PROCEEDINGS OF SCIENCE

Solving Universal Approximation Problem by a Class of Neural Networks based on Hankel Approximate Identity in Function Spaces

Saeed Panahian Fard[1]
School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM Penang, Malaysia
Email: saeedpanahian@usm.my

Zarita Zainuddin[2]
School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM Penang, Malaysia
Email: zarita@usm.my

Artificial neural networks has been effectively applied to numerous applications because of their universal approximation property. This work is grounded on two frameworks. Firstly, it is concerned with solving universal approximation problem by a class of neural networks based on Hankel approximate identity which is embedded in the space of continuous functions on real positive numbers. Secondly, this problem solving will be investigated in the Lebesgue spaces on real positive numbers. The methods are constructed on the notions of Hankel convolution linear operators, Hankel approximate identity, and epsilon-net.

ISCC 2015
18-19, December, 2015
Guangzhou, China

[1]We gratefully acknowledge Universiti Sains Malaysia for providing Postdoctoral Fellowship grant funding scheme to support our research.
[2]We thank the members of the organizing committee of ISCC 2015 for their sincere efforts and for their financial support. We also gratefully acknowledge Universiti Sains Malaysia for providing Postdoctoral Fellowship grant funding scheme to support our research.

PoS(ISCC2015)031

1. Introduction

It is well known that the problem of the universal approximation by feedforward neural networks has been solved in 1990s. Solving universal approximation problem by feedforward neural networks can be conceptually defined as an unlimited number of activation functions can theoretically approximate a continuous functions. Principe and Chen [1] surveyed the recent history of solving universal approximation problem by artificial neural networks.

On the other side, the study of the Hankel convolution linear operators was started by Zemanian [2]. Arteaga and Marrero [3], and Baddour [4] developed Hankel convolution theory. Recently, Hankel convolution linear operators have attracted many interests since these operators have many applications such as solving optic problems [5], radiation, diffraction, and field projection [6], and neutron transportation equation [7].

In this work, we are motivated to use Hankel convolution linear operators based on the Hankel approximate identity notion in order to construct a class of feedforward neural networks. Two aims are sought to be achieved in the present paper. The first aim is to apply a class of feedforward neural networks based on Hankel approximate identity to approximate continuous functions on real positive numbers. Subsequently, Lebesgue integrable functions on real positive numbers will be approximated by applying the same networks.

The approach of this work is as follows: we primary introduce the notion of Hankel approximate identity which is an extension of approximate identity. We apply Hankel approximate identity to prove the uniform convergence of a class of the Hankel convolution type operators in the space of continuous functions on real positive numbers. In the next step, we study the universal approximation by feedforward Hankel approximate identity neural networks in the space of continuous functions on real positive numbers. Moreover, we focus on the analysis of the uniform convergence of Hankel convolution type operators in the Lebesgue spaces on real positive numbers. Then, we prove the universal approximation by the above neural networks in the Lebesgue spaces on real positive numbers.

It is important to note that the advantage of this works is to construct some developments for theoretical research on solving universal approximation problem by feedforward artificial neural networks. The drawback of this work is the lack of numerical simulations.

The outline of this paper is as follows: Section 2 will focus on the introduction of Hankel approximate identity notion. Moreover, some basic definitions used in this paper will be presented in this section. In Section 3, we give two theoretical results concern with universal approximation by a class of feedforward neural networks based on Hankel approximate identity in the space of continuous functions on real positive numbers. In Section 4, we derive another two theoretical results in the Lebesgue spaces on real positive numbers. In Section 5, we give conclusion and future directions for research.

2. Notations and Definitions

Throughout this paper, the space of continuous functions on real positive numbers will be presented by $C(R_+)$. We also denote by $L_p(R_+)$, $1 \leq p < +\infty$, the Lebesgue spaces on real positive numbers. Before tackling our analyses, it is required to recall a number of basic notions. We start with the definition of the Hankel approximate identity.

Definition1: Let$\{\varphi_n(x)\}n\in N$, $\varphi_n(x): R+\rightarrow R$ be a Hankel approximate identity if the following properties hold:

1) $\int_R +\varphi_n(x)\, d\, x=1$;

2) let us assume that $\varepsilon>0$ and $\delta>0$, there is an N given that if $n\geq N$ then $\int_{x>\delta}(\varphi_n(x))\, d\, x\leq\varepsilon$.

Definition 2: The below equation is the definition of convolution in the Hankel sense

$f *_H \quad g$ of two given functions $f, g : R_{+\to} R$

$$f *_H g := \int_0^\infty f(u)g(xu)dx \tag{2.1}$$

Let us now state the epsilon-net notion in functions spaces.

