Communications in Mathematics and Applications
Vol. 13, No. 2, pp. 585–594, 2022
ISSN 0975-8607 (online); 0976-5905 (print)
Published by RGN Publications
DOI: 10.26713/cma.v13i2.1735

http://www.rgnpublications.com

Research Article

Neural Network of Multivariate Square Rational Bernstein Operators

Ibtihal J. Mohammad* and Ali J. Mohammad

Department of Mathematics, University of Basrah, Basrah, Iraq
***Corresponding author:** pgs2206@uobasrah.edu.iq

Received: November 12, 2021 **Accepted:** April 20, 2022

Abstract. This paper introduced a family of neural networks of multivariate square rational Bernstein operators defined by extending the artificial neural networks multivariate Bernstein by using square Bernstein polynomials and studied the behavior of this neural network. Also, gave application through some numerical examples.

Keywords. Multivariate neural network operators, Activation functions, Pointwise approximation theorems, Uniform approximation theorems, Space of Lipschitz classes on $\mathcal{R}$

Mathematics Subject Classification (2020). 41A25, 41A30, 47A58

1. Introduction

The problems in the approximation theory related to single-layer neural networks are discussed by Pinkus in 1999 [19] for an activation function $\sigma:\mathbb{R}\to\mathbb{R}$ which is expressed by the following formula:

$$N_n(\boldsymbol{x})=\sum_{i=0}^{n} c_i\sigma(\boldsymbol{a}_i\cdot\boldsymbol{x}+\theta_i), \quad n\in\mathbb{N}^+, \tag{1.1}$$

where $\boldsymbol{x}\in\mathbb{R}^s$, $s\in\mathbb{N}^+$, $0\le i\le n$, θ_i, $c_i\in\mathbb{R}$, $\boldsymbol{a}_i,\boldsymbol{a}_i\cdot\boldsymbol{x}\in\mathbb{R}^s$.

The symbols in eq. (1.1) θ_i, c_i, $\boldsymbol{a}_i$ and $\boldsymbol{a}_i\cdot\boldsymbol{x}$ denote to be threshold values, coefficients, weights, and the inner product, respectively.

Many papers are published in this branch, we refer here to some of them [2, 3, 11–16] and [17].

For $f : \mathcal{R} \to \mathbb{R}$ be a bounded function and $\boldsymbol{x}$ in $\mathcal{R} := [a_1, b_1] \times \cdots \times [a_s, b_s] \subset \mathbb{R}^s$, Costarelli and Spigler [4, 5] are introduced and studied the behavior of artificial neural networks in the case of the univariate and the multivariate Bernstein, given as:

$$F_n(f;\boldsymbol{x}) = \frac{\sum\limits_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor} \cdots \sum\limits_{k_s=\lceil na_s\rceil}^{\lfloor nb_s\rfloor} f\left(\frac{\boldsymbol{k}}{n}\right) \Psi_\sigma(n\boldsymbol{x}-\boldsymbol{k})}{\sum\limits_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor} \cdots \sum\limits_{k_s=\lceil na_s\rceil}^{\lfloor nb_s\rfloor} \Psi_\sigma(n\boldsymbol{x}-\boldsymbol{k})}, \quad n \in \mathbb{N}^+, \tag{1.2}$$

where Ψ_σ is a density function that is built from a sigmoidal function $\sigma, \boldsymbol{k} = (k_1, \dots, k_s) \in \mathbb{Z}^+$. As usual, the symbols $\lfloor\cdot\rfloor$, $\lceil\cdot\rceil$ denote taking the "floor" and the "ceiling" of a given number, respectively.

Costarelli and Spigler [6] used the structure of Kantorovich to the multivariate NN operators for eq. (1.2) and studied the approximation theorems to this new NN.

Costarelli and Vinti [7] introduced a neural network by using max-product and studied approximation theorems also estimates the rate of convergence to the multivariate max-product NN operators and the multivariate quasi-interpolation max-product NN operators.

Gavrea and Ivan [9] defined the square of Bernstein polynomials which is given as:

$$B_{n,2}(f;x) = \frac{\sum\limits_{k=0}^{n} b_{n,k}^2(x) f\left(\frac{k}{n}\right)}{\sum\limits_{k=0}^{n} b_{n,k}^2(x)}, \quad n = 1, 2, \dots, \tag{1.3}$$

where $b_{n,k}^2(x) = (b_{n,k}(x))^2$, $x \in [0,1]$, $f \in C[0,1]$.

