Republic of Iraq

Ministry of Higher Education

and Scientific Research

University of Babylon

College of Education for pure Sciences

Department of Mathematics

Neural Approximation in Orlicz Spaces

Research

Submitted to College of Education for Pure

Sciences of University of Babylon as Partial Fulfillment of the

Requirement for the Degree of

Higher Diploma Education / Mathematics

by

Ikram Abed Mezher Mukheef AL_Tamimi

Supervised by

Ass.Prof.Dr. Hawraa Abbas Fadhil Almurieb

2023 A.D *1445 A.H*

بسم الله الرحمن الرحيم

ٱقۡرَأۡ بِٱسۡمِ رَبِّكَ ٱلَّذِي خَلَقَ ١ خَلَقَ ٱلۡإِنسَٰنَ مِنۡ عَلَقٍ ٢ ٱقۡرَأۡ وَرَبُّكَ ٱلۡأَكۡرَمُ ٣ ٱلَّذِي
عَلَّمَ بِٱلۡقَلَمِ ٤ عَلَّمَ ٱلۡإِنسَٰنَ مَا لَمۡ يَعۡلَمۡ ٥

صدق الله العلي العظيم

من سورة العلق

1-5

Supervisor Certification

I certify that this research entitled **" Neural Approximation in Orlicz Spaces"** was prepared by the student " **Ikram Abed Mezher mukheef "** under my supervision at University of Babylon, College of Education for Pure Sciences, Department of Mathematics as a partial fulfillment of the requirement the Diploma Education / Mathematics.

Signature:
Name: Dr. Hawraa Abbas Fadhil Mohammed Almurieb
Scientific grade: Assistant Professor
Date: / / 2023

In view of available recommendation, I forward this research for debate by the examining committee.

Signature:
Name: Dr. Azal Jafaar Musa
Address: Head of Department of Mathematics
Scientific grade: Assistant Professor
Date: / / 2023

Scientific Certification

This is to certify that I have read this research which is entitled **" Neural Approximation in Orlicz Spaces "** by the student **" Ikram Abed Mezher Mukheef "** and I found that this research is qualified for debate.

Signature :

Name : Dr. Azal Jafaar Musa

Title : Assistant Professor

Address : Head of Department of Mathematics

Date: / /2023

Linguistic Certification

This is to certify that I have read this research which is entitled **"" Neural Approximation in Orlicz Spaces "** by the student **" Ikram Abed Mezher Mukheef "** and I found that this research is qualified for debate.

Signature :

Name : Amera Al-Funjan

Title : Assistant Professor

Address : Lecturer

Date: / /2023

Examing Committee Certification

We certify that we have read this research entitled "**Neural Approximation in Orlicz Spaces** " and as board, tested the student "**Ikram Abed Mezher Mukheef** " in its contents and that in our opinion it is adequate for the partial fulfillment of the requirement for the Degree of Higher Diploma Education\ Mathematics.

Signature:

Name: Dr.Zainab Abdulmunim Sharba

Title: Assistant Professor

Date: / / 2023

Signature:

Name: Mayada Ali Kareem

Title: Assistant Professor

Date: / / 2023

Signature:

Name: Dr. Hawraa Abbas Almurieb

Title: Assistant Professor

Date: / / 2023

I hereby certify the decision of college of the examining committee.

Signature:

Name: Dr. Bahaa Hussien Salih Rabee

Title: Professor

Address: Dean of College of Education for Pure Sciences

Date: / / 2023

Dedication

In the hope that this work in some way may contribute in adding many things to the space .

I dedicate my work to "my lovely family" especially Mom and Dad who support, Motivate and encourage me all this time .

"To my friends and colleagues"

To every one who contributed even a latter to my academic life. For all of you, I Iove you and thank you very much.

ACKNOWLEDGEMENTS

In the name of Allah, the most gracious and the most merciful. First and foremost, Alhamdulillah, who helped me to accomplish my work and letting me through all the difficulties. I would like to express my thanks to my teacher and supervisor Dr. Hawraa Abbas Fadhil Almurieb for her support, the completion of this study could not have been possible without her expertise.

Big thanks to my lovely parents, brother and sisters for encouraging me to finish this work, your prayer for me was what sustained me this far, I am so grateful for you .

I would like to than the staff of College of Education for Pure Sciences, especially those members of Department of Mathematics in their insightful comments and encouragement.

Final thanks to my friends and colleagues who provided me with emotional and moral support throughout this challenging journey.

Publication

“Best Neural Approximation in Orlicz Spaces with s-norm”, accepted for presentation in the First International Scientific Conference "The Arab's Role in Theeoretical and Applied Sciences" at AlZahraa University for Women, College of Education, for Oct, 29-30, 2023.

In cooperation with University of Milan (Bicocca), and University of Mosul /College of Education for Human Sciences.

Abstract

Orlicz spaces include measurable functions generated by Lebesgue functions and they form Banch spaces. There are many norms defined on Orlicz space, we deal here with s-norm, where s is an outer function that defines Orlicz functions.

The $s-$norm defines a suitable modulus smoothness that gives suitable properties for approximation porpuses. Moreover, the $K-$functional to study function approximation by proving equivalence between modulus of smoothness and $K-$ functional. Both of them to get quantative estimates for degree of approximation.

It's essential to determine the approximation space to be a vital applicable one. Therefore, a class of neural networks with an activation function of ReLU type are defined. In this research, the activation function on $s-$norm Orlicz space and its properties are explained. Our choice of this activation function goes to its different applications. In addition, the best approximation by modulus of smoothness is to prove total best approximation and estimate superimum and infimum bounds of best approximation.

Contents

Figures

Table of Symbols and Abbreviations

Symbol	Definition	Page
L_p	Lebesge space	4
L_Φ	Orlicz spaces	4
Φ	Orlicz function	4
I_Φ	Modular function	4
E_Φ	Köthe dual space	5
$\|x\|_\Phi$	Orlicz norm	7
s	Outer function of Orlicz space	8
$\|x\|_{\Phi,s}$	$s-$ norm	8
$\mathbb{R}$	Real numbers	12
$\omega_k(f,\delta)_{\Phi,s}$	$s-$ modulus of smoothnees	12
$\Delta_h^k f(x)$	Symmetric difference	12
$\mathbb{N}$	Natural numbers	15
$K(f,t^r)_{\Phi,s}$	$s-$K-functional	15
ANNs	Artificial Neural networks	20
N_d	Neural network function	21
UAT	Universal approximation theorem	23
ReLU	Rectified linear unites	25
$s-$ **ReLU**, R_S	Activation function of neural for $s-$Orlicz space	27
$\mathcal{N}$	Neural networks space	29
N_f	Neural function of $s-$Orlicz space	29
$E_n(f)_{\Phi,s}$	Degree of best approximation in $s-$Orlicz space for f	31

Introduction

One of the most important classes of Banach spaces is Orlicz space. Also, Orlicz spaces L_Φ are wider than the classical Lebesgue spaces.

In 1932, W. Orlicz was introduce Orlicz spaces [2] form a wide space of measurable functions (at atomless measure) or sequences (at counting measure) which called Banach spaces.

