where $k := (k_1, ..., k_n) \in \mathbb{Z}^N$, $\forall x \in \mathbb{R}^N$,

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_1(nx-k) = 1, \tag{15}$$

$\forall x \in \mathbb{R}^N$; $n \in \mathbb{N}$,

and

(iv)

$$\int_{\mathbb{R}^N} Z_1(x)\, dx = 1, \tag{16}$$

that is Z_1 is a multivariate density function.

(v) It is clear that

$$\sum_{\begin{cases} k = -\infty \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta} \end{cases}}^{\infty} Z_1(nx-k) < \frac{2}{\pi^2 (n^{1-\beta} - 2)} =: c_1(\beta, n), \tag{17}$$

$0 < \beta < 1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$.

Above it is $\|x\|_\infty := \max\{|x_1|, ..., |x_N|\}$, $x \in \mathbb{R}^N$, also set $\infty := (\infty, ..., \infty)$, $-\infty = (-\infty, ... -\infty)$ upon the multivariate context.

2.2 About the algebraic activation function

Here see also [17].

We consider the generator algebraic function

$$\varphi(x) = \frac{x}{\sqrt[2m]{1+x^{2m}}}, \quad m \in \mathbb{N}, x \in \mathbb{R}, \tag{18}$$

which is a sigmoidal type of function and is a strictly increasing function.

We see that $\varphi(-x) = -\varphi(x)$ with $\varphi(0) = 0$. We get that

$$\varphi'(x) = \frac{1}{(1+x^{2m})^{\frac{2m+1}{2m}}} > 0, \ \ \forall x \in \mathbb{R}, \tag{19}$$

proving φ as strictly increasing over $\mathbb{R}, \varphi'(x) = \varphi'(-x)$. We easily find that $\lim_{x \to +\infty} \varphi(x) = 1$, $\varphi(+\infty) = 1$, and $\lim_{x \to -\infty} \varphi(x) = -1$, $\varphi(-\infty) = -1$.

We consider the activation function

$$\psi_2(x) = \frac{1}{4}\left[\varphi(x+1) - \varphi(x-1)\right]. \tag{20}$$

Clearly it is $\psi_2(x) = \psi_2(-x)$, $\forall\ x \in \mathbb{R}$, so that ψ_2 is an even function and symmetric with respect to the y-axis. Clealry $\psi_2(x) > 0$, $\forall\ x \in \mathbb{R}$.

Also it is

$$\psi_2(0) = \frac{1}{2\sqrt[2m]{2}}. \tag{21}$$

By [13], we have that $\psi_2'(x) < 0$ for $x > 0$. That is ψ_2 is strictly decreasing over $(0, +\infty)$.

Clearly, ψ_2 is strictly increasing over $(-\infty, 0)$ and $\psi_2'(0) = 0$.

Furthermore we obtain that

$$\lim_{x\to+\infty}\psi_2(x) = \frac{1}{4}\left[\varphi(+\infty) - \varphi(+\infty)\right] = 0, \tag{22}$$

and

$$\lim_{x\to-\infty}\psi_2(x) = \frac{1}{4}\left[\varphi(-\infty) - \varphi(-\infty)\right] = 0. \tag{23}$$

That is the x-axis is the horizontal asymptote of ψ_2.

Conclusion, ψ_2 is a bell shape symmetric function with maximum

$$\psi_2(0) = \frac{1}{2\sqrt[2m]{2}}, \quad m \in \mathbb{N}. \tag{24}$$

We need

Theorem 6 *([13]) We have that*

$$\sum_{i=-\infty}^{\infty}\psi_2(x-i) = 1, \quad \forall\ x \in \mathbb{R}. \tag{25}$$

Theorem 7 *([13]) It holds*

$$\int_{-\infty}^{\infty}\psi_2(x)\,dx = 1. \tag{26}$$

Theorem 8 *([13]) Let $0 < \alpha < 1$, and $n \in \mathbb{N}$ with $n^{1-\alpha} > 2$. It holds*

$$\sum_{\left\{\begin{array}{c} k=-\infty \\ : |nx-k| \geq n^{1-\alpha}\end{array}\right.}^{\infty}\psi_2(nx-k) < \frac{1}{4m\left(n^{1-\alpha}-2\right)^{2m}}, \quad m \in \mathbb{N}. \tag{27}$$

We introduce (see also [18])

$$Z_2(x_1,...,x_N) := Z_2(x) := \prod_{i=1}^{N} \psi_2(x_i), \quad x = (x_1,...,x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{28}$$

It has the properties:

(i) $Z_2(x) > 0, \ \forall \, x \in \mathbb{R}^N$,

(ii)

$$\sum_{k=-\infty}^{\infty} Z_2(x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z_2(x_1-k_1,...,x_N-k_N) = 1, \tag{29}$$

where $k := (k_1,...,k_n) \in \mathbb{Z}^N$, $\forall \, x \in \mathbb{R}^N$,

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_2(nx-k) = 1, \tag{30}$$

$\forall \, x \in \mathbb{R}^N$; $n \in \mathbb{N}$,

and

(iv)

$$\int_{\mathbb{R}^N} Z_2(x)\, dx = 1, \tag{31}$$

that is Z_2 is a multivariate density function.