Mohammad and Mohammad [18] defined the neural network of type summation-integral Bernstein operators by using eq. (1.2), then studied pointwise and uniform approximation theorems for this neural network.

Hassan in 2018 introduce the new modified of Bernstein operators define in [10].

Bajpeyi and Kumar [1] introduced and studied a neural network of exponential type and studied its behavior in one- and multi-dimensional cases.

Costarelli *et al.* [8] have used the neural network in eq. (1.2) to introduce and study the multivariate max-product NN of Kantorovich type.

This paper extends the neural network in eq. (1.2) by using the square Bernstein polynomials in eq. (1.3) and studies the behavior of the family of the neural network of multivariate square rational Bernstein operators acting on the sigmoidal functions σ. Finally, gives two numerical examples for the NN operators $Q_n(\cdot;x,y)$ and the NN operators $F_n(\cdot;x,y)$ are applying for two test functions, it turns out from the figures and numerical results of the table in both examples that the NN operators $Q_n(\cdot;x,y)$ is better than the NN operators $F_n(\cdot;x,y)$.

2. Preliminary Results

Several preliminary results are recalled in this section.

The measurable functions like the Logistic function $\sigma_l(x) = (1+e^{-x})^{-1}$, Hyperbolic tangent $\sigma_h(x) = \frac{1}{2}[\tanh(x)-1]$, is called a sigmoidal function if satisfying $\lim\limits_{x\to-\infty} \sigma(x) = 0$ and $\lim\limits_{x\to+\infty} \sigma(x) = 1$. Also, the function $\Phi_\sigma(x)$is defined as

$$\Phi_\sigma(x) = \frac{1}{2}[\sigma(x+1)-\sigma(x-1)], \quad x \in \mathbb{R},$$

for every non-decreasing function σ satisfying assumptions ($\Sigma1$), ($\Sigma2$) and ($\Sigma3$) in [5].

(i) the odd function g_σ, such that $g_\sigma(x) = \sigma(x) - 1/2$;

(ii) the concave function σ is a function for $\sigma \in C^2(\mathbb{R})$, $x \geq 0$;

(iii) for some $\alpha > 0$, the function σ satisfying $\sigma(x) = \mathcal{O}(|x|^{-1-\alpha})$ as $x \to -\infty$.

We will give some definitions that we need:

Definition 2.1 ([5])**.** Any measurable function with the condition $\lim\limits_{x\to-\infty} \zeta(x) = 0$, $\lim\limits_{x\to+\infty} \zeta(x) = 1$, it's known as a sigmoidal function.

Definition 2.2 ([5])**.** Lipschitz classes are defined as:

$$\begin{aligned}\mathrm{Lip}(v) = \{f \in C^0(\mathcal{R}) : \exists\, \gamma > 0,\ M > 0 \text{ so that } \forall\ \boldsymbol{x} \in \mathcal{R}, |f(\boldsymbol{x}+\boldsymbol{y}) - f(\boldsymbol{x})| \leq M\|\boldsymbol{y}\|_2^v, \\ \forall\ \|\boldsymbol{y}\|_2 \leq \gamma \text{ with } (\boldsymbol{x}+\boldsymbol{y}) \in \mathcal{R},\ 0 < v \leq 1\}.\end{aligned}$$

3. Auxiliary Results

The multivariate NN operators $Q_n(f;\boldsymbol{x})$ is defined and studied follow:

Definition 3.1. For a bounded and continuous function $f : \mathcal{R} \to \mathbb{R}$, the linear positive multivariate (NN) operators of the multivariate square rational Bernstein operators of f, $Q_n(f;\boldsymbol{x})$ activated by the sigmoidal function σ acting on f, is defined by:

$$Q_n(f;\boldsymbol{x}) = \frac{\sum\limits_{\boldsymbol{k}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k}) f(\boldsymbol{k}/n)}{\sum\limits_{\boldsymbol{k}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})},$$

$$\sum_{\boldsymbol{k}} = \sum_{k_1=\lceil na_1\rceil}^{\lfloor nb_1\rfloor} \cdots \sum_{k_s=\lceil na_s\rceil}^{\lfloor nb_s\rfloor},$$

where the multivariate for the Φ_σ^2 define a function $\Psi_\sigma^2(\boldsymbol{x}) = \Phi_\sigma^2(x_1)\cdot\Phi_\sigma^2(x_2)\cdot\ldots\cdot\Phi_\sigma^2(x_s)$, observe that $Q_n(1;\boldsymbol{x}) = 1$, for every $\boldsymbol{x} \in \mathcal{R}$ and n tends to infinity.