Also in 1951, the norm $\|x\|_\Phi^A = \inf_{k>0} \frac{1}{k}\,(1 + I_\Phi(kx))$ considered by Amemiya formula [3].

In 1955, Luxemburg [4] introduced norm to the Orlicz, which is expressed by Amemiya formula [2,9]. Later, Hudzik and Maligrand [8] searched another norms from Amemiya formula norms which generating by using the s_p functions.

Two functions are used to define p −Amemiya norm [8,11]; outer and inner functions on $[0, \infty)$.

In [15], Wisla introduced the notion of outer function they investigated s-norms on Orlicz spaces. This space is well studied in this work.

For the importance of neural networks in approximation, we use neural networks from ReLU type activation function, that is called s- ReLU activation function. We conclude that $L_{\Phi,s}$ is a universal approximation using neural networks with the s −ReLU activation function can be estimated.

Our work consists of two chapters. In Chapter One, we introduce a new modulus of smoothness with respect to s −norm as follow

$$\omega_k(f,\delta)_{\Phi,s} = \sup_{0\le h\le t} \{ \|\Delta_h^k f(x)\|_{\Phi,s} : \delta \ge 0\},$$

As an alternative to $\omega_k(f,\delta)_{\Phi,s}$, one can measure smoothness using K −functionals. Moduli of smoothness $\omega_k(f,\delta)_{\Phi,s}$ and K −functionals $K(f,\delta)_{\Phi,s}$ are both functions of the real parameters, $\delta \ge 0$

$$K(f,t^r)_{\Phi,s} = \inf\left\{\|f-g\|_{\Phi,s} + t^r\left\|g^{(r)}\right\|_{\Phi,s}\right\}$$

The main theorem is to prove the equivalence between modulus of smoothness and K − functional to get

$$\omega_k(f,\delta)_{\Phi,s} \sim K(f,t^r)_{\Phi,s}$$

In Chapter Two, we define a new activation function, named, s −ReLU, that carries much similarities with ReLU activation function. Moreover , it is suitable to s-Orlicz space, so we define

$$R_s(x) = s(x)$$

with s −ReLU activation function, we construct a neural network in s-norm Orlicz space, then we prove the equivalence among degree of best approximation for functions from s- norm Orlicz space, s-Orlicz modulus of smoothness and $s - K$ − functional as follow

$$E_n(f)_{\Phi,s} \sim \omega_k(f,\delta)_{\Phi,s} \sim K(f,\delta)_{\Phi,s}$$

Chapter One:

s −norm Orlicz Space

1.1 Introduction to Orlicz Spaces

Orlicz spaces, introduced by W. Orlicz in 1932 (see [2]) form a wide class of Banach spaces of measurable functions (in the case of atomless measure) or sequences (in the case of counting measure). Also, Orlicz spaces L_Φ are generalizations of the classical Lebesgue spaces L_p for $p \geq 1$, the L_p space for $p \geq 1$ is given by

$$L_p\,([a,b]) = \{f:[a,b] \to R, \qquad f \text{ measurable and } \|f\|_p < \infty\},$$

where

$$\|f\|_p = \left(\int_a^b |f|^p\right)^{1/p}$$

A map $\Phi : R \to [0,\infty]$ is said to be an Orlicz function if $\Phi(0) = 0, \Phi$ is not identically equal to zero, it is even, convex on the interval $(-b_\Phi, b_\Phi)$ and left-continuous at b_Φ , i.e., $\lim_{u \to b_\Phi^-} \Phi(u) = \Phi(\ b_\Phi)$

For a given Orlicz function Φ, on the space $L_0(\mu)$we define a convex functional (called a pseudomodular [16]) by

$$I_\Phi\,(x) = \int_T \ \Phi\big(x(t)\big)d\mu$$

The Orlicz space L_Φ generated by an Orlicz function Φ is a linear space of measurable functions defined by the formula

$$L_\Phi \ = \ \{x \in L_0 \colon \exists \lambda > 0, I_\Phi(\lambda x) < \infty\}$$

By the space E_Φ we mean a linear subspace of L_Φ consisting of all defined as follow

$$E_\Phi = \{x \in L_\Phi : I_\Phi(\lambda_{x\chi T\setminus T_\lambda}) < \infty\}$$

where χ A denotes the characteristic function of the set A, i.e., $\chi(t) = 1$ if $t \in A$ and $\chi(t) = 0$ in the other case. Moreover, if the Orlicz function Φ takes finite values only, then we can take $T_\lambda = \emptyset$ as well. Let us note that the space E_Φ can degenerate to one element set $\{0\}$.

The definition of the norm introduced by Orlicz in 1932 was quite complicated - it was based on the modular unit ball generated by the function Ψ complementary to Φ in the sense of Young .

$$\|x\|_\Phi^0 = \sup\left\{\int_T |x(t)y(t)|d\mu : y \in L_\Psi, I_\Psi(y) \le 1\right\},$$

where Ψ is defined by

$$\Psi(u) = \sup\{|u|v - \Phi(v) : v \ge 0\},$$

for all $u \in R$. The reason for using the above formula was that it gives a simple and precise description of the space of regular linear continuous functionals on L_Φ, i.e., functionals possessing the integral representation.

A precise description of the conjugate norm to the Orlicz one is attributed to Nakano (1950), Morse-Transue (1950) and Luxemburg (1955) (see [3]). They investigated the norm, known today as the Luxemburg norm, that was defined, by use of the notion of Minikowski functional on a unit modular ball, as follows

$$\|x\|_\Phi = \inf\left\{\lambda > 0: \ I_\Phi\left(\frac{x}{\lambda}\right) \le 1\right\}$$

It was proved, under some minor assumptions, that the Köthe dual of the space

$$E_\Phi = \{x \in L_\Phi : I_\Phi(kx) < \infty \text{ for all } k > 0\},$$

equipped with the Luxemburg norm $\|\cdot\|_\Phi$ is equal to the Orlicz space L_Ψ generated by the function Ψ (complementary to Φ in the sense of Young) and equipped with the Orlicz norm $\|x\|_\Phi^0$.

In the fifties, Amemiya [4] considered the norm defined by the following formula

$$\|x\|_\Phi^A = \inf_{k>0} \frac{1}{k}\left(1 + I_\Phi(kx)\right)$$

Krasnoselskii and Rutickii [5], Nakano [6], Luxemburg and Zaanen [7] proved, under additional assumptions on the function Φ, that Orlicz norm can be expressed exactly by the Amemiya formula, i.e.

$$\|x\|_\Phi^0 = \|x\|_\Phi^A$$

In the most general case of Orlicz function Φ, the similar result was obtained by Hudzik and Maligranda [8]. Moreover, they verified easily that the Luxemburg norm can also be expressed by an Amemiya-like formula (see [2,9]), namely

$$\|x\|_{\Phi} = \inf_{k>0} \max\{1, I_{\Phi}(kx)\} \qquad (1)$$

In the paper [8], Hudzik and Maligranda proposed to investigate another class of norms given by the Amemiya formula - norms generated by the Function $s_p(u) = (1 + u^p)^{1/p}$ to define the norm

$$\|x\|_{\Phi,p} = \inf_{k>0} \frac{1}{k} \left(1 + I_{\Phi}^{p}(kx)\right)^{1/p} \qquad (2)$$

where $1 \leq p \leq \infty$ (if $p = \infty$ then we use the formula (1).