(v) It is clear that

$$\sum_{\substack{k=-\infty \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\infty} Z_2(nx-k) < \frac{1}{4m\left(n^{1-\beta}-2\right)^{2m}} =: c_2(\beta,n), \tag{32}$$

$0 < \beta < 1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$, $m \in \mathbb{N}$.

2.3 About the Gudermannian activation function

See also [29], [14].

Here we consider $gd(x)$ the Gudermannian function [29], which is a sigmoid function, as a generator function:

$$\sigma(x) = 2\arctan\left(\tanh\left(\frac{x}{2}\right)\right) = \int_0^x \frac{dt}{\cosh t} =: gd(x), \ x \in \mathbb{R}. \tag{33}$$

Let the normalized generator sigmoid function

$$f(x) := \frac{4}{\pi}\sigma(x) = \frac{4}{\pi}\int_0^x \frac{dt}{\cosh t} = \frac{8}{\pi}\int_0^x \frac{1}{e^t+e^{-t}} dt, \quad x \in \mathbb{R}. \tag{34}$$

Here

$$f'(x) = \frac{4}{\pi \cosh x} > 0, \quad \forall\, x \in \mathbb{R},$$

hence f is strictly increasing on $\mathbb{R}$.

Notice that $\tanh(-x) = -\tanh x$ and $\arctan(-x) = -\arctan x$, $x \in \mathbb{R}$.

So, here the neural network activation function will be:

$$\psi_3(x) = \frac{1}{4}\left[f(x+1) - f(x-1)\right], \; x \in \mathbb{R}. \tag{35}$$

By [14], we get that

$$\psi_3(x) = \psi_3(-x), \quad \forall\, x \in \mathbb{R}, \tag{36}$$

i.e. it is even and symmetric with respect to the y-axis. Here we have $f(+\infty) = 1$, $f(-\infty) = -1$ and $f(0) = 0$. Clearly it is

$$f(-x) = -f(x), \;\; \forall\, x \in \mathbb{R}, \tag{37}$$

an odd function, symmetric with respect to the origin. Since $x+1 > x-1$, and $f(x+1) > f(x-1)$, we obtain $\psi_3(x) > 0$, $\forall\, x \in \mathbb{R}$.

By [14], we have that

$$\psi_3(0) = \frac{2}{\pi} gd(1) \cong 0.551. \tag{38}$$

By [14] ψ_3 is strictly decreasing on $(0, +\infty)$, and strictly increasing on $(-\infty, 0)$, and $\psi_3'(0) = 0$.

Also we have that

$$\lim_{x \to +\infty} \psi_3(x) = \lim_{x \to -\infty} \psi_3(x) = 0, \tag{39}$$

that is the x-axis is the horizontal asymptote for ψ_3.

Conclusion, ψ_3 is a bell shaped symmetric function with maximum $\psi_3(0) \cong 0.551$.

We need

Theorem 9 *([14]) It holds that*

$$\sum_{i=-\infty}^{\infty} \psi_3(x-i) = 1, \;\; \forall\, x \in \mathbb{R}. \tag{40}$$

Theorem 10 *([14]) We have that*

$$\int_{-\infty}^{\infty} \psi_3(x)\, dx = 1. \tag{41}$$

So $\psi_3(x)$ is a density function.

Theorem 11 *([14]) Let $0 < \alpha < 1$, and $n \in \mathbb{N}$ with $n^{1-\alpha} > 2$. It holds*

$$\sum_{\begin{cases} k=-\infty \\ : |nx-k| \geq n^{1-\alpha} \end{cases}}^{\infty} \psi_3 (nx-k) < \frac{1}{\pi e^{(n^{1-\alpha}-2)}} = \frac{4e^2}{\pi e^{n^{1-\alpha}}}. \tag{42}$$

We introduce (see also [16])

$$Z_3 (x_1, ..., x_N) := Z_3 (x) := \prod_{i=1}^{N} \psi_3 (x_i), \quad x = (x_1, ..., x_N) \in \mathbb{R}^N, \ N \in \mathbb{N}. \tag{43}$$

It has the properties:

(i) $Z_3 (x) > 0, \ \forall \ x \in \mathbb{R}^N$,

(ii)

$$\sum_{k=-\infty}^{\infty} Z_3 (x-k) := \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} ... \sum_{k_N=-\infty}^{\infty} Z_3 (x_1 - k_1, ..., x_N - k_N) = 1, \tag{44}$$

where $k := (k_1, ..., k_n) \in \mathbb{Z}^N$, $\forall \ x \in \mathbb{R}^N$,

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_3 (nx-k) = 1, \tag{45}$$

$\forall \ x \in \mathbb{R}^N$; $n \in \mathbb{N}$,

and

(iv)

$$\int_{\mathbb{R}^N} Z_3 (x) \, dx = 1, \tag{46}$$

that is Z_3 is a multivariate density function.