Definition 3.2. For $v > 0$, the discrete absolutely moment of the function Φ_σ^2 of order v is defined as

$$m_v(\Phi_\sigma^2) = \sup_{x\in\mathbb{R}} \sum_{k\in\mathbb{Z}} \Phi_\sigma^2(x-k)|x-k|^v.$$

The properties of the functions Φ_σ and Ψ_σ in [4] and [5] are needed to give and prove the following Lammas 3.1-3.3 directly.

Lemma 3.1. *For the function $\Phi^2_\sigma(x)$, one has:*

(i) $\Phi^2_\sigma(x) \geq 0$, $\forall\ x \in \mathbb{R}$ *and* $\lim\limits_{x\to\pm\infty} \Phi^2_\sigma(x) = 0$, *as well* $\Phi^2_\sigma(1) > 0$;

(ii) *the function* $\Phi^2_\sigma(x)$ *is even;*

(iii) $\sum\limits_{k\in\mathbb{Z}} \Phi^2_\sigma(x-k) \simeq 0.156517$, $\forall\ x \in \mathbb{R}$;

(iv) *the series* $\sum\limits_{k\in\mathbb{Z}} \Phi^2_\sigma(x-k)$ *on the compact subset of* $\mathbb{R}$*is uniformly converged;*

(v) $\Phi^2_\sigma(x) = \mathcal{O}(|x|^{-2(1+\alpha)})$ *as* $x \to \pm\infty$.

Proof. One can easily prove this lemma by direct computation and the prove of properties the function Φ_σ in [4]. □

The next lemma gives some properties for the function $\Psi^2_\sigma(\boldsymbol{x}-\boldsymbol{k})$.

Lemma 3.2. *For the function $\Psi^2_\sigma(\boldsymbol{x}-\boldsymbol{k})$, one has:*

(i) $\sum\limits_{\boldsymbol{k}} \Psi^2_\sigma(\boldsymbol{x}-\boldsymbol{k}) \simeq (0.156517)^s$, *for all* $\boldsymbol{x} \in \mathbb{R}^s$;

(ii) *the series* $\sum\limits_{\boldsymbol{k}} \Psi^2_\sigma(\boldsymbol{x}-\boldsymbol{k})$ *on the compact subset of* $\mathbb{R}^s$ *are uniformly converged;*

(iii) $\lim\limits_{n\to\infty} \sum\limits_{\|\boldsymbol{x}-\boldsymbol{k}\|>\gamma n} \Psi^2_\sigma(\boldsymbol{x}-\boldsymbol{k}) = 0$ *are converges uniformly to* $\boldsymbol{x} \in \mathbb{R}^s$; *and*
$\sum\limits_{\|\boldsymbol{x}-\boldsymbol{k}\|>\gamma n} \Psi^2_\sigma(\boldsymbol{x}-\boldsymbol{k}) = \mathcal{O}(n^{-v})$ *in particularly for* $0 < v < \alpha$, *where* $\gamma, \alpha > 0$, α *is a constant and* $\|\boldsymbol{x}\|_\infty = \max\{|x_i|, i = 1,\dots,s\}$.

Proof. One can easily prove this lemma by direct computation and the prove of properties the function Ψ_σ in [5]. □

Lemma 3.3. (i) *For $x \in [a,b] \subset \mathbb{R}$, then*

$$\frac{1}{\sum\limits_{k=\lceil na\rceil}^{\lfloor nb\rfloor} \Phi^2_\sigma(nx-k)} \leq \frac{1}{\Phi^2_\sigma(1)};$$

(ii) *for $\boldsymbol{x} \in \mathcal{R}$ then*

$$\frac{1}{\prod\limits_{i=1}^{s} \sum\limits_{k_i=\lceil na_i\rceil}^{\lfloor nb_i\rfloor} \Phi^2_\sigma(nx_i-k_i)} \leq \frac{1}{[\Phi^2_\sigma(1)]^s}.$$

Proof. One can easily prove this lemma by direct computation and using the prove of Lemma 2.7 in [5]. □

The following theorem studies the pointwise and the uniform convergence for the NN, $Q_n(f;\boldsymbol{x})$.