Definition 3: [8] Let $\varepsilon > 0$. A set $V_\varepsilon \subset C(R_{+})$ is called epsilon-net of a set V in $C(R_{+})$, if $\tilde{f} \in V_\varepsilon$ can be found for $\forall \quad f \in V$ such that $(f - \tilde{f})_{C(R_+)} < \quad \varepsilon$.

Definition 4: [8] Let $\varepsilon > 0$. A set $V_\varepsilon \in L_p(R_+)$ is referred to as epsilon-net of a set V in $L_P \ (R_+)$ if $\tilde{f} \in V_\varepsilon$ can be obtained for $\forall \quad f \in V$ provided that

$$\|f - \tilde{f}\|_{L_P(\mathbb{R}_+)} < \varepsilon .$$

Definition 5: [8] If the epsilon-net includes a finite set of elements, it can be said to be finite epsilon-net.

3. Solving Universal Approximation by a Class of Neural Networks Based on Hankel Approximate Identity in the Space of Continuous Functions on Real Positive Numbers

This section gets Theorem 3.1. According to Theorem 3.1, Hankel convolution linear operators based on Hankel approximate identity have uniform convergence property in the space of continuous functions on real positive numbers.

Theorem 3.1 Suppose $C \ (R_+)$ be the space of continuous functions with a compact support on R_+. Suppose$\{\varphi_n(x)\}n \in \mathrm{N}$, $\varphi_n : R_{+\to} R$ be a Hankel approximate identity. Suppose f be a function in $C(R_{+})$. Then $\varphi_n *_H \ f$ uniformly converges to f on $C \ (R_+)$.

Proof. Let $x \in (0, \quad \delta)$ and $\varepsilon > 0$. There is a $\eta > 0$ provided that $(f(x) - f(y)) < \varepsilon$ for all $y, \quad (x - y) < \eta$. Then,

$$\varphi_n *_H f(x) - f(x) = \int_0^{+\infty} \varphi_n(y) f(xy) dy - f(x) \tag{3.1}$$

$$= \int_0^{+\infty} \varphi_n(y) f(xy)dy - f(x) \int_0^{+\infty} \varphi_n(y)dy$$

$$= \int_0^{+\infty} \varphi_n(y)\{f(xy)dy - f(x)\}\, dy \tag{3.2}$$

$$=(\int_0^\delta + \int_\delta^{+\infty}) \ \varphi_n(y)\{f(xy) - f(x)\}dy \tag{3.3}$$

$$= I_1 + I_2 \tag{3.4}$$

where I_1, I_2 are as follows:

$$(I_1) \le \int_0^\delta \varphi_n(y)(f(xy) - f(x))dy \tag{3.5}$$

$$\le \frac{\varepsilon}{2} \int_0^\delta \varphi_n(y)dy \tag{3.6}$$

$$\leq \frac{\varepsilon}{2} \int_{0}^{+\infty} \varphi_n(y)\,d\,y$$

(3.7)

$$\leq \frac{\varepsilon}{2}$$

(3.8)

For I_2, we have

$$|I_2| \leq 2\|f\|_{C(\mathbb{R}_+)} \int_{\delta}^{+\infty} |\varphi_n(y)|\,dy$$

(3.9)

Since

$$\lim_{n\to\infty} \int_{\delta}^{+\infty} |\varphi_n(y)|d\,y = 0, \quad (3.10)$$

there exists an $n_0 \in N$ such that for all $n \geq n_0$,

$$\int_{\delta}^{+\infty} |\varphi_n(y)|\,dy < \frac{\varepsilon}{2\|f\|_{C(\mathbb{R}_+)}}.$$

(3.11)

Combining I_1 and I_2 for $n \geq n_0$, we have

$$\|\varphi_n *_H f(x) - f(x)\|_{C(\mathbb{R}^d)} < \varepsilon.$$

(3.2)

Using Theorem 3.1, we immediately have Theorem 3.2 as the fundamental outcome of this section. For proving Theorem 3.2, we use the same method which was used by Wu et 'al.[9]. Theorem 3.2 illustrates that feedforward Hankel approximate identity neural networks have universal approximation property in the space of continuous functions on real positive numbers.

Theorem 3.2 Suppose $C(R_+)$ be the space of continuous functions with a compact support on R_+, and $V \subset C(R_+)$ a compact set. Suppose $\{\varphi_n(x)\}n \in \mathrm{N}$, $\varphi_n : R_+ \to R$ be a Hankel approximate identity. Suppose $\{ \sum_{j=1}^{M} \lambda_j \varphi_j(x) | \lambda_j \in R,\ x \in R_+,\ \mathrm{M} \in \mathrm{N}\}$ be dense in C(R_+), and given $\varepsilon > 0$. Then there is $N \in N$ which depends on V and ε but not on f, such that for any $f \in V$ there are weights $c_k = c_k(f, V, \varepsilon)$ satisfying

$$\left(f(x) - \sum_{i=1}^{N} c_k \varphi_k(x) \right)_{C(R_+)} < \varepsilon$$

(3.13)

Moreover, every c_k is a continuous function of $f \in V$.

