Asian Journal of Mathematics and Computer Research

21(2): 74-86, 2017
ISSN: 2395-4205 (P), ISSN: 2395-4213 (O)

SUMMATION-INTEGRAL BERNSTEIN TYPE OF NEURAL NETWORK OPERATORS

ALI J. MOHAMMAD[1*] AND IBTIHAL JASSIM MOHAMMAD[1]

[1]Department of Mathematics, College of Education for Pure Science, University of Basrah, Basrah, Iraq.

AUTHORS' CONTRIBUTIONS

This work was carried out in collaboration between both authors. Author AJM designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Author IJM managed the analyses of the study and managed the literature searches. Both authors read and approved the final manuscript.

ARTICLE INFORMATION

Editor(s):
(1) Dariusz Jacek Jakóbczak, Technical University of Koszalin, Poland.
Reviewers:
(1) Bamadio Boureima, University of Social Scientists and Management of Bamako, Mali.
(2) Dalibor Petkovic, University in Nis, Serbia.

Received: 3rd October 2017
Accepted: 18th October 2017
Published: 25th October 2017

Original Research Article

ABSTRACT

In this paper, we introduce a family of neural network operators of summation-integral Bernstein type, which are define by using some sigmoidal functions. We give pointwise and uniform approximation theorems for these operators when are applied for continuous functions. In addition, we discuss the approximation for these operators in L^p-spaces with $1 \leq p < \infty$. Then we give some applications of the sequences of a family of linear positive multivariate neural network operators $B_n(.;\mathbf{x})$, then we describe the results by graphics of the error function for some particular values of $n = 10, 20, 30$ and for two test functions in 2-dimensional.

Keywords: Sigmoidal functions; multivariate neural network; uniform approximation; L^p-approximation; Summation-integral Bernstein-type.

AMC 2010: 41A25, 41A30, 47A58.

1 Introduction

In 1912 [1], Bernstein defined a sequence of linear positive operator called the Bernstein polynomials written as:

$$M_n(f;x) = \sum_{k=0}^{n} b_{n,k}(x) f\left(\frac{k}{n}\right), \qquad x \in [0,1], f \in C[0,1]$$

*Corresponding author: Email: alijasmoh@gmail.com;

where

$$b_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}.$$

In 1967 [2], Durrmeyer defined a sequence of summation-integral of Bernstein polynomial written as:

$$SIB_n(f;x) = (n+1)\sum_{k=0}^{n} b_{n,k}(x)\int_0^1 b_{n,k}(t)f(t)dt,$$

to approximate a function f on $C[0,1]$.

Neural networks (NNs) with one hidden layer can represented as

$$N_n(\mathbf{x}) = \sum_{j=0}^{n} c_j\sigma(\mathbf{a}_j.\mathbf{x} + b_j), \mathbf{x} \in \mathbb{R}^s, s \in \mathbb{N}^+,$$

where, for $0 \le j \le n$, the $b_j \in \mathbb{R}$, are the threshold values, the $\mathbf{a}_j \in \mathbb{R}^s$, are the weights , and the c_j are the coefficients. Here $\mathbf{a}_j.\mathbf{x}$ is the inner product in $\mathbb{R}^s$, and σ is the activation function of the network, see [3, 1996], [4,1998] and [5,2013]. The activation function is usually a sigmoidal function. Neural networks are extensively used in Approximation Theory [6,1989],[7,2003],[8,1992],[9,2009],[10,2011] and [11,2013].

Costarelli and Spigler in 2013, introduced a family of linear positive multivariate neural network (NN) operators with sigmoidal activation functions, have the form: [12,13].

$$F_n(f;\mathbf{x}) = \frac{\sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor} \dots \sum_{k_s=\lceil na_s\rceil}^{\lfloor nb_s\rfloor} f\left(\frac{\mathbf{k}}{n}\right) \Psi_\sigma(n\mathbf{x}-\mathbf{k})}{\sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor} \dots \sum_{k_s=\lceil na_s\rceil}^{\lfloor nb_s\rfloor} \Psi_\sigma(n\mathbf{x}-\mathbf{k})}.$$

For $f: \mathcal{R} \to \mathbb{R}$ be a bounded function and $n \in \mathbb{N}^+$ such that $\lceil na_i\rceil \le \lfloor nb_i\rfloor$ for every $i = 1, \dots, s$, $\mathbf{x} \in \mathcal{R} \subset \mathbb{R}^s$, $\mathbf{k}/n := (k_1/n, \dots, k_s/n)$ and $\mathcal{R} := [a_1, b_1] \times \dots \times [a_s, b_s]$. In usual, we denote to $\mathbb{R}$ be the real numbers and the symbols $\lfloor . \rfloor$ and $\lceil . \rceil$ denote taking the "integer part" and the "ceiling" of a given number, respectively. The convergence and order of approximation was study as well as pointwise and uniform convergence for the special case of (NN) operators, activation by logistic, hyperbolic tangent, and ramp sigmoidal functions.

In 2014 [14], Costarelli and Spigler introduced and studied pointwise and uniform approximation theorems for the family of neural network operators of the Kantorovich type $K_n(f, \mathbf{x})$ have the form:

$$K_n(f;\mathbf{x}) = \frac{\sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor} \dots \sum_{k_s=\lceil na_s\rceil}^{\lfloor nb_s\rfloor} \left[n^s \int_{R_{\mathbf{k}}^n} f(\mathbf{u})d\mathbf{u}\right] \Psi_\sigma(n\mathbf{x}-\mathbf{k})}{\sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor} \dots \sum_{k_s=\lceil na_s\rceil}^{\lfloor nb_s\rfloor} \Psi_\sigma(n\mathbf{x}-\mathbf{k})}.$$

For $f: \mathcal{R} \to \mathbb{R}$ be a locally integrable function and $n \in \mathbb{N}^+$ such that $\lceil na_i\rceil \le \lfloor nb_i\rfloor - 1$, $i = 1, \dots, s$, where $R_{\mathbf{k}}^n := \left[\frac{k_1}{n}, \frac{k_1+1}{n}\right] \times \dots \times \left[\frac{k_s}{n}, \frac{k_s+1}{n}\right]$, $\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{Z}^s$ and $\mathbf{x} \in \mathcal{R} \subset \mathbb{R}^s$.