Theorem 3.1. *For $f:\mathcal{R}\to\mathbb{R}$bounded and continuous function,*

$$\lim_{n\to\infty} Q_n(f;\boldsymbol{x}) = f(\boldsymbol{x}),$$

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

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

Proof. Suppose $\boldsymbol{x}\in\mathcal{R}$ is a point of continuity of f we have

$$|Q_n(f;\boldsymbol{x})-f(\boldsymbol{x})| = \left|\frac{\sum_{\boldsymbol{k}}\Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})f(\boldsymbol{k}/n)}{\sum_{\boldsymbol{k}}\Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})} - f(\boldsymbol{x})\right|$$

and by using Lemma 3.3, we get

$$|Q_n(f;\boldsymbol{x})-f(\boldsymbol{x})| \le \frac{1}{[\Phi_\sigma^2(1)]^s}\sum_{\boldsymbol{k}}\Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})|f(\boldsymbol{k}/n)-f(\boldsymbol{x})|$$

$\forall\ n\to\infty$, $n\in N^+$, $\boldsymbol{x}\in\mathbb{R}^s$ are arbitrary but fixed. Suppose for a fixed $\varepsilon>0$, and from the continuity of f at $\boldsymbol{x}$, $\exists\ \gamma>0$: $|f(\boldsymbol{y})-f(\boldsymbol{x})|<\varepsilon$, $\forall\ \boldsymbol{y}\in\mathcal{R}$ with $\|\boldsymbol{y}-\boldsymbol{x}\|_2<\gamma$, the symbol $\|\cdot\|_2$ denote to Euclidean norm.

Now,one gets

$$\begin{aligned}|Q_n(f;\boldsymbol{x})-f(\boldsymbol{x})| &\le \frac{1}{[\Phi_\sigma^2(1)]^s}\Bigg\{\sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\|<\frac{\gamma}{\sqrt{s}}}\Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})|f(\boldsymbol{k}/n)-f(\boldsymbol{x})| \\ &\qquad + \sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\|\ge\frac{\gamma}{\sqrt{s}}}\Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})|f(\boldsymbol{k}/n)-f(\boldsymbol{x})|\Bigg\} \\ &:= \frac{1}{[\Phi_\sigma^2(1)]^s}(I_1+I_2).\end{aligned}$$

Now using the continuity of f and Lemma 3.2, we get that $\|\boldsymbol{k}/n-\boldsymbol{x}\|_2\le\sqrt{s}\|\boldsymbol{k}/n-\boldsymbol{x}\|\le\gamma$. So estimation I_1 is,

$$I_1<\varepsilon\sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\|\le\frac{\gamma}{\sqrt{s}}}\Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})\le\varepsilon.$$

From the boundedness of f and Lemma 3.2, for sufficiently large n, we have

$$I_2\le 2\|f\|_\infty\sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\|>\frac{\gamma}{\sqrt{s}}}\Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})<2\|f\|_\infty\varepsilon,$$

uniformly $\forall\ \boldsymbol{x}\in\mathbb{R}^s$. The first direction of the theorem holds because ε arbitrarily. When $f\in C^0(\mathcal{R})$, the proof of the other direction is readily followed in the same way by exchange $\gamma>0$ with the parameter of the uniform continuity of f on $\mathcal{R}$. □

Now, in the following, the order of approximation of (NN) operators in $C^0(\mathcal{R})$ is studied.

Theorem 3.2. *Suppose $f\in\mathrm{Lip}(v)$ for some $0<v\le 1$, and let the sigmoidal function σ satisfy the condition $(\Sigma 3)$ in [5] for some $\alpha>1$. Then,*

$$\|Q_n(f;\boldsymbol{x})-f(\cdot)\|_\infty = \mathcal{O}(n^{-v})\ \ as\ \ n\to\infty.$$

Proof. Let $f \in \mathrm{Lip}(v)$, $\forall\ \boldsymbol{x} \in \mathcal{R}$, for some $v \in (0,1]$, by using Lemma 3.3, one obtains

$$|Q_n(f;\boldsymbol{x}) - f(\boldsymbol{x})| \le \frac{1}{[\Phi_\sigma^2(1)]^s} \sum_{\boldsymbol{k}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})|f(\boldsymbol{k}/n) - f(\boldsymbol{x})|.$$