In that case, we obtain a family of topologically equivalent norms (called $p -$ Amemiya norms and denot by $\|x\|_{\Phi,p}$ indexed by $1 \leq p \leq \infty$ and satisfying the inequalities

$$\|x\|_{\Phi} = \|x\|_{\Phi,\infty} \leq \|x\|_{\Phi,p} \leq \|x\|_{\Phi,q} \leq \|x\|_{\Phi,1} = \|x\|_{\Phi}^{0} \leq 2\|x\|_{\Phi},$$

$$\|x\|_{\Phi} \leq \|x\|_{\Phi,p} \leq 2^{\frac{1}{p}} \|x\|_{\Phi}$$

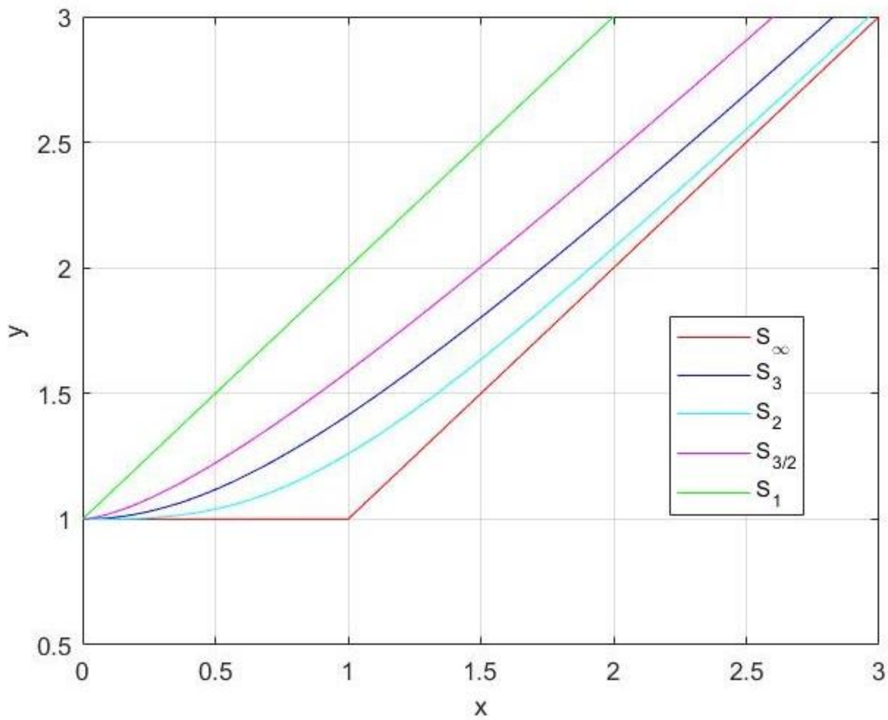

Figure 1.1 The functions $s_p = (1 + u^p)^{1/p}$

In the Figure1.1, we represent some curves of s_p by MATLAB.

for all $1 \leq q \leq p \leq \infty$ under some additional conditions (e.g., in the atomless measure case, Φ, Ψ take only finite values and Ψ is not linear on (0, ∞)) there is a natural Köthe duality between these norms, that is, the Köthe dual of $\left(E_\Phi, \|x\|_{\Phi,p}\right)$ coincides with $\left(L_\Psi, \|x\|_{\Psi,p}\right)$, where $1/p + 1/q = 1$ and Ψ denotes the complementary function to Φ in the sense of Young.

Further details about Orlicz spaces equipped with the Luxemburg or the Orlicz norm, can be found in [1,10]. Basic results on Orlicz spaces equipped with p −Amemiya norms have been presented in [11], while an overview of the resent results on that spaces has been presented in [12]. Note that these results can be applied, among others, to dominated best approximation problems [13] and fixed point property (see e.g. [14]).

In [15], Wisla introduced the notion of outer function, he investigated s-norms on Orlicz spaces defined by the formula

$$\|x\|_{\Phi,s} = \inf_{k>0} \frac{1}{k}\, s(I_\Phi(kx))$$

The class of outer functions is very wide and covers all the functions s_p.

The aim of Wisla's work is to present the basics of the theory of s −norms. He presented the basic properties of outer functions and s −norms.

1.2 Outer Functions and s −Norms

The p −Amemiya norm (2)(see [11,8]) was defined by using of two functions: the (inner) Orlicz function Φ (more precisely: the modular I_Φ) and the outer function s_p defined on half-line $[0.\infty)$ by

$$s_p(u) = \begin{cases} (1 + u^p\ )^{\frac{1}{p}} & \text{for } 1 \le p < \infty \\ max\{1, u\} & \text{for } p = \infty \end{cases}.$$

The family $\{s_p: 1 \le p \le \infty\}$ consists of functions that satisfy the following properties:

1. convex.

2. non-decreasing on $[0.\infty)$ with exactly one common point (knot) at 0 (i.e., $s_p(0) = 1$ for all $1 \le p \le \infty$).

3. On the half-line $[0.\infty)$, the functions s_p are strictly increasing for $1 \le p < \infty$, strictly convex for $1 < p < \infty$ and $s_p(u) < s_{p'}(u)$ for every $1 \le p' < p \le \infty$ and $u > 0$.

Definition 1.2.1[15] A function $s : [0\ ,\infty) \to [1.,\infty)$ is called an outer function, if it is convex and

$$\max\{u, 1\} \le s(u) \le u + 1, \text{ for all } u \ge 0$$

In order to simplify notations, Wilsa extended the domain and range of s to the interval $[0, \infty]$ by setting $s(\infty) = \infty$.

Note that the family of outer functions has common properties: for each outer function s.

1. $s(0) = 1$.

2. s admits an asymptote at infinity with the slope 1 .

3. $\inf_{k>0} \frac{1}{k} s(k) = 1$, since

$$1 = \inf_{k>0} \frac{1}{k} \max\{1, k\} \leq \inf_{k>0} \frac{1}{k} s(k) \leq \inf_{k>0} \frac{1}{k} (1 + k) = 1$$

Now, we are ready to interduce the norm of Wisla [15] in the following definition**.**

Definition 1.2.2 [15] If s and Φ are outer and Orlicz functions respectively, then the functional

$$\|x\|_{\Phi,s} = \inf_{k>0} \frac{1}{k} s(I_{\Phi}(kx)) \qquad (3)$$

is a norm on the Orlicz space.

For any outer function s and any Orlicz function Φ the norm given by the formula (3) is called the s −norm.

In the following lemma that is proved by wisla [15], we interduce some of the main properties of s −norm.