(v) It is also clear that

$$\sum_{\begin{cases} k=-\infty \\ \left\| \frac{k}{n} - x \right\|_{\infty} > \frac{1}{n^{\beta}} \end{cases}}^{\infty} Z_3 (nx-k) < \frac{4e^2}{\pi e^{n^{1-\beta}}} = c_3 (\beta, n), \tag{47}$$

$0 < \beta < 1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$, $m \in \mathbb{N}$.

2.4 About the generalized symmetrical activation function

Here we consider the generalized symmetrical sigmoid function ([15], [23])

$$f_1 (x) = \frac{x}{\left(1 + |x|^{\mu}\right)^{\frac{1}{\mu}}}, \quad \mu > 0, \ x \in \mathbb{R}. \tag{48}$$

This has applications in immunology and protection from disease together with probability theory. It is also called a symmetrical protection curve.

The parameter μ is a shape parameter controling how fast the curve approaches the asymptotes for a given slope at the inflection point. When $\mu = 1$ f_1 is the absolute sigmoid function, and when $\mu = 2$, f_1 is the square root sigmoid function. When $\mu = 1.5$ the function approximates the arctangent function, when $\mu = 2.9$ it approximates the logistic function, and when $\mu = 3.4$ it approximates the error function. Parameter μ is estimated in the likelihood maximization ([23]). For more see [23].

Next we study the particular generator sigmoid function

$$f_2(x) = \frac{x}{\left(1 + |x|^{\lambda}\right)^{\frac{1}{\lambda}}}, \quad \lambda \text{ is an odd number, } x \in \mathbb{R}. \tag{49}$$

We have that $f_2(0) = 0$, and

$$f_2(-x) = -f_2(x), \tag{50}$$

so f_2 is symmetric with respect to zero.

When $x \geq 0$, we get that ([15])

$$f_2'(x) = \frac{1}{(1 + x^{\lambda})^{\frac{\lambda+1}{\lambda}}} > 0, \tag{51}$$

that is f_2 is strictly increasing on $[0, +\infty)$ and f_2 is strictly increasing on $(-\infty, 0]$. Hence f_2 is strictly increasing on $\mathbb{R}$.

We also have $f_2(+\infty) = f_2(-\infty) = 1$.

Let us consider the activation function ([15]):

$$\psi_4(x) = \frac{1}{4}\left[f_2(x+1) - f_2(x-1)\right] =$$

$$\frac{1}{4}\left[\frac{(x+1)}{\left(1 + |x+1|^{\lambda}\right)^{\frac{1}{\lambda}}} - \frac{(x-1)}{\left(1 + |x-1|^{\lambda}\right)^{\frac{1}{\lambda}}}\right]. \tag{52}$$

Clearly it holds ([15])

$$\psi_4(x) = \psi_4(-x), \quad \forall\, x \in \mathbb{R}. \tag{53}$$

and

$$\psi_4(0) = \frac{1}{2\sqrt[\lambda]{2}}, \tag{54}$$

and $\psi_4(x) > 0$, $\forall\, x \in \mathbb{R}$.

Following [15], we have that ψ_4 is strictly decreasing over $[0,+\infty)$, and ψ_4 is strictly increasing on $(-\infty,0]$, by ψ_4-symmetry with respect to y-axis, and $\psi_4'(0)=0$.

Clearly it is

$$\lim_{x\to+\infty}\psi_4(x)=\lim_{x\to-\infty}\psi_4(x)=0, \tag{55}$$

therefore the x-axis is the horizontal asymptote of $\psi_4(x)$.

The value

$$\psi_4(0)=\frac{1}{2\sqrt[\lambda]{2}},\ \lambda \text{ is an odd number}, \tag{56}$$

is the maximum of ψ_4, which is a bell shaped function.

We need

Theorem 12 *([15]) It holds*

$$\sum_{i=-\infty}^{\infty}\psi_4(x-i)=1,\ \ \forall\ x\in\mathbb{R}. \tag{57}$$

Theorem 13 *([15]) We have that*

$$\int_{-\infty}^{\infty}\psi_4(x)\,dx=1. \tag{58}$$

So that $\psi_4(x)$ is a density function on $\mathbb{R}$.

We need

Theorem 14 *([15]) Let $0<\alpha<1$, and $n\in\mathbb{N}$ with $n^{1-\alpha}>2$. It holds*

$$\sum_{\begin{cases} j=-\infty \\ :|nx-j|\geq n^{1-\alpha}\end{cases}}^{\infty}\psi_4(nx-j)<\frac{1}{2\lambda\left(n^{1-\alpha}-2\right)^{\lambda}}, \tag{59}$$

where $\lambda\in\mathbb{N}$ is an odd number.