Now by using the definition of $\mathrm{Lip}(v)$, where γ, $C > 0$ are constants relative to f one obtains

$$\begin{aligned}|Q_n(f;\boldsymbol{x}) - f(\boldsymbol{x})| &\le \frac{1}{[\Phi_\sigma^2(1)]^s}\left\{ \sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\| \le \frac{\gamma}{\sqrt{s}}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})|f(\boldsymbol{k}/n) - f(\boldsymbol{x})| \right.\\ &\qquad \left. + \sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\| > \frac{\gamma}{\sqrt{s}}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})|f(\boldsymbol{k}/n) - f(\boldsymbol{x})| \right\}\\ &:= \frac{1}{[\Phi_\sigma^2(1)]^s}(J_1 + J_2).\end{aligned}$$

Since $f \in \mathrm{Lip}(v)$, we get for $\|\boldsymbol{k}/n - \boldsymbol{x}\|_2 \le \sqrt{s}\|\boldsymbol{k}/n - \boldsymbol{x}\| \le \gamma$, and hence $|f(\boldsymbol{k}/n) - f(\boldsymbol{x})| < C\|\boldsymbol{k}/n - \boldsymbol{x}\|_2^v \le Cs^{\frac{v}{2}}\|\boldsymbol{k}/n - \boldsymbol{x}\|^v$.

$$\begin{aligned}J_1 &\le n^{-v} C s^{v/2} \sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\| \le \frac{\gamma}{\sqrt{s}}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})\|n\boldsymbol{x}-\boldsymbol{k}\|^v\\ &\le n^{-v} C s^{v/2} \sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\| \le \frac{\gamma}{\sqrt{s}}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})\|n\boldsymbol{x}-\boldsymbol{k}\|^v\end{aligned}$$

for fixed $0 < v_i < \alpha$, by using Lemma 3.2, for a compact subset $K \subset \mathbb{R}^s$. $\forall\ \boldsymbol{x} \in \mathbb{R}^s$, if $n \to \infty$ implies the following:

$$\begin{aligned}J_1 &\le n^{-v} C s^{v/2} \sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\| \le \frac{\gamma}{\sqrt{s}}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k})\|n\boldsymbol{x}-\boldsymbol{k}\|^v\\ &\le n^{-v} C s^{v/2} \sum_{j=1}^{s} \left\{ \sum_{k_j \in \mathbb{Z}} \Phi_\sigma^2(nx_j - k_j)|nx_j - k_j|^v \left[\sum_{\boldsymbol{k}_{[j]} \in \mathbb{Z}^{s-1}} \Psi_\sigma^{2[j]}(n\boldsymbol{x}_{[j]} - \boldsymbol{k}_{[j]}) \right] \right\},\end{aligned}$$

where

$$\Psi_\sigma^{2[j]}(n\boldsymbol{x}_{[j]} - \boldsymbol{k}_{[j]}) = \Phi_\sigma^2(nx_1 - k_1) \cdot \ldots \cdot \Phi_\sigma^2(nx_{j-1} - k_{j-1}) \cdot \Phi_\sigma^2(nx_{j+1} - k_{j+1}) \cdot \ldots \cdot \Phi_\sigma^2(nx_s - k_s).$$

Notice that, for every $j = 1,\ldots,s$, $\boldsymbol{x}_{[j]} = (x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_s) \in \mathbb{R}^{s-1}$, $\boldsymbol{k}_{[j]} = (k_1,\ldots,k_{j-1},k_{j+1},\ldots,k_s) \in \mathbb{Z}^{s-1}$. Now, let $k_{[j]} \subset \mathbb{R}$ the set of j-th projection of a compact set K for all elements. By using Lemma 3.2 and for all sufficiently large $N \in \mathbb{N}^+$, then

$$\begin{aligned}J_1 &\le (0.156517)^{s-1} n^{-v} C s^{v/2} \sum_{j=1}^{s} \left\{ \sum_{k_j \in \mathbb{Z}} \Phi_\sigma^2(nx_j - k_j)|nx_j - k_j|^v \right\}\\ &\le (0.156517)^{s-1} n^{-v} C s^{1+v/2} m_v(\Phi_\sigma^2).\end{aligned}$$