Lemma 1.2.3 [15] For every outer function s and Orlicz function Φ, we have

(i) $\|x\|_{\Phi,\infty} \leq \|x\|_{\Phi,s} \leq \|x\|_{\Phi,1} \leq 2\|x\|_{\Phi,\infty}$ for $all\ x \in L_{\Phi}$

(ii) The norm topology on L_{Φ} generated by s −norm $\|x\|_{\Phi,s}$ is equivalent to the topology generated by the Orlicz or Luxemburg norm, *i.e.*, for every sequence x_n of elements of L_{Φ} and every x,

$$\|x_n - x\|_{\Phi,s} \to 0 \Leftrightarrow \forall\lambda > 0,\ I_\Phi(\lambda(x_n - x)) \to 0$$

(iii) The s −norm $\|x\|_{\Phi,s}$ is order continuous on the space E_Φ.

1.3 Moduli of Smoothness

Moduli of smoothness represent important tools for obtaining quantitative estimates of the error of approximation for many processes. There are various such special functions associated with wide classes of function spaces. For example, most theoretical estimates for the order of approximation of functions by say, polynomials or splines or operators, such as waveletes or neural networks, are now given in terms of such a modulus. Such theoretically elegant results are, however, not always easy to use in practice. Very often, the potential user will have a particular function and will require a quantitative estimate of its order of approximation. Such an estimate depends in turn on an accurate calculation of the order of the modulus of smoothnessof the given function [17]. Our porpuse of this section is to define a modulus of smoothness with respect to s-norm Definition 1.4.1.

1.4 s −Orlicz Modulus of Smoothnees

In this section, we introduce a modulus of smoothness with respect to s −norm , the purpose is estimating the dgree of approximation as good as possible, so we define a new modulus of smoothness.

Definition 1.4.1

The k $- th$ order L_Φ -modulus of smoothness of any function

$$f \in L_\Phi[a,b], \qquad a,b \in \mathbb{R}$$

is given by

$$\omega_k(f,\delta)_{\Phi,s} = \sup_{0\le h\le\delta} \{ \|\Delta_h^k f(x)\|_{\Phi,s} : \delta \ge 0\},$$

where

$$\Delta_h^k f(x) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} f(x+jh)$$

$\|x\|_{\Phi,s}$ is the s −norm on $[a,b]$ defined in Definition 1.2.2

Moreover, we study the main properties any modulus should have for approximation purposes. In the following theorem we prove the main results about the so-called s −modulus of smoothness.

1.5 Properties of s −Modulus of Smoothness

In this section, we study and prove the main properties of s −modulus of smoothness

Theorem 1.5.1

Let $k \in \mathbb{N}_0, f,g \in L_{\Phi,s}[a,b]$, then for all $\delta > 0$, we have

(a) $\lim_{\delta\to 0^+} \omega_k(f,\delta)_{\Phi,s} = 0$

(b) $\omega_k(f+g,\delta)_{\Phi,s} \le \omega_k(f,\delta)_{\Phi,s} + \omega_k(g,\delta)_{\Phi,s}$

(c) $\omega_k(cf,\delta)_{\Phi,s} = |c|\omega_k(f,\delta)_{\Phi,s}$

(d) $\omega_k(f,\delta)_{\Phi,s} \le \omega_k(f,\delta')_{\Phi,s}$, $\delta < \delta'$

(e) $\omega_k(f,\lambda\delta)_{\Phi,s} \le \omega_k(f,\delta)_{\Phi,s} \le (1+\lambda)^k\, \omega_k(f,\delta)_{\Phi,s}$, for $0 < \lambda < 1$

(f) $\omega_k(f, \delta)_{\Phi,s} \le c(j,k)\|f\|_{\Phi,s}$

Proof:

(a) $\lim_{\delta\to 0^+} \omega_k(f,\delta)_{\Phi,s} = \lim_{\delta\to 0^+} \sup_{0\le h\le\delta}\{\|\Delta_h^k f(x)\|_{\Phi,s}\} = 0$

(if $\delta \to 0^+$ then $f(x) \to 0^+$, f is continiuos) $\delta \nrightarrow 0^-$

(b) $\omega_k(f+g,\delta)_{\Phi,s} = \sup_{0\le h\le\delta}\{\|\Delta_h^k (f+g)(x)\|_{\Phi,s}\}$

$$\le \sup_{0\le h\le\delta}\{\|\Delta_h^k f(x)\|_{\Phi,s}\} + \sup_{0\le h\le\delta}\{\|\Delta_h^k g(x)\|_{\Phi,s}\}$$

(Minkowisk inequality)

$$\le \omega_k(f,\delta)_{\Phi,s} + \omega_k(g,\delta)_{\Phi,s}$$

(c) $\omega_k(cf,\delta)_{\Phi,s} = \sup_{0\le h\le\delta}\{\|\Delta_h^k cf(x)\|_{\Phi,s}\}$

$$= \sup_{0\le h\le\delta}\{|c|\,\|\Delta_h^k f(x)\|_{\Phi,s}\}$$

$$= |c| \sup_{0\le h\le\delta}\{\|\Delta_h^k f(x)\|_{\Phi,s}\}$$

$$= |c|\omega_k(f,\delta)_{\Phi,s}$$

(d) $\omega_k(f,\delta)_{\Phi,s} = \sup_{0\le h\le\delta}\{\|\Delta_h^k f(x)\|_{\Phi,s}\}$

$$\le \sup_{0\le h\le \delta'}\{\|\Delta_h^k f(x)\|_{\Phi,s}\} \quad ,\delta < \delta'$$

$$= \omega_k(f,\delta')_{\Phi,s}$$

(e) If $0 < \lambda < 1$, then $\lambda\delta < \delta$, so by (d), we get

$\omega_k(f,\lambda\delta)_{\Phi,s} \le \omega_k(f,\delta)_{\Phi,s}$

$$\le (1+\lambda)^k\, \omega_k(f,\delta)_{\Phi,s}$$

(f) $\omega_k(f,\delta)_{\Phi,s} = \sup_{0\le h\le\delta}\left\{\left\|\Delta_h^k f(x)\right\|_{\Phi,s}\right\}$

$$= \sup_{0\le h\le\delta}\left\{\inf_{k>0}\frac{1}{k}\, s\left(I_\Phi\left(k\left(\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j} f(x+jh)\right)\right)\right)\right\}$$

$$= \sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\sup_{0\le h\le\delta}\left\{\inf_{k>0}\frac{1}{k}\, s\left(I_\Phi\left(k\big(f(x+jh)\big)\right)\right)\right\}$$

$$\le \sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\sup_{0\le h\le\delta}\left\{\|f\|_{\Phi,s}\right\}$$

$$= c(j,k)\,\|f\|_{\Phi.s}$$ ■

1.6 K − Functional

As an alternative to $\omega_k(f,\delta)_{\Phi.s}$ one can measure smoothness using K −functional. Moduli of smoothness $\omega_k(f,\delta)_{\Phi.s}$ and K −functionals $K(f,t)$ are both functions of the real parameters, $\delta \ge 0$, $t \ge 0$ which express some intrinsic properties of function f.