We introduce (see also [19])

$$Z_4(x_1,...,x_N):=Z_4(x):=\prod_{i=1}^{N}\psi_4(x_i),\quad x=(x_1,...,x_N)\in\mathbb{R}^N,\ N\in\mathbb{N}. \tag{60}$$

It has the properties:

(i) $Z_4(x)>0,\ \ \forall\ x\in\mathbb{R}^N$,

(ii)

$$\sum_{k=-\infty}^{\infty}Z_4(x-k):=\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}...\sum_{k_N=-\infty}^{\infty}Z_4(x_1-k_1,...,x_N-k_N)=1, \tag{61}$$

where $k := (k_1, ..., k_n) \in \mathbb{Z}^N$, $\forall\ x \in \mathbb{R}^N$,

hence

(iii)

$$\sum_{k=-\infty}^{\infty} Z_4 (nx - k) = 1, \tag{62}$$

$\forall\ x \in \mathbb{R}^N$; $n \in \mathbb{N}$,

and

(iv)

$$\int_{\mathbb{R}^N} Z_4 (x)\, dx = 1, \tag{63}$$

that is Z_4 is a multivariate density function.

(v) It is clear that

$$\sum_{\substack{k = -\infty \\ \left\| \frac{k}{n} - x \right\|_\infty > \frac{1}{n^\beta}}}^{\infty} Z_4 (nx - k) < \frac{1}{2\lambda \left(n^{1-\beta} - 2\right)^\lambda} =: c_4 (\beta, n)\,, \tag{64}$$

$0 < \beta < 1$, $n \in \mathbb{N} : n^{1-\beta} > 2$, $x \in \mathbb{R}^N$, λ is odd.

For $f \in C_B^+ \left(\mathbb{R}^N\right)$ (continuous and bounded functions from $\mathbb{R}^N$ into $\mathbb{R}_+$), we define the first modulus of continuity

$$\omega_1 (f, \delta) := \sup_{\substack{x, y \in \mathbb{R}^N \\ \|x - y\|_\infty \leq h}} |f (x) - f (y)|\,, \quad h > 0. \tag{65}$$

Given that $f \in C_U^+ \left(\mathbb{R}^N\right)$ (uniformly continuous from $\mathbb{R}^N$ into $\mathbb{R}_+$, same definition for ω_1), we have that

$$\lim_{h \to 0} \omega_1 (f, h) = 0. \tag{66}$$

When $N = 1$, ω_1 is defined as in (65) with $\|\cdot\|_\infty$ collapsing to $|\cdot|$ and has the property (66).

3 Main Results

We need

Definition 15 *Let $\mathcal{L}$ be the Lebesgue σ-algebra on $\mathbb{R}^N$, $N \in \mathbb{N}$, and the set function $\mu : \mathcal{L} \to [0, +\infty)$, which is assumed to be monotone, submodular and strictly positive. For $f \in C_B^+ \left(\mathbb{R}^N\right)$, we define the general multivariate Kantorovich-Choquet type neural network operators for any $x \in \mathbb{R}^N$ $(j = 1, 2, 3, 4)$:*

$${}_jK_n^\mu (f, x) = {}_jK_n^\mu (f, x_1, ..., x_N) := \tag{67}$$

$$\sum_{k=-\infty}^{\infty}\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)}\right) Z_j(nx-k) =$$

$$\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\cdots\sum_{k_N=-\infty}^{\infty}\left(\frac{(C)\int_0^{\frac{1}{n}}\cdots\int_0^{\frac{1}{n}} f\left(t_1+\frac{k_1}{n}, t_2+\frac{k_2}{n},\ldots,t_N+\frac{k_N}{n}\right)d\mu(t_1,\ldots,t_N)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)}\right)$$

$$\left(\prod_{i=1}^{N}\psi_j(nx_i-k_i)\right),$$

where $x=(x_1,\ldots,x_N)\in\mathbb{R}^N$, $k=(k_1,\ldots,k_N)$, $t=(t_1,\ldots,t_N)$, $n\in\mathbb{N}$.

Clearly here $\mu\left(\left[0,\frac{1}{n}\right]^N\right)>0$, $\forall\ n\in\mathbb{N}$.

Above we notice that

$$\left\|{}_jK_n^{\mu}(f)\right\|_{\infty}\leq\|f\|_{\infty}, \tag{68}$$

so that ${}_jK_n^{\mu}(f,x)$ *is well-defined*, $j=1,2,3,4$.