Note that $m_v(\Phi_\sigma^2) < \infty$, where $m_v(\Phi_\sigma^2)$ give in Definition 3.2 since $v < \alpha$, therefore

$$J_1 = \mathcal{O}(n^{-v}), \quad n \to \infty.$$

Now, the estimation of $J_2$is done by using the other direction of Lemma 3.2, i.e.

$$J_2 \le 2\|f\|_\infty \sum_{\|\boldsymbol{k}/n-\boldsymbol{x}\| > \frac{\gamma}{\sqrt{s}}} \Psi_\sigma^2(n\boldsymbol{x}-\boldsymbol{k}) = \mathcal{O}(n^{-v}), \quad \text{as } n \to \infty. \qquad \square$$

Theorem 3.3. *Let the function* σ *for some* $\alpha \in (0,1]$ *satisfy the condition* $(\Sigma 3)$ *in* [5], *and let* $f \in \mathrm{Lip}(v)$ *for some* $v \in (0,1]$. *Then,*

(i) $\|Q_n(f;\cdot) - f(\cdot)\|_\infty = \mathcal{O}(n^{-v})$, *as* $n \to \infty$, *if* $v < \alpha$.

(ii) $\|Q_n(f;\cdot) - f(\cdot)\|_\infty = \mathcal{O}(n^{-\alpha+\varepsilon})$, *as* $n \to \infty$, $\forall\ 0 < \varepsilon < \alpha$, *if* $\alpha \le v < 1$.

Proof. (i) Using the same step of Theorem 3.2 one obtains proving

$$\|Q_n(f;\cdot) - f(\cdot)\|_\infty = \mathcal{O}(n^{-v}), \quad \text{as } n \to \infty$$

for function $f \in \mathrm{Lip}(v)$ at $0 < v < \alpha$.

(ii) As a special case for all $f \in \mathrm{Lip}(v)$ with $\alpha \le v \le 1$, with ε is fixed but arbitrary choose $\beta := \alpha - \varepsilon$, and get $0 < \beta < \alpha$, by based on part (i), then

$$\|Q_n(f;\cdot) - f(\cdot)\|_\infty = \mathcal{O}(n^{-\beta}) = \mathcal{O}(n^{-\alpha+\varepsilon}), \quad \text{as } n \to \infty$$

for function $f \in \mathrm{Lip}(\beta)$, at $0 < \varepsilon < \alpha$. □

4. Numerical Examples

In this part, two numerical examples for the NN operators $Q_n(\cdot;x,y)$ and the NN operators $F_n(\cdot;x,y)$ are applying for two test functions $f(x,y) = \cos(9xy) + 2\sin(x+y)$ and $g(x,y) = (2x-1)^2 - (2y-1)^2$, $(x,y) \in [0,1] \times [0,1]$ for the values of $n = 10, 30, 60$. The numerical results obtained are described in the figures and compared with the convergence of the two NN. Also, at able of the maximum error function for the two NN is given. It turns out from the figures and numerical results of the table in both examples that the NN operators $Q_n(\cdot;x,y)$ is better than the NN operators $F_n(\cdot;x,y)$.

Example 4.1. For $n = 10, 30, 60$, the convergence of NN operators $Q_n(f;x,y)$, $F_n(f;x,y)$ to test function $f(x,y)$ can be described in Figure 1.