The simple K- functional has been introduced by Peetre in [18] is given by

$$K(f,t) = K(f,t;X_0,X_1) := \inf \|f - g\|_{X_0} + t\|g\|_{X_1}, t \ge 0, g \in X_1$$

Where X_i , $i = 0,1$ be two Banach spaces, with X_1 continuously embedded in X_0: $X_1 \subset X_0$

This quantity expresses some approximation properties of f. That's the inequality $K(f,t) < \varepsilon$ in X_0 by an element $g \in X_1$, whose norm is not too large $\|g\|_{X_1} < \varepsilon t - 1$.

For $s-$ Orlicz norm spaces, and an interval $[a,b]$. We define K-functional as follow,

Definition 1.6.1

For any $f \in L_{\Phi,s}$, K- functional is given by

$$K(f,t^r)_{\Phi,s} = \inf\left\{\|f-g\|_{\Phi,s} + t^r\left\|g^{(r)}\right\|_{\Phi,s}\right\}, g^{(r)} \in L_{\Phi,s},$$

where $r \in \mathbb{N}$.

1.7 Equivalence of Modulus of Smoothness with K −functional

The main result here is to study the equivalance between s −modulus of smoothness and K −functional. It is benefit for studying the degree of best approximation later.

Theorem 1.7.1

For any $f \in L_{\Phi,s}[a,b]$,

$$\omega_k(f,\delta)_{\Phi,s} \sim K(f,t^r)_{\Phi,s}$$

Proof: For any $\varepsilon > 0$, there exists $g^{(r)} \in L_{\Phi,s}$, such that

$$\|f-g\|_{\Phi,s} + t^r\left\|g^{(r)}\right\|_{\Phi,s} \leq K(f,t^r)_{\Phi,s} + \varepsilon$$

and

$$\omega_k(f,\delta)_{\Phi,s} \leq c(r)t^r\left\|g^{(r)}\right\|_{\Phi,s},$$

for some $\delta \leq t^r$

Applying some properties of s −modulus of smoothness from Theorem 1.5.1 yields

$$\omega_k(f,\delta)_{\Phi,s} \le \omega_k(f-g,\delta)_{\Phi,s} + \omega_k(g,\delta)_{\Phi,s}$$

$$\le 2^r\|f-g\|_{\Phi,s} + c(r)t^r\left\|g^{(r)}\right\|_{\Phi,s}$$

$$\le c\ \left(K(f,t^r)_{\Phi,s} + \varepsilon\right)$$

By letting $\varepsilon \to 0$, we obtain

$$\omega_k(f,\delta)_{\Phi,s} \le c(r)K(f,t^r)_{\Phi,s}$$

On the other hand, we have by Definition 1.6.1 and Lemma 1.2.3

$$K(f,t^r)_{\Phi,s} = \inf\left\{\|f-g\|_{\Phi,s} + t^r\left\|g^{(r)}\right\|_{\Phi,s}\right\}$$

$$\le 2\ \inf\left\{\|f-g\|_{\Phi,\infty} + t^r\left\|g^{(r)}\right\|_{\Phi,\infty}\right\}$$

$$\le 2\ K(f,t^r)_{\Phi,\infty}$$

$$\le 2\ \omega_k(f,t^r)_{\Phi,\infty}$$

$$= 2\sup\{\|\Delta_h^r(f)\|_{\Phi,\infty}\}$$

$$\le 2\sup\{\|\Delta_h^r(f)\|_{\Phi,s}\}$$

$$\le 2\ \omega_k(f,t^r)_{\Phi,s}$$

we get

$$K(f,t^r)_{\Phi,s} \le 2\ \omega_k(f,t^r)_{\Phi,s}$$

Thus,

$$\omega_k(f,\delta)_{\Phi,s} \sim K(f,t^r)_{\Phi,s} \qquad \blacksquare$$

In the following theorem, we prove property (e) from Theorem 1.5.1 for the case $\lambda \ge 1$.

Theorem 1.7.2

$\omega_k(f,\lambda\delta)_{\Phi,s} \leq 2c\lambda\, \omega_k(f,\delta)_{\Phi,s}$, for $\lambda \geq 1$

Proof:

If $\lambda \geq 1$, then $\delta \leq \lambda\delta$, by Theorem 1.5.1 (d), we get

$$\omega_k(f,\delta)_{\Phi,s} \leq \omega_k(f,\lambda\delta)_{\Phi,s}$$

Otherwise, by Theorem 1.7.1, we get

$$\omega_k(f,\lambda\delta)_{\Phi,s} \leq c(r)\, K(f,\lambda t^r)_{\Phi,s} \leq c(r)\lambda\, K(f,t^r)_{\Phi,s}$$

$$\leq 2c(r)\lambda\, \omega_k(f,\delta)_{\Phi,s}$$

where c is constant. ■

Chapter Two:

Approximation by Neural Networks in s −Orlicz Spaces

In this chapter, we study best approximation in the space of study in chapter One, that is s −Orlicz space with neural networks. We begin with brief introduction to approximation by neural networks.

2.1 Introduction to Best Approximations

In this section, we interduce the main concepts, theorems and definitions. Moreover, we project the concepts, definitions and theorems directly into s −Orlicz space $L_{\Phi,S}$. The proofs comes later.

Definition 2.1.1.[21]

Let $\mathcal{N}$ be a nonempty subset of the space $L_{\Phi,S}$, and let $f \in L_{\Phi,S}$. An element $g_o \in \mathcal{N}$ is called a best approximation, or nearest point, to f from $\mathcal{N}$ if

$$\|f - g_o\| = d(f,g),$$

where $d(f,g) \coloneqq \inf_{g\in\mathcal{N}} \|f-g\|$. The number $d(f,g)$ is called the distance from f to $\mathcal{N}$, or the error in approximating f by $\mathcal{N}$. The (possibly empty) set of all best approximations from f to $\mathcal{N}$ is denoted by $\mathcal{N}_f$. Thus

$$\mathcal{N}_f := \{g \in \mathcal{N} \ : \ \|f-g\| = d(f,\mathcal{N})\}.$$

2.1.2 Existence Theorem

Let $\mathcal{N}$ be a finite-dimensional subspace of a normed linear space $L_{\Phi,S}$ and let $f \in L_{\Phi,S}$. Then, there exists a (not necessarily unique) $g^* \in \mathcal{N}$ such that

$$\|f - g^*\| = \min_{g\in\mathcal{N}} \|f - g\|$$

for all $g \in \mathcal{N}$. That is, there is a best approximation to f by elements of $\mathcal{N}$.

One of the earliest cases of existence of approximation is the following one that was originally discovered by Weierstrass.

2.1.3 Uniqueness Theorem

$L_{\Phi,S}$ has a strictly convex norm if and only if the triangle inequality is strict on non-parallel vectors; that is, if and only if

$f \neq \alpha g, g \neq \alpha f,$ for all $\alpha \in \mathbb{R}$ Then $\|f + g\| < \|f\| + \|g\|$.

2.2 Approximation Theory in Neural Networks

Artificial Neural Networks (ANNs) contains several layers of neurons, an input layer $\boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$, one or more hidden layers with an activation function σ, and an output layer. The general mathematical formula of any neural network is:

$$N_d(\boldsymbol{x}) = \sum_{i=1}^{d} c_i\sigma(w_i x_i + b_i) \quad ,$$

where $w_i \in \mathbb{R}$ are the weights, $c_i \in \mathbb{R}$ are coefficients and $b_i \in \mathbb{R}$ are bases. Figure 2.1 shows the procedure of nueral network algorithm.