For more application, we refer to [15, 16, 17 and 18].

In this paper, we study pointwise and uniform convergence of sequence $B_n(f;\mathbf{x})$ of the summation-integral Bernstein-type neural network operators.

2 Preliminary Result

In this part, we give some preliminary results where are need in our study. First, we should know a measurable function $\sigma: \mathbb{R} \to \mathbb{R}$ is a sigmonidal if and only if $\lim_{x\to-\infty} \sigma(x) = 0$ and $\lim_{x\to+\infty} \sigma(x) = 1$, some examples on sigmodal function: [19]

1. Logistic function $\sigma_l(x) = (1 + e^{-x})^{-1}$, $x \in \mathbb{R}$,
2. Hyperbolic tangent $\sigma_h(x) = \frac{1}{2}[\tanh(x) + 1]$, $x \in \mathbb{R}$,
3. Gompertz function $\sigma_{\alpha\beta}(x) = e^{\alpha e - \beta x}$, $x \in \mathbb{R}, \alpha, \beta > 0$.

In addition, in [13], for any non-decreasing function, and $\sigma(2) > \sigma(0)$ "which is a merely technical condition", also satisfying some assumptions, we state here the following**:**

1. $g_\sigma(x) = (x) - 1/2$, is an odd function,
2. $\sigma \in C^2(\mathbb{R})$ is concave for $x \geq 0$,
3. $\sigma(x) = O(|x|^{-1-\alpha})$ as $x \to -\infty$, for some $\alpha > 0$.

For every given non-decreasing function σ, satisfying all such assumptions, is defined $\Phi_\sigma(x) = \frac{1}{2}[\sigma(x+1) - \sigma(x-1)]$, $x \in \mathbb{R}$.

The following lemmas give us, a number of important properties for Φ_σ.

***Lemma* 2.1:** [14]

For the function, $\Phi_\sigma(x)$ we have:

1. $\Phi_\sigma(x) \geq 0$ for every $x \in \mathbb{R}$ and $\lim_{x\to\pm\infty} \Phi_\sigma(x) = 0$,
2. $\Phi_\sigma(x)$ is an even function,
3. For every $x \in \mathbb{R}$, $\sum_{k\in\mathbb{Z}} \Phi_\sigma(x-k) = 1$,
4. $\Phi_\sigma(x)$ is non-decreasing for $x < 0$ and non-increasing for $x \geq 0$,
5. $\Phi_\sigma(x) = O(|x|^{-1-\alpha})$ as $x \to \pm\infty$,
6. The series $\sum_{k\in\mathbb{Z}} \Phi_\sigma(x-k)$ converges uniformly on compact subsets of $\mathbb{R}$.

Note that, the multivariate function $\Psi_\sigma(\mathbf{x}) = \Phi_\sigma(x_1).\Phi_\sigma(x_1)$ ... $\Phi_\sigma(x_s)$ for the Φ_σ the next lemma, contains some properties of the function $\Psi_\sigma(\mathbf{x})$, we denote with $\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{Z}^s$.

***Lemma* 2.2:** [13]

1. For every $\mathrm{x} \in \mathbb{R}^s$, $\sum_{\mathrm{k}} \Psi_\sigma(\mathrm{x} - \mathrm{k}) = 1$.
2. The series $\sum_{\mathrm{k}} \Psi_\sigma(\mathrm{x} - \mathrm{k})$ converges uniformly on compact subsets of $\mathbb{R}^s$.
3. Denote by $\|.\|$ the usual maximum norm of $\mathbb{R}^s$, i.e., $\|\mathrm{x}\|_\infty = \max\{|\mathrm{x_i}|, \mathrm{i} = 1, \dots, \mathrm{s}\}$, with $\mathrm{x} \in \mathbb{R}^s$. For every $\gamma > 0$, we have
$$\lim_{n\to\infty} \sum_{\|x-k\|>\gamma^n} \Psi_\sigma(x-k) = 0,$$

uniformly with respect to $x \in \mathbb{R}^s$. In particular, for every $\gamma > 0$ and $0 < v < \alpha$,
$$\sum_{\|x-k\|>\gamma^n} \Psi_\sigma(x-k) = O(n^{-v}),$$

where $\alpha > 0$ is the constant in condition (3).

***Lemma* 2.3:** [12, 13]

1. Let $x \in [a, b] \subset \mathbb{R}$, $n \in \mathbb{N}^+$, so that $\lceil na \rceil \leq \lfloor nb \rfloor$, then:

$$\frac{1}{\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \Phi_\sigma(nx-k)} \leq \frac{1}{\Phi_\sigma(1)}$$

2. Let $x \in [a_1, b_1] \times \dots \times [a_s, b_s] \subset \mathbb{R}^s$ and $n \in \mathbb{N}^+$ so that $\lceil na \rceil \leq \lfloor nb \rfloor$ for every $i = 1, \dots, s$ then:

$$\frac{1}{\prod_{i=1}^{s} \sum_{k_i=\lceil na_i \rceil}^{\lfloor nb_i \rfloor} \Phi_\sigma(nx_i - k_i)} \leq \frac{1}{[\Phi_\sigma(1)]^s}.$$

***Lemma* 2.4:** [14]

We have the following facts:

1. $\int_{\mathbb{R}} \Phi_\sigma(t_i) dt_i = 1$, $t_i \in \mathbb{R}$, for all $i = 1, \dots, s$,
2. $\int_{\mathbb{R}^s} \Psi_\sigma(\mathbf{t})\, d\mathbf{t} = 1$, $\mathbf{t} \in \mathbb{R}^s$.