We make

Remark 16 *Let* $f\in C_B^{+}\left(\mathbb{R}^N\right)$, $t\in\left[0,\frac{1}{n}\right]^N$ *and* $x\in\mathbb{R}^N$, *then*

$$f\left(t+\frac{k}{n}\right)=f\left(t+\frac{k}{n}\right)-f(x)+f(x)\leq\left|f\left(t+\frac{k}{n}\right)-f(x)\right|+f(x),$$

hence

$$(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)d\mu(t)\leq$$

$$(C)\int_{\left[0,\frac{1}{n}\right]^N}\left|f\left(t+\frac{k}{n}\right)-f(x)\right|d\mu(t)+(C)\int_{\left[0,\frac{1}{n}\right]^N} f(x)\,d\mu(t)= \tag{69}$$

$$(C)\int_{\left[0,\frac{1}{n}\right]^N}\left|f\left(t+\frac{k}{n}\right)-f(x)\right|d\mu(t)+f(x)\,\mu\left(\left[0,\frac{1}{n}\right]^N\right).$$

That is

$$(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)d\mu(t)-f(x)\,\mu\left(\left[0,\frac{1}{n}\right]^N\right)\leq$$

$$(C)\int_{\left[0,\frac{1}{n}\right]^N}\left|f\left(t+\frac{k}{n}\right)-f(x)\right|d\mu(t). \tag{70}$$

Similarly, we have that

$$f(x)=f(x)-f\left(t+\frac{k}{n}\right)+f\left(t+\frac{k}{n}\right)\leq\left|f\left(t+\frac{k}{n}\right)-f(x)\right|+f\left(t+\frac{k}{n}\right).$$

Hence

$$(C)\int_{\left[0,\frac{1}{n}\right]^N} f(x)\,\mu(dt) \leq$$

$$(C)\int_{\left[0,\frac{1}{n}\right]^N} \left| f\left(t+\frac{k}{n}\right) - f(x)\right| d\mu(t) + (C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)\mu(dt),$$

and

$$f(x)\,\mu\left(\left[0,\frac{1}{n}\right]^N\right) - (C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)\mu(dt) \leq$$

$$(C)\int_{\left[0,\frac{1}{n}\right]^N} \left| f\left(t+\frac{k}{n}\right) - f(x)\right| d\mu(t). \tag{71}$$

By (70) and (71) we derive that

$$\left|(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)\mu(dt) - f(x)\,\mu\left(\left[0,\frac{1}{n}\right]^N\right)\right| \leq$$

$$(C)\int_{\left[0,\frac{1}{n}\right]^N} \left| f\left(t+\frac{k}{n}\right) - f(x)\right| d\mu(t). \tag{72}$$

In particular, it holds

$$\left|\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} f\left(t+\frac{k}{n}\right)\mu(dt)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)}\right) - f(x)\right| \leq \frac{(C)\int_{\left[0,\frac{1}{n}\right]^N} \left|f\left(t+\frac{k}{n}\right) - f(x)\right| d\mu(t)}{\mu\left(\left[0,\frac{1}{n}\right]^N\right)}. \tag{73}$$

We present the following approximation result.

Theorem 17 *Let $f \in C_B^+\left(\mathbb{R}^N\right)$, $0<\beta<1$, $x \in \mathbb{R}^N$, $N, n \in \mathbb{N}$ with $n^{1-\beta} > 2$; $j = 1,2,3,4$. Then*

i)

$$\sup_{\mu} \left| {}_jK_n^{\mu}(f,x) - f(x)\right| \leq \omega_1\left(f, \frac{1}{n} + \frac{1}{n^{\beta}}\right) + 2\left\|f\right\|_{\infty} c_j(\beta,n) =: \rho_{jn}, \tag{74}$$

and

ii)

$$\sup_{\mu} \left\| {}_jK_n^{\mu}(f) - f\right\|_{\infty} \leq \rho_{jn}. \tag{75}$$

Given that $f \in \left(C_U^+\left(\mathbb{R}^N\right) \cap C_B^+\left(\mathbb{R}^N\right)\right)$, we obtain $\lim_{n\to\infty} {}_jK_n^k(f) = f$, uniformly.

Above $c_j(\beta,n)$ are as in (17), (32), (47) and (64), respectively.

Proof. We observe that

$$\left|{}_{j}K_{n}^{\mu}\left(f,x\right)-f\left(x\right)\right|=$$

$$\left|\sum_{k=-\infty}^{\infty}\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^{N}}f\left(t+\frac{k}{n}\right)d\mu\left(t\right)}{\mu\left(\left[0,\frac{1}{n}\right]^{N}\right)}\right)Z_{j}\left(nx-k\right)-\sum_{k=-\infty}^{\infty}f\left(x\right)Z_{j}\left(nx-k\right)\right|$$

$$=\left|\sum_{k=-\infty}^{\infty}\left(\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^{N}}f\left(t+\frac{k}{n}\right)d\mu\left(t\right)}{\mu\left(\left[0,\frac{1}{n}\right]^{N}\right)}\right)-f\left(x\right)\right)Z_{j}\left(nx-k\right)\right|\leq \tag{76}$$