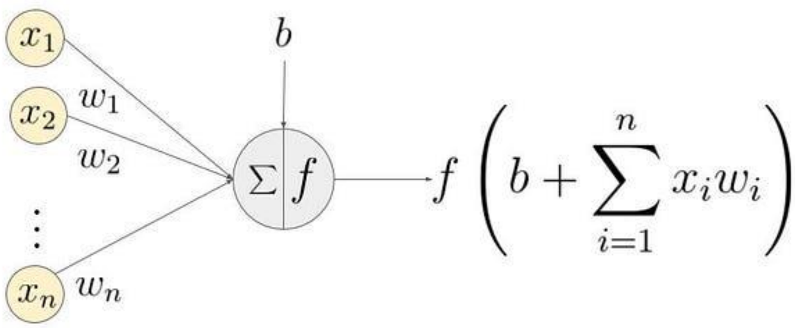

Figure 2.1 Neural Network mathematically

The range of functions that are approximated by neural networks is very wide, due to their applications in different fields, they can approximate any function with some conditions, such as continuity,integrability and sufficient training sets.[22], [23].

To understand more about the relationship of neural networks in the approximation of functions, we first need to talk about the activation function and its importance in neural networks. In neural networks we have neurons, each neuron receives inputs and performs weighted summation operations on them, then passes the resulted summation into the activation function. That turns it into an output.

The main part of approximation theory of artificial neural networks is that it concerns with universal approximation theorems, they are results that establish the density of an algorithmically generated class of functions within a given function space of interest. Typically, these results concern the approximation capabilities of the feedforward architecture on the space of continuous functions between two Euclidean spaces, and the approximation is with respect to the compact convergence topology.

However, there are also a variety of results between non-Euclidean spaces and other commonly used architectures and, more generally, algorithmically generated sets of functions, such as the convolutional neural network (CNN) architecture, radial basis-functions, or neural networks with specific properties. Most universal approximation theorems can be parsed into two classes, the first quantifies the approximation capabilities of neural networks with an arbitrary number of artificial neurons ("*arbitrary width*" case) and the second focuses on the case with an arbitrary number of hidden layers, each containing a limited number of artificial neurons ("*arbitrary depth*" case). In addition to these two classes, there are also universal approximation theorems for neural networks with bounded number of hidden layers and a limited number of neurons in each layer ("*bounded depth and bounded width*" case). [23]

Universal approximation theorems imply that neural networks can represent a wide variety of interesting functions when given appropriate weights. On the other hand, they typically do not provide a construction for the weights, but merely state that such a construction is possible.

2.2.1 Universal Approximation Theorem (UAT)

For any continuous function f defined on a compact space $X \subseteq \mathbb{R}$, there exists a neural network of single hidden layer $N_n f$ that satisfies

$$\|f - N_n f\| < \varepsilon$$

The combination of the two topics (Approximation and NNs) supply strength not only to approximation aspects but also to the neural network topic itself. Because of the accurate convergence to the value function, the approximated neural network can be an excellent replacement for the original function.

"Neural-Networks are Universal Function Approximators" means that almost any function represents some neural network that computes and learns any process at all.

In spite of the importance of the UAT, it suffered from a number of limitations; it was primitive for the following aspects; the used function space, the degree of approximation, and the nature of the neural network itself.

However, under UAT, thousands of papers were published in the field of neural approximation. Tens of neural approximation theorems were proved in this manner. These theorems simulate Cybenko’s Theorem implicitly, but they are unlike in details. Such details as function spaces, mathematical structures of neural networks including activation functions and others.

Figure2.2 shows neural network of the most important approximation theorems of that period. Each color refers to a layer of the neural network, the input layer is of the color yellow, the green is for the hidden layers, and the red is for the output layers.

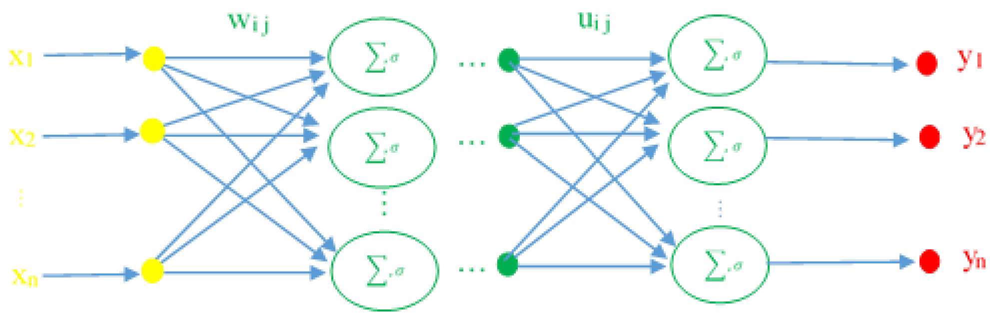

Figure 2.2 n-layer Feedforward Neural Network Approximate Continuous Functions

In addition, the activation function of the neural network defined in UAT is not specified. He proved the theorem for any neural network with any arbitrary activation function. The importance of the activation function is not only to to create a relationship between inputs and outputs but also to add the ability to the network to learn any type of data. To build a more powerful network, it is essential to choose a suitable activation function depending on various issues such as; type of data, number of hidden layers and the network model.

The question now is, what if no activation function is used and the neurons are allowed to give the sum of data to the inputs as well as to the outputs. In this case, the calculation will be very simple because the weighted sum of the inputs has no range. Hence, an important use of activation functions is to keep the output data restricted to a certain range. Its activation functions help neural networks learn complex relationships in data. Another use of the activation function is to add nonlinearity to the data. Nonlinear functions are always chosen as the activation functions.

Hereafter, we use a new-defined ReLU type activation function with special chractaristics concerns with the space of study.

2.3 Rectified Activation Function(ReLU) in s −Orlicz Space

In the last couple of decades, ReLU is used to generate powerful models of artificial neural networks by avoiding vanishing gradient problems and to reduce the slowness of convergence. They were V. Nair and E. Hinton [33] who use ReLU for the first time to construct a more efficient learning model. The mathematical formula of ReLU is given by

$$\sigma(x) = max\{0, x\}$$

σ combines linearity and nonlinearity, i.e. it is linear for positive values, so it could be a benefit in backpropagation. Yet, it is a nonlinear function as the output of negative values is always “zero”.

In spite of its simplicity and efficiency, ReLU is not used alone in deep learning problems for several reasons. It is used only in the hidden layers. In problems of classification, SoftMax should be used in the output layer in order to classify outputs to different categories.