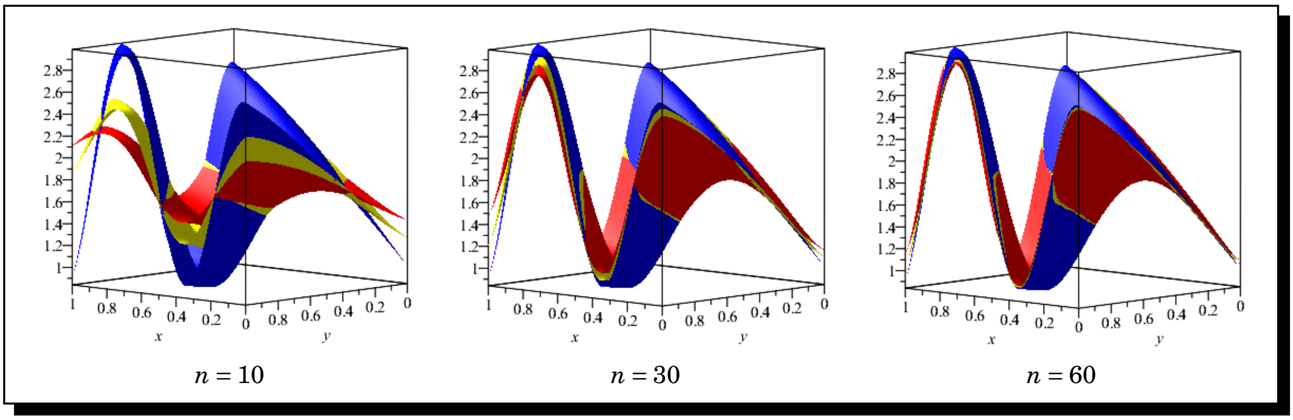

Figure 1. The numerical convergence of NN operators $F_n(f;x,y)$ (red) and $Q_n(f;x,y)$ (yellow) to $f(x,y)$ (blue)

Example 4.2. For $n = 10, 30, 60$, the convergence of NN operators $Q_n(g;x,y)$, $F_n(g;x,y)$ to test function $g(x,y)$ can be described in Figure 2.

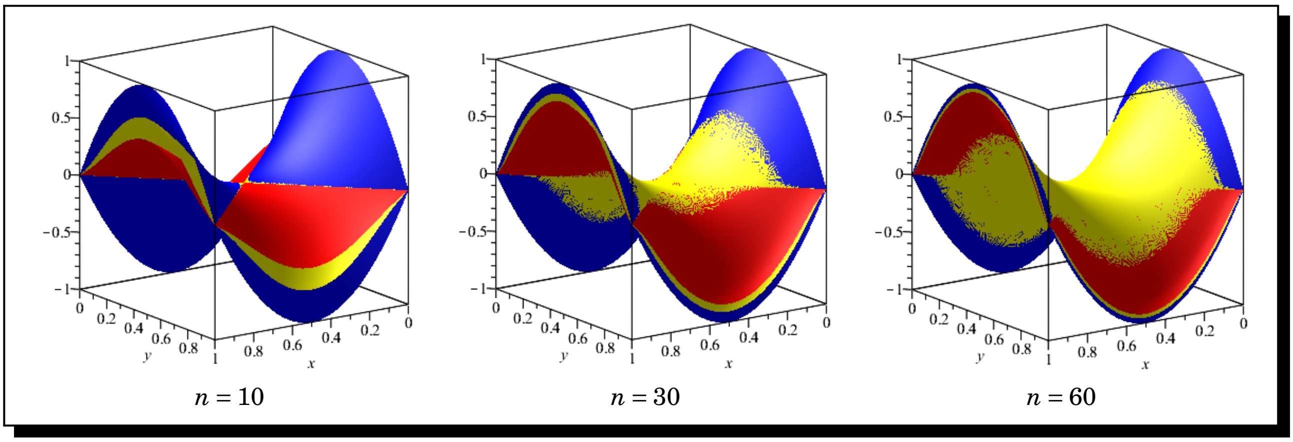

Figure 2. The numerical convergence of NN operators $F_n(g;x,y)$ (red) and $Q_n(g;x,y)$ (yellow) to $g(x,y)$ (blue)

Now, the following table calculation maximum error values between the test function and NN in $\mathbb{R}^2$, by using test functions $f(x,y)$, $g(x,y)$:

Table 1. The maximum error

NN	$n=10$	$n=30$	$n=60$
$F_n(f;x,y)$	0.221650351	0.050239325	0.031124401
$Q_n(f;x,y)$	0.490870074	0.091746325	0.042499216
$F_n(g;x,y)$	$1.082000283\times10^{-10}$	$2.034676091\times10^{-10}$	$4.215129055\times10^{-10}$
$Q_n(g;x,y)$	0.2810197385	0.0888466031	0.0428823462

5. Conclusions

The two numerical examples above and Table 1, are shown that the NN operators $Q_n(\cdot;x,y)$ gives better numerical results with smaller maximum error than the classical NN operators $F_n(\cdot;x,y)$ for the two test functions f and g.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All the authors contributed significantly in writing this article. The authors read and approved the final manuscript.

References

[1] S. Bajpeyi and A.S. Kumar, Approximation by exponential type neural network operators, *Analysis and Mathematical Physics* **11** (2021), Article number: 108, DOI: 10.1007/s13324-021-00543-y.