We need to state the following definitions:

***Definition* 2.1:** [19]

A measurable function $s: \mathbb{R} \to \mathbb{R}$ is called "activation function" whenever $\lim_{x \to -\infty} s(x) = a$ and $\lim_{x \to +\infty} s(x) = b$ with $a \neq b$.

The space of all continuous real-value functions define on $\mathcal{R}$, equipped with the sup-norm $\|.\|_\infty$ denoted by $C^0(\mathcal{R})$. [13].

3 The Main Results

As usual, we will define the operator $SIB_n(f; x)$ "summation-integral of Bernstein polynomial" in s-dimensional in the following definition:

***Definition* 3.1:**

Let $f: R \to \mathbb{R}$ be a bounded and continuous functions. The linear positive operators $SIB_n^s(f; \mathbf{x})$ for the function f in s-dimensional define as:

$$SIB_n^s(f; \mathbf{x}) = (n+1)^s \sum_{\underline{\mathbf{k}}} \mathbf{b}_{n,\mathbf{k}}(\mathbf{x}) \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) f(\mathbf{u}) d\mathbf{u}$$

where $R = [0,1] \times \dots \times [0,1]$ (s-times), $\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{Z}^s$, $\mathbf{x} \in R \subset \mathbb{R}^s$.

The following symbols in the above definition means:

$$\sum_{\underline{\mathbf{k}}} = \sum_{k_1=0}^{n} \ldots \sum_{k_s=0}^{n}$$
$$\int_R = \int_0^1 \ldots \int_0^1 \quad , (s - times)$$
$$f(\mathbf{u}) = \big(f(u_1), \ldots, f(u_s)\big),$$
$$\mathbf{b}_{n,\mathbf{k}}(\mathbf{x}) = b_{n,k_1}(x_1) \times \ldots \times b_{n,k_s}(x_s),$$

Now, we will make some modifications on the operator $F_n(f;\mathbf{x})$ by replacing the terms $f\left(\frac{\mathbf{k}}{n}\right)$ in operator $F_n(f;\mathbf{x})$ by the terms $\left((n+1)^s \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) f(\mathbf{u}) \mathrm{d}\mathbf{u}\right)$ from the operators $SIB_n^s(f;\mathbf{x})$, as in the following definition:

***Definition* 3.2**

Let $f: \mathcal{R} \to \mathbb{R}$ be a bounded and continuous function, the "Linear positive multivariate Bernstein of summation-integral-type (NN) operators, $B_n(f;.)$ activated by the sigmoidal function σ acting on, are defined by:

$$B_n(f;\mathbf{x}) = \frac{(n+1)^s \sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x} - \mathbf{k}) \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) f(\mathbf{u}) d\mathbf{u}}{\sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x} - \mathbf{k})}$$

where $\mathcal{R} = [a_1, b_1] \times \ldots \times [a_s, b_s]$, $R = [0,1] \times \ldots \times [0,1]$, $\mathbf{k} = (k_1, \ldots, k_s) \in \mathbb{Z}^s$, $n \in \mathbb{N}^+$,

$$\sum_{\mathbf{k}} = \sum_{k_1=\lceil na_1 \rceil}^{\lfloor nb_1 \rfloor} \ldots \sum_{k_s=\lceil na_s \rceil}^{\lfloor nb_s \rfloor}$$
$$\int_R = \int_0^1 \ldots \int_0^1 \quad , (s - times)$$

we observe that $B_n(1;\mathbf{x}) = 1$, for every $\mathbf{x} \in \mathcal{R} \subset \mathbb{R}^s$ and n tends to infinity, such that $\lceil na_i \rceil \leq \lfloor nb_i \rfloor$, $i = 1, \ldots, s$.

We will study pointwise and uniform convergence of $B_n(f;\mathbf{x})$, as the following:

***Theorem* 3.1**

Let $f: \mathcal{R} \to \mathbb{R}$ be a bounded and continuous function. Then $B_n(f;\mathbf{x})$ convergence to $f(\mathbf{x})$ as n tends to infinity i.e.

$$\left(\lim_{n\to\infty} B_n(f;\mathbf{x}) = f(\mathbf{x})\right)$$

at each point $\mathbf{x} \in \mathcal{R}$ where f is continuous, if $f \in C^0(\mathcal{R})$, then

$$\lim_{n\to\infty} \sup_{\mathbf{x}\in\mathcal{R}} |B_n(f;\mathbf{x}) - f(\mathbf{x})| = \lim_{n\to\infty} \|B_n(f;.) - f(.)\|_\infty = 0.$$