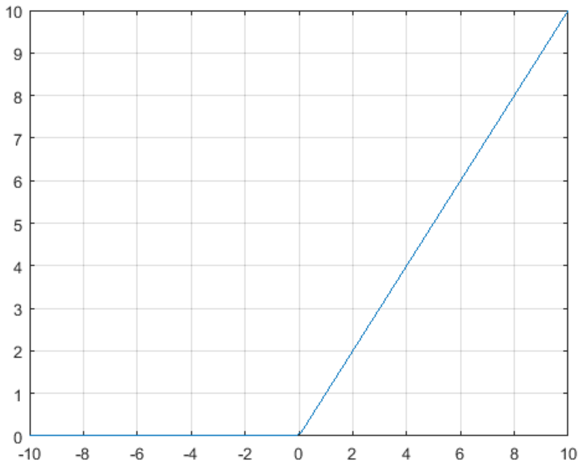

Figure 2.3 ReLU function

On the other hand, linear functions are simply enough to be activated in the output layer for regression problems.

Moreover, some fragile gradients of ReLU may kill some neurons during training, which causes disables in the next data update. To solve this problem, Leaky ReLU is defined to replace ReLU in situations that need alive gradients. They sometimes work together in the same network under what is called "Mixout" activation function. Mathematically, the leaky ReLU is given by

$$\sigma(x) = max\{\propto, \propto x\}$$

where $\propto$ is a small value that introduces a small slope to keep the weights alive through backpropagation. Unlike ReLU, leaky ReLU is differentiable at zero and everywhere[23].

Even it should be only used within hidden layers; it is still one of the most popular used activation functions in different fields until now.

Before we explain the approximation capabilities of neural, we need to determine the function space and the criterion of approximation that makes the rate of approximation as much accurate as we can.

2.4 Neural Approximation in s −Orlicz Space

In this section, we define a new activation function that carries much similarities with ReLU activation function. Moreover, it is suitable to s −Orlicz space .

Definetion 2.4.1

Let s any outer function of the formula in Definetion 1.2.1 then the s −ReLU activation function is given by

$$R_s: (-\infty, \infty) \to [0,\infty) : \; R_s(x) = s(x) \qquad (4)$$

where

$$\max\{u, 1\} \leq s(u) \leq u + 1$$

2.5 Properties of s −ReLU Activation Function

As mentioned before, any activation function should carry some main properties,such as, mathematical complexity, applicape simplicity, differentiability and combination between linearity and nonlinearity. The last property is specific to the ReLU type function

In the following theorem, we given some properties of s −ReLU function that any ReLU function has, with other more essential properties that comes frome the outer function s in Definetion 1.2.1, such as, convexity, finiteness, being within curves of ReLU function $\sigma(x) = max\{0, x\}$

Theorem 2.5.1

s −ReLU activation function from (4) in Definition 2.4.1satisfies

1. Convexity: $\forall x \in (-\infty, \infty)$: $R_s(x)$ is convex since s is convex.

2. Curve of R_s lie within curve of R: $\forall x \in (-\infty, \infty)$; $0 \leq R_s(x) \leq x$, see Figure 1.1 and Figure 2.3.

3. Finiteness: $\forall x \in (-\infty, \infty)$; $\|R_s(x)\|_{\Phi,s} < \infty$

4. Differentation: $R'_s(x) = \begin{cases} 0 & if\ x \le 0 \\ s'(x) & if\ x > 0 \end{cases}$, conditioning that s' exists.

2.6 Construction of NNs in s-Orlicz normed Space

In this section, we are ready to define a neural network with s −ReLU activation function.

Definition 2.6.1

Let $\mathcal{N}$ be the space of all neural networks of the form

$$N_f(\boldsymbol{x}) = \sum_{i=1}^{d} \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} f(x + jh) R_s(x) \qquad (5)$$

where $x \in [a, b]$, $f \in L_{\Phi,s}$, $h < \delta$

Now, the main theorem of approximation are being proved, we being with direct theorem that evaluate the degree of best approximation in terms of neural networks of the form (5) in Definition 2.6.1

2.6.2 Direct Theorem

For any function $f \in L_{\Phi,s}\,[a, b]$, there exists a neural network $N_f \in \mathcal{N}$ that satisfies :

$$\|f - N_f\|_{\Phi,s} \le \omega_k(f, \delta)_{\Phi,s}$$

Proof :

By Definition 2.6.1, Theorem 1.5.1 and Theorem 2.5.1

$$\left\|N_f - f\right\|_{\Phi,s} = \left\|\sum_{i=1}^{d}\sum_{j=1}^{k}(-1)^{k-j}\binom{k}{j} f(x+jh)R_s(x) - f(x)\right\|_{\Phi,s}$$

$$\leq \sum_{i=1}^{d}\left\|\sum_{j=1}^{k}(-1)^{k-j}\binom{k}{j} f(x+jh)R_s(x) - f(x)\right\|_{\Phi,s}$$

$$\leq \sum_{i=1}^{d}\left\|\sum_{j=1}^{k}(-1)^{k-j}\binom{k}{j} R_s(x)\,[f(x+jh) - f(x)]\right\|_{\Phi,s}$$

$$\leq c\sum_{i=1}^{d}\omega_k(f.\delta)_{\Phi.s}$$

$$\leq c(d)\omega_k(f.\delta)_{\Phi.s}$$

Hence

$$\left\|f - N_f\right\|_{\Phi,s} \leq \omega_k(f,\delta)_{\Phi,s} \qquad \blacksquare$$

On the other hand, inverse theorem grantee that the degree of approximation is equivalent to $s -$ modulus of smoothness.

2.6.3 Inverse Theorem

For every $f \in L_{\Phi,s}$, $\exists\ N_f \in \mathcal{N}$ that satisfies

$$\omega_k(f,\delta)_{\Phi,s} \leq \left\|f - N_f\right\|_{\Phi,s}$$

Proof :

$$\|f(x)\|_{\Phi,s} \leq \left\|f - N_f\right\|_{\Phi,s} + \left\|N_f\right\|_{\Phi,s}$$

$$\leq \left\|f - N_f\right\|_{\Phi,s} + c\omega_k(f,\delta\,)_{\Phi.s}$$

Since

$$K(f,t^r)_{\Phi,s} \leq \left\|f - N_f\right\|_{\Phi,s} + t^r\left\|f - N_f\right\|_{\Phi,s} + c\omega_k(f,\delta\,)_{\Phi.s}$$

$$\omega_k(f,\delta\,)_{\Phi.s} \leq \left\|f - N_f\right\|_{\Phi,s} + \omega_k(f,\delta\,)_{\Phi.s}$$

$$\omega_k(f,\delta\,)_{\Phi.s} \leq \left\|f - N_f\right\|_{\Phi,s} \qquad \blacksquare$$

Now by Direct Theorem 2.6.2 and Inverse Theorem 2.6.3 we get

$$E_n(f)_{\Phi,s} \sim \omega_k(f,\delta\,)_{\Phi.s}$$

CONCLUSION

To obtainment best approximation for any function in $L_{\Phi,s}$, we define $s-$moduly of smoothnees. The definition of modulus of smoothnees has some functional properties and relations equivalent to the K-functional, which define on the s-Orlicz spaces.

In our work, we present function approximation on $L_{\Phi,s}$ by using neural networks from ReLU type, called $s-$ ReLU activation function. We conclude that neural networks with $s-$ ReLU activation function are universal approximation in $L_{\Phi,s}$.