[2] C.K. Chui and X. Li, Approximation by ridge functions and neural networks with one hidden layer, *Journal of Approximation Theory* **70**(2) (1992), 131 – 141, DOI: 10.1016/0021-9045(92)90081-X.

[3] D. Costarelli and R. Spigler, Approximation by series of sigmoidal functions with applications to neural networks, *Annali di Matematica Pura ed Applicata* **194**(1) (2015), 289 – 306, DOI: 10.1007/s10231-013-0378-y.

[4] D. Costarelli and R. Spigler, Approximation results for neural network operators activated by sigmoidal functions, *Neural Networks* **44** (2013), 101 – 106, DOI: 10.1016/j.neunet.2013.03.015.

[5] D. Costarelli and R. Spigler, Multivariate neural network operators with sigmoidal activation functions, *Neural Networks* **48** (2013), 72 – 77, DOI: 10.1016/j.neunet.2013.07.009.

[6] D. Costarelli and R. Spigler, Convergence of a family of neural network operators of the Kantorovich-type, *Journal of Approximation Theory* **185** (2014), 80 – 90, DOI: 10.1016/j.jat.2014.06.004.

[7] D. Costarelli and G. Vinti, Pointwise and uniform approximation by multivariate neural network operators of the max-product type, *Neural Networks* **81** (2016), 81 – 90, DOI: 10.1016/j.neunet.2016.06.002.

[8] D. Costarelli, A.R. Sambucini and G. Vinti, Convergence in Orlicz spaces by means of the multivariate max-product neural network operators of the Kantorovich type and applications, *Neural Computing and Applications* **31**(9) (2020), 5069 – 5078, DOI: 10.1007/s00521-018-03998-6.

[9] I. Gavrea and M. Ivan, On a new sequence of positive linear operators related to squared Bernstein polynomials, *Positivity* **21** (2017), 911 – 917, DOI: 10.1007/s11117-016-0442-0.

[10] A.K. Hassan, On generalized Szasz-Bernstein–type Operators, *Journal of University of Babylon for Pure and Applied Sciences* **26**(4) (2018), 36 – 44, URL: https://journalofbabylon.com/index.php/JUBPAS/article/view/682.

[11] L.K. Jones, Constructive approximations for neural networks by sigmoidal functions, *Proceedings of the IEEE* **78**(10) (1990), 1586 – 1589, DOI: 10.1109/5.58342.

[12] B. Lenze, Constructive multivariate approximation with sigmoidal functions and applications to neural networks, in: D. Braess and L.L. Schumaker (eds.), *Numerical Methods in Approximation Theory*, Vol. **9**, ISNM 105: International Series of Numerical Mathematics / Internationale Schriftenreihe zur Numerischen Mathematik / Série Internationale d'Analyse Numérique, Vol. **105**, Birkhäuser, Basel (1992), DOI: 10.1007/978-3-0348-8619-2_9.

[13] X. Li, Simultaneous approximations of multivariate functions and their derivatives by neural networks with one hidden layer, *Neurocomputing* **12**(4) (1996), 327 – 343, DOI: 10.1016/0925-2312(95)00070-4.

[14] X. Li and C.A. Micchelli, Approximation by radial basis and networks, *Numerical Algorithms* **25** (2000), 241 – 262, DOI: 10.1023/A:1016685729545.

[15] W. Light, Ridge function, sigmoidal functions and neural networks, in *Approximation Theory* VII (Austin, TX, 1992) (1993) 163 – 206.

[16] Y. Makovoz, Uniform approximation by neural networks, *Journal of Approximation Theory* **95**(2) (1998), 215 – 228, DOI: 10.1006/jath.1997.3217.

[17] H.N. Mhaskar and C.A. Micchelli, Degree of approximation by neural and translation networks with a single hidden layer, *Advances in Applied Mathematics* **16**(2) (1995), 151 – 183, DOI: 10.1006/aama.1995.1008.

[18] A.J. Mohammad and I.J. Mohammad, Summation-integral Bernstein type of neural network operators, *Asian Journal of Mathematics and Computer Research* **21**(2) (2017), 74 – 86, URL: https://ikprress.org/index.php/AJOMCOR/article/view/1120.

[19] A. Pinkus, Approximation theory of the MLP model in neural networks, *Acta Numerica* **8** (1999), 143 – 195, DOI: 10.1017/S0962492900002919.