Proof: Let $\mathbf{x} \in \mathcal{R}$ the point of continuity of f

$$|B_n(f;\mathbf{x}) - \mathrm{f}(\mathbf{x})| = \left| \frac{(n+1)^s \sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) f(\mathbf{u}) d\mathbf{u}}{\sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x}-\mathbf{k})} - f(\mathbf{x}) \right|$$

$$= \left| \frac{\sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \left[(n+1)^s \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) f(\mathbf{u}) - f(\mathbf{x}) \right]}{\sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x}-\mathbf{k})} \right|$$

by using ***lemma* 2.3(2)**,we get

$$\leq \frac{1}{[\Phi_\sigma(1)]^s} \left| \sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \left[(n+1)^s \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) f(\mathbf{u}) d\mathbf{u} - f(\mathbf{x}) \right] \right|$$

$$\leq \frac{1}{[\Phi_\sigma(1)]^s} \sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \left| (n+1)^s \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) [f(\mathbf{u}) - f(\mathbf{x})] d\mathbf{u} \right|$$

$$\leq \frac{1}{[\Phi_\sigma(1)]^s} \sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \left[(n+1)^s \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) |f(\mathbf{u}) - f(\mathbf{x})| d\mathbf{u} \right].$$

By the continuity of f at $\mathbf{x}$**,** given $\varepsilon > 0$, $\exists\, \gamma > 0$ such that $|f(\mathbf{u}) - f(\mathbf{x})| < \varepsilon$ for every $\mathbf{u} \in \mathcal{R}$, with $\|\mathbf{u} - \mathbf{x}\|_2 < \gamma$,where $\|.\|_2$ denoted the Euclidean norm.

$$\begin{aligned}
&|B_n(f;\mathbf{x}) - f(\mathbf{x})| \\
&\leq \frac{1}{[\Phi_\sigma(1)]^s} \left\{ \sum_{\|\mathbf{u}-\mathbf{x}\|_2 < \gamma} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \left[(n+1)^s \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) |f(\mathbf{u}) - f(\mathbf{x})| d\mathbf{u} \right] \right. \\
&\left. + \sum_{\|\mathbf{u}-\mathbf{x}\|_2 \geq \gamma} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \left[(n+1)^s \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u}) |f(\mathbf{u}) - f(\mathbf{x})| d\mathbf{u} \right] \right\} \\
&:= \frac{1}{[\Phi_\sigma(1)]^s} \, (I_1 + I_2).
\end{aligned}$$

Let us estimate I_1, for $n \in \mathbb{N}^+$sufficiently large, then by the continuity of f and

***Lemma* 2.2(1),** we get

$$I_1 < \varepsilon \sum_{\|\mathbf{u}-\mathbf{x}\|_2 < \gamma} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \leq \varepsilon.$$

Moreover, by the boundedness of f and by ***Lemma* 2.2(2),** we get:

$$\begin{aligned}
&I_2 \leq 2\|f\|_\infty \sum_{\|\mathbf{u}-\mathbf{x}\|_2 \geq \gamma} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \\
&< 2\|f\|_\infty \sum_{\|\mathbf{u}-\mathbf{x}\|_2 \geq \gamma} \Psi_\sigma(n\mathbf{x}-\mathbf{k}) \\
&< 2\|f\|_\infty \varepsilon
\end{aligned}$$

uniformly with respect to $\mathbf{x} \in \mathbb{R}^s$, since ε arbitrarily we obtain the first part of the theorem. When $f \in C^0(\mathcal{R})$, the second part is similarly, by replacing γ with the parameter of the uniform continuity of f in $\mathcal{R}$. ■

Our next theorem is a result of the previous theorem**:**

***Theorem* 3.2**

For every $f \in C^0(\mathcal{R})$, we have

$$\lim_{n\to\infty} \|B_n(f;.) - f(.)\|_p = 0$$

where $\|.\|_p$ denotes the usual $L^p(\mathcal{R})$ norm, with $1 \leq p < \infty$.

Proof: Let $\varepsilon > 0$ be fixed, we have

$$\|B_n(f;.) - f(.)\|_p = \left(\int_{\mathcal{R}} |B_n(f;\mathbf{x}) - f(\mathbf{x})|^p d\mathbf{x}\right)^{\frac{1}{p}}$$

$$\leq \left(\int\limits_{\mathcal{R}} \|B_n(f;.) - f(.)\|_\infty^p d\mathbf{x}\right)^{\frac{1}{p}}$$

$$\leq \frac{1}{|\mathcal{R}|^p} \|B_n(f;.) - f(.)\|_\infty < \varepsilon.$$

For $n \in \mathbb{N}^+$ sufficiently large, as result of ***Theorem* 3.1**. Here $|\mathcal{R}|$ denotes the Lebesgue measure of $\mathcal{R}$ in $\mathbb{R}^s$, by ε arbitrarily the proof complete. ■

After the following theory, we prove convergence of the family of our operators in L^p .

***Theorem* 3.3**

The inequality

$$\|B_n(f;.) - B_n(g;.)\|_p \leq \frac{1}{[\Phi_\sigma(1)]^{s/p}} \|f - g\|_p$$

holds, for every $f, g \in L^p(\mathcal{R})$,$1 \leq p < \infty$, where $\|.\|_p$ is the usual $L^p(\mathcal{R})$ norm.