Also, it is important in the field of approximation theory, to prove inverse and direct theorems and that is important to increase the trust in our work.

FUTURE WORK

Although Orlicz spaces have been extensively studied, there is still room for further investigation into their properties and characterizations. The study of Orlicz spaces can contribute to understanding the generalization ability of neural networks. By establishing suitable bounds in Orlicz spaces, researchers can analyze the trade-off between model complexity and generalization performance. This can guide the development of new regularization techniques specific to neural networks.

Conducting theoretical analysis on the expressive power and approximation properties of neural networks in Orlicz spaces can deepen our understanding of their capabilities and limitations. This can involve investigating the approximation rate of neural networks with Orlicz norms and exploring the associated error bounds.

REFERNECES

[1] K. Conrad, "Lp-Spaces for $0 < p < 1$," Lecture Notes. pp. 1–15, 2003.

[2] W. Orlicz, Über eine gewisse Klasse von Räumen vom Typus B, Bull. Intern. Pol. Ser. A Kraków (1932) 207–220.

[3] H. Nakano, Topology and Linear Topological Spaces, Tokyo Math. Book Series, vol.3, Maruzen, Tokyo, 1951

[4] W.A.J. Luxemburg, Banach Function Spaces, Thesis, Delft, 1955

[5] M.A. Krasnoselskii, Y.B. Rutickii, Convex Function and Orlicz Spaces, Nordhoff, Groningen, 1961.

[6] H. Nakano, Topology and Linear Topological Spaces, Tokyo Math. Book Series, vol.3, Maruzen, Tokyo, 1951

[7] W.A.J. Luxemburg, A.C. Zaanen, Conjugate spaces of Orlicz spaces, Indag. Math. 18 (1956) 217–228.

[8] H. Hudzik, L. Maligranda, Amemiya norm equals Orlicz norm in general, Indag. Math. 11(4) (2000) 573–585.

[9] W. Orlicz, A note on modular spaces, Bull. Acad. Polon. Sci. Math. 9 (1961) 157–162.

[10] M.M. Rao, Z.D. Ren, Theory of Orlicz Spaces, Dekker Inc., New York, 1991

[11] Y. Cui, L. Duan, H. Hudzik, M. Wisła, Basic theory of p-Amemiya norm in Orlicz spaces (1 ≤p ≤∞): extreme points and rotundity in Orlicz spaces equipped with these norms, Nonlinear Anal. 69 (2008) 1796–1816.

[12] M. Wisła, Geometric properties of Orlicz spaces equipped with p-Amemiya norms -results and open questions, Comm. Math. 55(2) (2015) 183–209

[13] Y. Cui, H. Hudzik, M. Wisła, Monotonicity properties and dominated best approximation problems in Orlicz spaces equipped with the p-Amemiya norm, J. Math. Anal. Appl. 432 (2015) 1095–1105

[14] Y. Cui, H. Hudzik, M. Wisła, K. Wlaźlak, Non-squareness properties of Orlicz spaces equipped with the p-Amemiya norm, Nonlinear Anal. 75 (2012) 3973–3993.

[15] M . Wisla , spaces equipped with S-norms, Journal of Mathematical Analysis and Applications , 483 (2020) 123659.

[16] J. Musielak, Orlicz Spaces and Modular Spaces, Lecture Notes in Math., vol.1034, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983

[17] Z . A . Abdulhady & E . S . Bhaya , Versions of Moduli of Smoothness , PhD Thesis , University of Babylon , 2021.

[18] K. G. Ivanov, "A Characterization of Weighted Peetre K-Functionals." Journal of Approximation Theory, vol. 56, no .2, pp. 185-211, 1989.

[19] A . S . Mnahi & H . A . Almurieb , Functions Approximation by Spectral Graph Waveles . Dip. Thesis , University of Babylon , 2022.

[20] N. L. Carothers, A Short Course on Approximation Theory, Math 682 Summer ,Citeseer , 1998.

[21] F. Deutsch, Best Approximation in Inner Product Spaces, Canadian Mathematical society (societe Mathematique du Canada) . Springer-Verlag New York, Inc. 2001.

[22] G. Cybenko, Continuous Valued Neural Networks: Approximation Theoretic Results, in Computer Science and Statistics: proceedings of the 20th Symposium on the Interface, 1988, pp. 174–183.

[23] H . A . Almureib & E . S . Bhaya, The Degree of Best Neural Networks Approximation With Application, PhD Thesis, University of Babylon, 2020.

[24] S . Dekel, On the equivalence of the modulus of smoothness and the K – functional over convex domains, Journal of Approximation Theory, 26 August 2008.

[25] E . S . Bhaya & N . A . Hadi, Jacobi Weighted Moduli of Smoothness for Approximation by Neural Networks Application, Journal of University of Babylon / Pure and Applied sciences, vol.(28), No.(3): 2020.

[26] A . A . Mardan & H . A . Almurieb ,Functions Approximation by Learing Neural Networks, M.S.c Thesis, University of Babylon, 2022.

[27] S. Chen, Geometry of Orlicz spaces, Dissertationes Math. 356 (1996).

[28] Johnen, Hans, and Karl Scherer. On the equivalence of the K-functional and moduli of continuity and some applications. Constructive Theory of Functions of Several Variables: Proceedings of a Conference Held at Oberwolfach April 25–May 1, 1976. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.

المستخلص

من المعلوم ان فضاء اورلكز فضاء يشمل الدوال القابله للقياس حيث انه تعميم لفضاء ليبيك ويمثل فضاء بناخ.

تم تعريف عدة أشكال من المعاييرعلى فضاء اورلكز، وهنا يكون عملنا على المعيار s.في هذا العمل قمنا بتعريف مقياس نعومه مناسب للمعيار s وذو خواص مناسبه. وكذلك عرّفنا عليه الدالي-K وبيّنا التكافؤ بين مقياس النعومه و الدالي-K واستخدمنا كليهما للحصول على تقديرات كميه لدرجه التقريب الأفضل.

من المعروف انه من الضروري اختيار فضاء التقريب ليكون فضاءً قابلا للتطبيق .

لذلك تم تعريف شبكات عصبيه من نوع داله التفعيل ReLU وعرّفناهاعلى فضاء s-Orlicz space ووضحنا خواصها ،لما لها من تطبيقات مختلفه بالاضافه لسهولتها وكفاءتها وللحصول على افضل تقريب بدلالة مقياس النعومه وتم اثبات مبرهنة التقريب الأفضل الكلي ،كما تم تقدير الحدود العليا والدنيا لدرجة التقريب الافضل .

جمهورية العراق
وزارة التعليم العالي والبحث والعلمي
جامعة بابل
كلية التربية للعلوم الصرفة
قسم الرياضيات

التقريب العصبي في فضاءات اورليكز

بحث مقدم إلى

مجلس كلية التربية للعلوم الصرفة – جامعة بابل

كجزء من متطلبات نيل شهادة الدبلوم العالي تربية / الرياضيات

من قبل

اكرام عبد مزهر مخيف

بإشراف

أ.م.د. حــــوراء عبــــاس فاضــل المرعب

2023 م **1445 هـ**