$$\sum_{k=-\infty}^{\infty}\left|\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^{N}}f\left(t+\frac{k}{n}\right)d\mu\left(t\right)}{\mu\left(\left[0,\frac{1}{n}\right]^{N}\right)}\right)-f\left(x\right)\right|Z_{j}\left(nx-k\right)\overset{(73)}{\leq}$$

$$\sum_{k=-\infty}^{\infty}\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^{N}}\left|f\left(t+\frac{k}{n}\right)-f\left(x\right)\right|d\mu\left(t\right)}{\mu\left(\left[0,\frac{1}{n}\right]^{N}\right)}\right)Z_{j}\left(nx-k\right)=$$

$$\sum_{\substack{k=-\infty\\ \left\|\frac{k}{n}-x\right\|_{\infty}\leq\frac{1}{n^{\beta}}}}^{\infty}\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^{N}}\left|f\left(t+\frac{k}{n}\right)-f\left(x\right)\right|d\mu\left(t\right)}{\mu\left(\left[0,\frac{1}{n}\right]^{N}\right)}\right)Z_{j}\left(nx-k\right)+ \tag{77}$$

$$\sum_{\substack{k=-\infty\\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}}}^{\infty}\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^{N}}\left|f\left(t+\frac{k}{n}\right)-f\left(x\right)\right|d\mu\left(t\right)}{\mu\left(\left[0,\frac{1}{n}\right]^{N}\right)}\right)Z_{j}\left(nx-k\right)\leq$$

$$\sum_{\substack{k=-\infty\\ \left\|\frac{k}{n}-x\right\|_{\infty}\leq\frac{1}{n^{\beta}}}}^{\infty}\left(\frac{(C)\int_{\left[0,\frac{1}{n}\right]^{N}}\omega_{1}\left(f,\|t\|_{\infty}+\left\|\frac{k}{n}-x\right\|_{\infty}\right)d\mu\left(t\right)}{\mu\left(\left[0,\frac{1}{n}\right]^{N}\right)}\right)Z_{j}\left(nx-k\right)+$$

$$2\left\|f\right\|_{\infty}\left(\sum_{\substack{k=-\infty\\ \left\|\frac{k}{n}-x\right\|_{\infty}>\frac{1}{n^{\beta}}}}^{\infty}Z_{j}\left(\left|nx-k\right|\right)\right)\text{ (by (17), (32), (47), (64))}$$

$$\leq\omega_{1}\left(f,\frac{1}{n}+\frac{1}{n^{\beta}}\right)+2\left\|f\right\|_{\infty}c_{j}\left(\beta,n\right), \tag{78}$$

proving the claim. ■

Additionally we give

Definition 18 *Denote* $C_B^+(\mathbb{R}^N,\mathbb{C}) = \{f : \mathbb{R}^N \to \mathbb{C} | f = f_1 + if_2$*, where* $f_1, f_2 \in C_B^+(\mathbb{R}^N)\}$*. We set for* $f \in C_B^+(\mathbb{R}^N,\mathbb{C})$ *that*

$$ {}_jK_n^{\mu}(f,x) := {}_jK_n^{\mu}(f_1,x) + i\ {}_jK_n^{\mu}(f_2,x), \tag{79} $$

$\forall\ n \in \mathbb{N},\ x \in \mathbb{R}^N;\ j = 1,2,3,4;\ i = \sqrt{-1}.$

We give

Theorem 19 *Let* $f \in C_B^+(\mathbb{R}^N,\mathbb{C})$*,* $f = f_1 + if_2$*,* $N \in \mathbb{N}$*,* $0 < \beta < 1$*,* $x \in \mathbb{R}^N$*,* $n \in \mathbb{N}$ *with* $n^{1-\beta} > 2$*;* $j = 1,2,3,4$*. Then*

i)

$$ \sup_{\mu} \left|{}_jK_n^{\mu}(f,x) - f(x)\right| \leq \left(\omega_1\left(f_1, \frac{1}{n} + \frac{1}{n^{\beta}}\right) + \omega_1\left(f_2, \frac{1}{n} + \frac{1}{n^{\beta}}\right)\right) $$

$$ +2\left(\|f_1\|_{\infty} + \|f_2\|_{\infty}\right) c_j(\beta,n) =: \gamma_{jn}, \tag{80} $$

and

ii)

$$ \sup_{\mu} \left\|{}_jK_n^{\mu}(f) - f\right\|_{\infty} \leq \gamma_{jn}. $$

Proof. We have that

$$ \left|{}_jK_n^{\mu}(f,x) - f(x)\right| = \left|{}_jK_n^{\mu}(f_1,x) + i\ {}_jK_n^{\mu}(f_2,x) - f_1(x) - if_2(x)\right| = $$

$$ \left|\left({}_jK_n^{\mu}(f_1,x) - f_1(x)\right) + i\left({}_jK_n^{\mu}(f_2,x) - f_2(x)\right)\right| \leq $$