Proof: For every $f, g \in L^p(\mathcal{R})$,$1 \leq p < \infty$, we have

$$\|B_n(f;.) - B_n(g;.)\|_p = \left(\int\limits_{\mathcal{R}} |B_n(f;\mathbf{x}) - B_n(g;\mathbf{x})|^p d\mathbf{x}\right)^{\frac{1}{p}}$$

$$= \left(\int\limits_{\mathcal{R}} \left|\frac{\sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x} - \mathbf{k}) \left[(n+1)^s \int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u})[f(\mathbf{u}) - g(\mathbf{u})]\right] d\mathbf{u}}{\sum_{\mathbf{k}} \Psi_\sigma(n\mathbf{x} - \mathbf{k})}\right|^p d\mathbf{x}\right)^{\frac{1}{p}}.$$

Being $|.|^p$ convex, we infer form Jensen's inequality and ***lemma* 2.3(2)**, that

$$\|B_n(f;.)-B_n(f;.)\|_p$$

$$\leq\left(\int_{\mathcal{R}}\frac{\sum_{\mathbf{k}}\Psi_\sigma(n\mathbf{x}-\mathbf{k})\left|(n+1)^s\int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u})[f(\mathbf{u})-g(\mathbf{u})]\,d\mathbf{u}\right|^p}{\sum_{\mathbf{k}}\Psi_\sigma(n\mathbf{x}-\mathbf{k})}d\mathbf{x}\right)^{\frac{1}{p}}$$

$$\leq\frac{1}{[\Phi_\sigma(1)]^{s/p}}\left(\int_{\mathcal{R}}\sum_{\mathbf{k}}\Psi_\sigma(n\mathbf{x}-\mathbf{k})\left|(n+1)^s\int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u})[f(\mathbf{u})-g(\mathbf{u})]d\mathbf{u}\right|^p \mathrm{d}\mathbf{x}\right)^{\frac{1}{p}}$$

$$\leq\frac{1}{[\Phi_\sigma(1)]^{\frac{s}{p}}}\left(\sum_{\mathbf{k}}\int_{\mathbb{R}^s}\Psi_\sigma(n\mathbf{x}-\mathbf{k})\left|(n+1)^s\int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u})[f(\mathbf{u})-g(\mathbf{u})]d\mathbf{u}\right|^p d\mathbf{x}\right)^{\frac{1}{p}}.$$

Changing variables, setting $\mathbf{x}=(\mathbf{t}+\mathbf{k})/n$, we obtain

$$\|B_n(f;.)-B_n(g;.)\|_p$$

$$\leq\frac{1}{[\Phi_\sigma(1)]^{\frac{s}{p}}}\left(\sum_{\mathbf{k}}\frac{1}{(n+1)^s}\int_{\mathbb{R}^s}\Psi_\sigma(\mathbf{t})\left|(n+1)^s\int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u})[f(\mathbf{u})-\mathrm{g}(\mathbf{u})]d\mathbf{u}\right|^p d\mathbf{t}\right)^{\frac{1}{p}}$$

$$=\frac{1}{[\Phi_\sigma(1)]^{s/p}}\left(\frac{1}{(n+1)^s}\int_{\mathbb{R}^s}\Psi_\sigma(\mathbf{t})\sum_{\mathbf{k}}\left|(n+1)^s\int_{\mathcal{R}} \mathbf{b}_{n,\mathbf{k}}(\mathbf{u})[f(\mathbf{u})-g(\mathbf{u})]d\mathbf{u}\right|^p d\mathbf{t}\right)^{\frac{1}{p}}.$$

Using Jensen′s inequality, again, we obtain

$$\|B_n(f;.)-B_n(g;.)\|_p$$

$$\leq\frac{1}{[\Phi_\sigma(1)]^{s/p}}\left(\frac{1}{(n+1)^s}\int_{\mathbb{R}^s}\Psi_\sigma(\mathbf{t})\mathrm{d}\mathbf{t}\sum_{\mathbf{k}}(n+1)^s\int_R \mathbf{b}_{n,\mathbf{k}}(\mathbf{u})|f(\mathbf{u})-\mathrm{g}(\mathbf{u})|^p d\mathbf{u}\right)^{\frac{1}{p}}$$

$$\leq\frac{1}{[\Phi_\sigma(1)]^{s/p}}\left(\int_{\mathbb{R}^s}\Psi_\sigma(\mathbf{t})\mathrm{d}\mathbf{t}\int_{\mathcal{R}} \mathbf{b}_{n,\mathbf{k}}(\mathbf{u})|f(\mathbf{u})-\mathrm{g}(\mathbf{u})|^p d\mathbf{u}\right)^{\frac{1}{p}}$$

$$=\frac{1}{[\Phi_\sigma(1)]^{s/p}}\|f-g\|_p\left(\int_{\mathbb{R}^s}\Psi_\sigma(\mathbf{t})d\mathbf{t}\right)^{\frac{1}{p}}.$$

Now, from ***Lemma* 2.4 (2)** we get,

$$\|B_n(f;.)-B_n(g;.)\|_p\leq\frac{1}{[\Phi_\sigma(1)]^{s/p}}\|f-g\|_p. \quad ■$$

From ***Theorem* 3.3**, we conclude that the maps $B_n\colon L^p(\mathcal{R})\to L^p(\mathcal{R})$ are well-defined. Moreover displays a continuity property for the family of (NN) operators of the Bernstein type, $(B_n)_{n\in\mathbb{N}^+}$, in $L^p(\mathcal{R})$, $1\leq p<\infty$.

Now we can show the following theorem:

***Theorem* 3.4**

For every $f \in L^p(\mathcal{R})$, $1 \leq p < \infty$, we have

$$\lim_{n\to\infty} \|B_n(f;.) - f\|_p = 0$$

Proof: Let be $f \in L^p(\mathcal{R})$, and $\varepsilon > 0$ be fixed. Since the space $C^0(\mathcal{R})$ is dense in $L^p(\mathcal{R})$ with respect to norm$\|.\|_p$, there exists $g \in C^0(\mathcal{R})$ such that $\|f(.) - g(.)\|_p < \left(\Phi_\sigma(1)^{-s/p} + 1\right)^{-1}\varepsilon/2$ then by ***Theorem* 3.3** we write

$$\begin{aligned}
&\|B_n(f;.) - f(.)\|_p \\
&\leq \|B_n(f;.) - B_n(g;.)\|_p + \|B_n(g;.) - g(.)\|_p + \|g(.) - f(.)\|_p \\
&\leq \frac{1}{[\Phi_\sigma(1)]^{s/p}}\|g(.) - f(.)\|_p + \|B_n(g;.) - g(.)\|_p + \|g(.) - f(.)\|_p \\
&\leq \left(\frac{1}{[\Phi_\sigma(1)]^{s/p}} + 1\right)\|g(.) - f(.)\|_p + \|B_n(g;.) - g(.)\|_p \\
&\qquad\qquad < \frac{\varepsilon}{2} + \|B_n(g;.) - g(.)\|_p.
\end{aligned}$$

By ***Theorem* 3.2**

$\|B_n(f;.) - f(.)\|_p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

The proof end, for $n \in \mathbb{N}^+$sufficiently large**,** and ε arbitrary. ■

4 Application

In this part, we give some applications of the sequences of a family of linear positive multivariate neural network operators $B_n(.;\mathbf{x})$, then we analyze the results of these sequences with the sequences of a family of linear positive multivariate neural network operators $F_n^s(.;\mathbf{x})$ and $K_n(.;\mathbf{x})$.

We describe the results by graphics of absolute value of the error function for some particular values of $n = 10, 20, 30$ and for two test functions in 2-dimensional $f(x_1, x_2) = x_1^2 - x_2^2$ and $g(x_1, x_2) = \frac{2}{3}\cos(4x_1x_2) + 2\sin(x_1 + x_2)$, we define the error function between any test function h and operator A_n as follows: $E(\mathbf{x}) = |A_n(\mathbf{x}) - h(\mathbf{x})|$, $\forall \mathbf{x} \in \mathbb{R}^s$.

Example 4.1

For $n = 10, 20, 30$ the sequence of a family of linear positive multivariate neural network operators $B_n(f;\mathbf{x})$ convergence to the test function $f(x_1, x_2) = x_1^2 - x_2^2$, with error given in the Figs. (4.1-4.3) respectively.

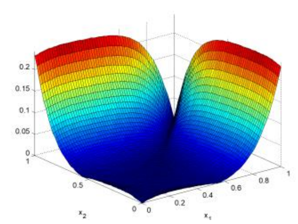

Fig. 4.1. The error function $|B_n(f;\mathrm{x}) - f(\mathrm{x})|$, as $n = 10$. CPU time is 17.706587 seconds

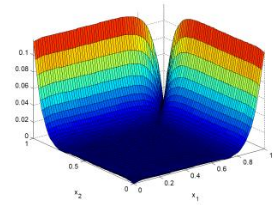

Fig. 4.2. The error function$|B_n(f;\mathrm{x}) - f(\mathrm{x})|$, as $n = 2\ 0$ CPU time is 55.197249 seconds

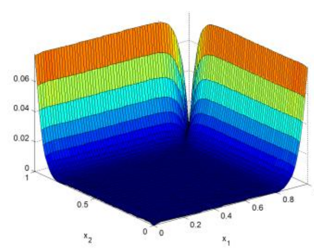

Fig. 4.3. The error function$|B_n(f;\mathrm{x}) - f(\mathrm{x})|$, as $n = 3\ 0$ CPU time is 215.532575 seconds

Example 4.2

For $n = 10, 20, 30$, the sequence of a family of linear positive multivariate neural network operators $F_n^s(f;\mathbf{x})$ convergence to the test function $f(x_1, x_2) = x_1^2 - x_2^2$, with error given in the Figs. (4.4-4.6) respectively.

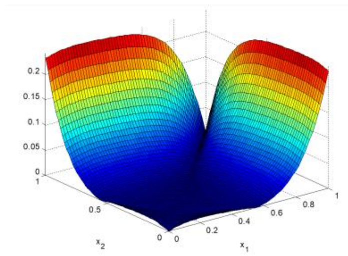

**Fig. 4.4. The error function $|\boldsymbol{F_n^s(f;\mathrm{x}) - f(\mathrm{x})}|$, as $\boldsymbol{n = 10}$.
CPU time is 12.730782 seconds**

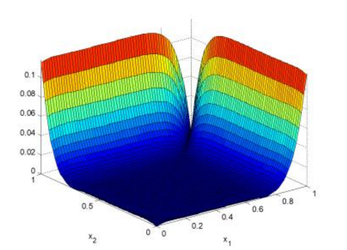

**Fig. 4.5. The error function$|\boldsymbol{F_n^s(f;\mathrm{x}) - f(\mathrm{x})}|$,as $\boldsymbol{n = 20}$
CPU time is 93.307534 seconds**

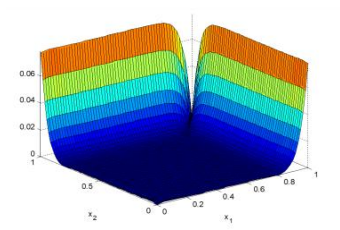

**Fig. 4.6. The error function$|\boldsymbol{F_n^s(f;\mathrm{x}) - f(\mathrm{x})}|$, as $\boldsymbol{n = 30}$
CPU time is 227.053357 seconds**

Example 4.3

For $n = 10, 20, 30$, the sequence of a family of linear positive multivariate neural network operators $K_n(f;\mathbf{x})$, convergence to the test function $f(x_1, x_2) = x_1^2 - x_2^2$, with error given in the Figs. (4.7-4.9) respectively.