$$ \left|{}_jK_n^{\mu}(f_1,x) - f_1(x)\right| + \left|{}_jK_n^{\mu}(f_2,x) - f_2(x)\right| \overset{(74)}{\leq} $$

$$ \left(\omega_1\left(f_1, \frac{1}{n} + \frac{1}{n^{\beta}}\right) + 2\|f_1\|_{\infty} c_j(\beta,n)\right) + \tag{81} $$

$$ \left(\omega_1\left(f_2, \frac{1}{n} + \frac{1}{n^{\beta}}\right) + 2\|f_2\|_{\infty} c_j(\beta,n)\right), $$

proving the claim. ■

We need

Definition 20 *Let* $\mathcal{L}^*$ *be the Lebesgue* σ*-algebra on* $\mathbb{R}$*, and the set function* $\mu^* : \mathcal{L}^* \to [0,+\infty]$*, which is assumed to be monotone, submodular and strictly positive. For* $f \in C_B^+(\mathbb{R})$*, we define the general univariate Kantorovich-Choquet type neural network operator for any* $x \in \mathbb{R}$ *(*$j = 1,2,3,4$*):*

$$ {}_jM_n^{\mu^*}(f,x) = \sum_{k=-\infty}^{\infty} \left(\frac{(C)\int_0^{\frac{1}{n}} f\left(t + \frac{k}{n}\right) d\mu^*(t)}{\mu^*\left(\left[0,\frac{1}{n}\right]\right)}\right) \psi_j(nx - k). \tag{82} $$

Clearly here $\mu^*\left(\left[0,\frac{1}{n}\right]\right)>0$, $\forall\ n\in\mathbb{N}$.

Above we notice that

$$\left\| {}_{j}M_{n}^{\mu^*}(f)\right\|_{\infty}\leq\|f\|_{\infty}, \tag{83}$$

so that ${}_{j}M_{n}^{\mu^*}(f,x)$ *is well-defined,* $j=1,2,3,4$.

Notice that ${}_{j}K_{n}^{\mu}$*, when* $N=1$*, collapses to* ${}_{j}M_{n}^{\mu^*}$, $j=1,2,3,4$.

It follows another appropiate result.

Corollary 21 *(to Theorem 17 when* $N=1$*)*

Let $f\in C_{B}^{+}(\mathbb{R})$*,* $0<\beta<1$*,* $x\in\mathbb{R}$*;* $n\in\mathbb{N}$ *with* $n^{1-\beta}>2$*;* $j=1,2,3,4$*. Then*

i)

$$\sup_{\mu^*}\left| {}_{j}M_{n}^{\mu^*}(f,x)-f(x)\right|\leq\omega_1\left(f,\frac{1}{n}+\frac{1}{n^{\beta}}\right)+2\|f\|_{\infty}c_j(\beta,n)=:\varepsilon_{jn}, \tag{84}$$

and

ii)

$$\sup_{\mu}\left\| {}_{j}M_{n}^{\mu^*}(f)-f\right\|_{\infty}\leq\varepsilon_{jn}. \tag{85}$$

Given that $f\in\left(C_{U}^{+}(\mathbb{R})\cap C_{B}^{+}(\mathbb{R})\right)$*, we obtain* $\lim_{n\to\infty}{}_{j}M_{n}^{\mu^*}(f)=f$*, uniformly.*

Above $c_j(\beta,n)$ *are as in (17), (32), (47) and (64), respectively.*

Proof. As similar to Theorem 17 is omitted. ■

We need

Definition 22 *Let* $f\in C_{B}^{+}(\mathbb{R},\mathbb{C})$ *where* $f=f_1+if_2$ *with* $f_1,f_2\in C_{B}^{+}(\mathbb{R})$*. We set*

$${}_{j}M_{n}^{\mu^*}(f,x):={}_{j}M_{n}^{\mu^*}(f_1,x)+i\ {}_{j}M_{n}^{\mu^*}(f_2,x), \tag{86}$$

$\forall\ n\in\mathbb{N}$, $x\in\mathbb{R}$; $j=1,2,3,4$.

We finish with

Corollary 23 *(to Theorem 19 when* $N=1$*) Let* $f\in C_{B}^{+}(\mathbb{R},\mathbb{C})$*,* $f=f_1+if_2$*,* $0<\beta<1$*,* $x\in\mathbb{R}$*,* $n\in\mathbb{N}$ *with* $n^{1-\beta}>2$*;* $j=1,2,3,4$*. Then*

i)

$$\sup_{\mu^*}\left| {}_{j}M_{n}^{\mu^*}(f,x)-f(x)\right|\leq\left(\omega_1\left(f_1,\frac{1}{n}+\frac{1}{n^{\beta}}\right)+\omega_1\left(f_2,\frac{1}{n}+\frac{1}{n^{\beta}}\right)\right)+$$
$$2\left(\|f_1\|_{\infty}+\|f_2\|_{\infty}\right)c_j(\beta,n)=:\delta_{jn}, \tag{87}$$

and

ii)

$$\sup_{\mu^*}\left\| {}_{j}M_{n}^{\mu^*}(f)-f\right\|_{\infty}\leq\delta_{jn}.$$

Proof. As similar to Theorem 19 is omitted. ■

References

[1] G.A. Anastassiou, *Rate of convergence of some neural network operators to the unit-univariate case*, J. Math. Anal. Appli. 212 (1997), 237-262.