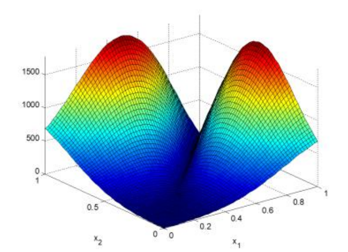

**Fig. 4.7. The error function$|\boldsymbol{K_n(f;\mathrm{x}) - f(\mathrm{x})}|$, as $\boldsymbol{n = 10}$.
CPU time is 10.798002 seconds**

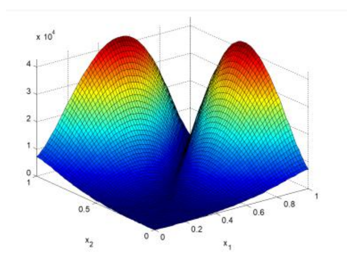

**Fig. 4.8. The error function$|\boldsymbol{K_n(f;\mathrm{x}) - f(\mathrm{x})}|$, as $\boldsymbol{n = 20}$
CPU time is 46.376399 seconds**

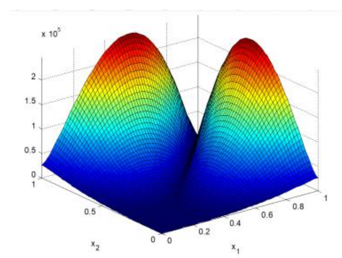

**Fig. 4.9. The error function$|\boldsymbol{K_n(f;\mathrm{x}) - f(\mathrm{x})}|$, as $\boldsymbol{n = 30}$
CPU time is 211.845499 seconds**

Example 4.4

For $n = 10, 20, 30$ the sequence of a family of linear positive multivariate neural network operators $B_n(g;\mathbf{x})$ convergence to the test function $g(x_1, x_2) = \frac{2}{3}\cos(4x_1x_2) + 2\sin(x_1 + x_2)$, with error given in the Figs. (4.10-4.12) respectively.

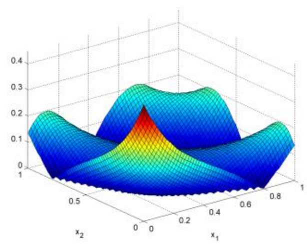

Fig. 4.10. The error function$|B_n(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 10$.
CPU time is 20.416746 seconds

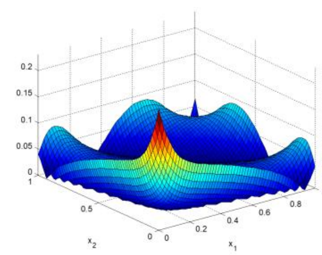

Fig. 4.11. The error function$|B_n(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 20$
CPU time is 101.361674 seconds

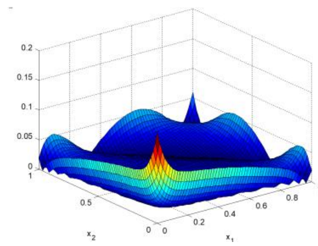

Fig. 4.12. The error function$|B_n(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 30$
CPU time is 445.471572 seconds

Example 4.5

For $n = 10, 20, 30$, the sequence of a family of linear positive multivariate neural network operators $F_n^s(g;\mathbf{x})$ convergence to the test function $g(x_1, x_2) = \frac{2}{3}\cos(4x_1x_2) + 2\sin(x_1 + x_2)$,with error given in the Figs. (4.13-4-15) respectively.

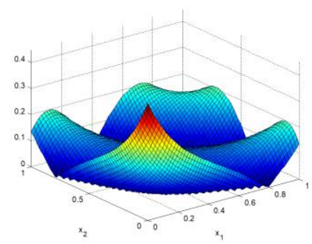

Fig. 4.13. The error function$|F_n^s(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 10$.
CPU time is 9.581222 seconds

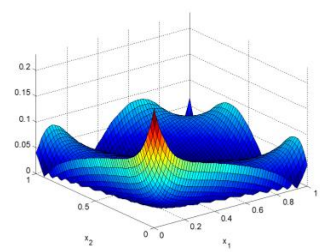

Fig. 4.14. The error function$|F_n^s(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 20$
CPU time is 51.313056 seconds

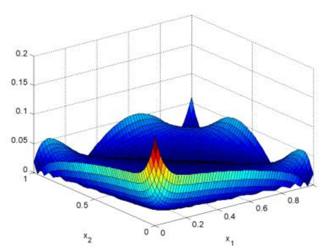

Fig. 4.15. The error function$|F_n^s(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 30$
CPU time is 272.267879 seconds

Example 4.6

For $n = 10, 20, 30$, the sequence of a family of linear positive multivariate neural network operators $K_n(g;\mathbf{x})$, convergence to the test function $g(x_1, x_2) = \frac{2}{3}\cos(4x_1x_2) + 2\sin(x_1 + x_2)$, with error given in the Figs. (4.16-4.18) respectively.

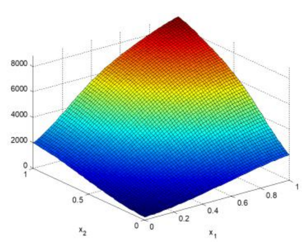

Fig. 4.16. The error function$|K_n(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 10$.
CPU time is 8.827533 seconds

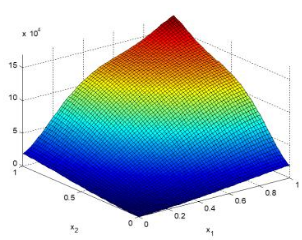

Fig. 4.17. The error function$|K_n(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 20$
CPU time is 88.360289 seconds

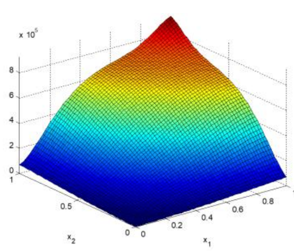

Fig. 4.18. The error function$|K_n(g;\mathrm{x}) - g(\mathrm{x})|$,as $n = 30$
CPU time is 373.654205 seconds

The results of applications previous obtained are shown in table next.

Table 4.1. Shows us when n change 10, 20, 30 the CPU time values and the maximum error in figures

Operators	n	CPU time values	Maximum error in figures (approximately)	Figure
$B_n(f;\mathbf{x})$	10	17.706587 s	0.2	4.1
	20	55.197249 s	0.1	4.2
	30	215.532575 s	0.06	4.3
$F_n^s(f;\mathbf{x})$	10	12.730782 s	0.2	4.4
	20	93.307534 s	0.1	4.5
	30	227.053357 s	0.06	4.6
$K_n(f;\mathbf{x})$	10	10.798002 s	1500	4.7
	20	46.376399 s	4×10^4	4.8
	30	211.845499 s	2×10^5	4.9
$B_n(g;\mathbf{x})$	10	20.416746 s	0.2	4.10
	20	101.361674 s	0.1	4.11
	30	445.471572 s	0.05	4.12
$F_n^s(g;\mathbf{x})$	10	9.581222 s	0.2	4.13
	20	51.313056 s	0.1	4.14
	30	272.267879 s	0.05	4.15
$K_n(g;\mathbf{x})$	10	8.827533 s	2000	4.16
	20	88.360289 s	15×10^4	4.17
	30	373.654205 s	8×10^5	4.18

5 Conclusions

We give some application of the sequences of linear positive multivariate neural network operators $B_n(.;\mathbf{x})$, $F_n^s(.;\mathbf{x})$ and $K_n(.;\mathbf{x})$. We take two test functions in 2- dimensional $f(x_1,x_2)$ and $g(x_1,x_2)$, The graphics of absolute value of the error function for some particular values of $n = 10, 20, 30$, we get the following results.

The best convergence sequence in $B_n(.;\mathbf{x})$ for the two test function f and g also when $n = 10, 20, 30$. On the other hand the worst result we got in $K_n(.;\mathbf{x})$ when $n = 10, 20, 30$, for the two test function f and g, due to the large proceedings required for these effects.

Note that these result for these applications were obtained when $n = 30$ or less due to slow computer or inability to implement higher values to n. These results can not be relied on for n values greater than 30 after application and observation of results.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Bernstein SN. Démonstration du théoréme de Weierstrass fondée sur le calcul de probabilities. Comm. Soc. Math. Kharkow. 1912/13;13:1-2.

[2] Durrmeyer JL. Une formula d ‘inversion de la transformée de laplace: Applications á la théorie des moment. Thése de 3e Cycle, Faculte des Sciences de I‘Unversite de Paris; 1967.

[3] Makovoz Y. Random approximants and neural networks. J. Approx. Theory. 1996;85:98-109.

[4] Makovoz Y. Uniform approximation by neural networks. J. Approx. Theory. 1998;95(2):215-228.

[5] Costarelli D, Spigler R. Approximation by series of sigmoidal functions with applications to neural networks. Ann. Mat. Appl.; 2013. (in press)
Available:http://dx.doi.org/10.1007/s10231-013-0378-y

[6] Cybenko G. Approximation by superpositions of a sigmoidal function. Math. Control Signals Systems. 1989;2:303-314.

[7] Lewicki G, Marino G. Approximation by superpositions of a sigmoidal function. Z. Anal. Anwend. 2003;22(2):463-470.

[8] Mhaskar HN, Micchelli CA. Approximation by superposition of sigmoidal and radial basis functions. Adv. Appl. Math. 1992;13:350-373.

[9] Kainen PC, Kurková V. An integral upper bound for neural network approximation. Neural Comput. 2009;21:2970-2989.

[10] Gnecco G, Sanguineti M. On a variational norm tailored to variable-basis approximation schemes. IEEE Trans. Inform. Theory. 2011;57:549-558.

[11] Costarelli D, Spigler R. Constructive approximation by superposition of sigmoidal functions. Anal. Theory Appl. 2013;29(2):169-196.

[12] Costarelli D, Spigler R. Approximation results for neural network operators activated by sigmoidal functions. Neural Netw. 2013;44:101-106.

[13] D. Costarelli, Spigler R. Multivariate neural network operators with sigmoidal activation functions. Neural Netw. 2013;48:72-77.

[14] Costarelli D, Spigler R. Convergence of a family of neural network operators of the Kantorovich-type. Journal of Approximation Theory. 2014;185:80-90.

[15] Zhaowei Cai, Quanfu Fan, Rogerio Feris, Nuno Vasconcelos, UC San Diego, IBM Watson Research. A unified multi-scale deep convolutional neural network for fast object detection. ECCV 2016: Computer Vision – ECCV. 2016;354-370. Cite as.

[16] Cong Wang, Hongli Zhang, Wenhui Fan, Xiaochao Fan. A new wind power prediction method based on chaotic theory and Bernstein Neural Network; Energy. 2016;117:259-271.

[17] Jafari R, Yu W, Li X. Solving fuzzy differential equation with bernstein neural networks 2016 IEEE International Conference on Systems, Man, and Cybernetics • SMC 2016.| October 9-12, 2016 • Budapest, Hungary.

[18] Shi W, Caballero J, Husz´ar F, Totz J, Aitken AP, Bishop R, Rueckert D, Wang Z, Pony M. Technology and I. College London. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network; CVPR 2016 paper.

[19] Costarelli D, Spigler R. Sigmoidal functions approximation and applications, Ph.D. thesis, UNIVERSITA̓ DEGLI STUDI; 2013.