[2] G.A. Anastassiou, *Quantitative Approximations*, Chapman & Hall/CRC, Boca Raton, New York, 2001.

[3] G.A. Anastassiou, *Intelligent Systems: Approximation by Artificial Neural Networks*, Intelligent Systems Reference Library, Vol. 19, Springer, Heidelberg, 2011.

[4] G.A. Anastassiou, *Univariate hyperbolic tangent neural network approximation*, Mathematics and Computer Modelling, 53 (2011), 1111-1132.

[5] G.A. Anastassiou, *Multivariate hyperbolic tangent neural network approximation*, Computers and Mathematics 61 (2011), 809-821.

[6] G.A. Anastassiou, *Multivariate sigmoidal neural network approximation*, Neural Networks 24 (2011), 378-386.

[7] G.A. Anastassiou, *Univariate sigmoidal neural network approximation*, J. of Computational Analysis and Applications, Vol. 14, No. 4, 2012, 659-690.

[8] G.A. Anastassiou, *Fractional neural network approximation*, Computers and Mathematics with Applications, 64 (2012), 1655-1676.

[9] G.A. Anastassiou, *Univariate error function based neural network approximation*, Indian J. of Math., Vol. 57, No. 2 (2015), 243-291.

[10] G.A. Anastassiou, *Intelligent Systems II: Complete Approximation by Neural Network Operators*, Springer, Heidelberg, New York, 2016.

[11] G.A. Anastassiou, *Intelligent Computations: Abstract Fractional Calculus, Inequalities, Approximations*, Springer, Heidelberg, New York, 2018.

[12] G.A. Anastassiou, *Ordinary and Fractional Approximation by Non-additive Integrals: Choquet, Shilkret and Sugeno Integral Approximators*, Springer, Heidelberg, New York, 2019.

[13] G.A. Anastassiou, *Algebraic function based Banach space valued ordinary and fractional neural network approximations*, New Trends in Mathematical Sciences, 10 special issues (1) (2022), 100-125.

[14] G.A. Anastassiou, *Gudermannian function activated Banach space valued ordinary and fractional neural network approximation*, Advances in Nonlinear Variational Inequalities, 25 (2) (2022), 27-64.

[15] G.A. Anastassiou, *Generalized symmetrical sigmoid function activated Banach space valued ordinary and fractional neural network approximation*, Analele Universităţii Oradea, Fasc. Matematica, accepted for publication, 2022.

[16] G.A. Anastassiou, *Abstract multivariate Gudermannian function activated neural network approximations*, Panamerican Mathematical Journal, accepted, 2022.

[17] G.A. Anastassiou, *General multivariate arctangent function activated neural network approximations*, submitted, 2022.

[18] G.A. Anastassiou, *Abstract multivariate algebraic function activated neural network approximations*, submitted, 2022.

[19] G.A. Anastassiou, *Generalized symmetrical sigmoid function activated neural network multivariate approximation*, submitted, 2022.

[20] Z. Chen and F. Cao, *The approximation operators with sigmoidal functions*, Computers and Mathematics with Applications, 58 (2009), 758-765.

[21] G. Choquet, *Theory of capacities*, Ann. Inst. Fourier (Grenoble), 5 (1954), 131-295.

[22] D. Denneberg, *Non-additive Measure and Integral*, Kluwer, Dordrecht, 1994.

[23] A.J. Dunning, J. Kensler, L. Goudeville, F. Bailleux, *Some extensions in continuous methods for immunological correlates of protection*, BMC Medical Research Methodology 15 (107) (28 Dec. 2015), doi:10.1186/s12874-015-0096-9.

[24] S. Gal, *Uniform and Pointwise Quantitative Approximation by Kantorovich-Choquet type integral Operators with respect to monotone and submodular set functions*, Mediterranean Journal of Mathematics, 14 (2017), no. 5, Art. 205, 12 pp.

[25] I.S. Haykin, *Neural Networks: A Comprehensive Foundation* (2 ed.), Prentice Hall, New York, 1998.

[26] W. McCulloch and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, Bulletin of Mathematical Biophysics, 7 (1943), 115-133.

[27] T.M. Mitchell, *Machine Learning*, WCB-McGraw-Hill, New York, 1997.

[28] Z. Wang, G.J. Klir, *Generalized Measure Theory*, Springer, New York, 2009.

[29] E.W. Weisstein, *Gudermannian*, MathWorld.