Proof. For any assumed ε>0, there is a finite $\frac{\varepsilon}{2}$-net $\left(f^1,\dots f^M\right)$ for a compact set V. This indicates that for any $f\in V$, there is an f^j such that $\left(f-f^j\right)_{C(R_+)}<\frac{\varepsilon}{2}$. For any given f^j, considering Theorem 3.2 's assumption, there are $\lambda_i^j\in R$, $N_j\in N$, and $\varphi_i^j(x)$ such that

$$\left(f^j(x)-\sum_{i=1}^{N_j}\lambda_i^j\varphi_i^j(x)\right)_{C(R_+)}<\frac{\varepsilon}{2}. \tag{3.14}$$

For any given $f\in V$, we define

$$F-(f)=\{j\mid \left(f-f^j\right)_{C(R_+)}<\frac{\varepsilon}{2}\}, \tag{3.15}$$

$$F_0(f)=\{j\mid \left(f-f^j\right)_{C(R_+)}=\frac{\varepsilon}{2}\}, \tag{3.16}$$

$$F_{+(f)}=\{j\mid \left(f-f^j\right)_{C(R_+)}>\frac{\varepsilon}{2}\}. \tag{3.17}$$

As the result, based on the definition of $\frac{\varepsilon}{2}$-net, $F_{-}(f)$ is not empty. If $\tilde{f}\in V$ limits f provided that $\left(\tilde{f}-f\right)_{C(R_+)}$ is small, hence we get $F_{-(f)}\subset F_{-(\tilde{f})}$ and $F_{+(f)}\subset F_{+(\tilde{f})}$. Thus which

$$F_-(\tilde{f})\cap F_+(f)\subset F_-(\tilde{f})\cap F_+(\tilde{f})=\emptyset$$

indicates $F_{-(\tilde{f})}\subset F_{-(f)}\cup F_0(f)$. The ending part is the following result.

$$F_{-(f)}\subset F_{-(\tilde{f})}\subset F_{-(f)}\cup F_0(f). \tag{3.18}$$

Define

$$d(f)=\left(\sum_{j\in F_{-(f)}}\left(\frac{\varepsilon}{2}-\left(f-f^j\right)_{C(R_+)}\right)\right)^{-1} \tag{3.19}$$

and

$$f_h=\sum_{j\in F_{-(f)}}\sum_{i=1}^{N_j}d(f)\left(\frac{\varepsilon}{2}-\left(f-f^j\right)_{L_p(R^d)}\right)\lambda_i^j\varphi_i^j(x) \tag{3.20}$$

Then $f_h\in\{\sum_{j=1}^{M}\lambda_j\varphi_j(x)\}$ approximates f:

$$\left(f-f_h\right)_{C(R_+)} = \left\|\sum_{j\in F_-(f)}d(f)\left(\frac{\varepsilon}{2}-\left\|f-f^j\right\|_{C(\mathbb{R}_+)}\right)\left(f-\sum_{i=1}^{N_j}\lambda_i{}^j\varphi_i{}^j(x)\right)\right\|_{C(\mathbb{R}_+)} \tag{3.21}$$

$$=\left\| \sum_{j\in F_{-}(f)} d(f)\left(\frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R}_+)}\right)(f - f^j + f^j - \sum_{i=1}^{N_j} \lambda_i{}^j \varphi_i{}^j(x))\right\|_{C(\mathbb{R}_+)} \tag{3.22}$$

$$\leq \sum_{j\in F_{-}(f)} d(f)\left(\frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R}_+)}\right)\left(\|f - f^j\|_{C(\mathbb{R}_+)} + \left\|f^j - \sum_{i=1}^{N_j} \lambda_i{}^j \varphi_i{}^j(x)\right\|_{C(\mathbb{R}_+)}\right) \tag{3.23}$$

$$< \sum_{j\in F_{-(f)}} d(f)\left(\frac{\varepsilon}{2} - (f - f^j)_{C(R_+)}\right)\left(\frac{\varepsilon}{2} + \frac{\varepsilon}{2}\right) = \varepsilon. \tag{3.24}$$

Next, the following part is dedicated to the proving continuity of c_k. We apply (3.18) to get

$$\sum_{j\in F_{-(f)}} \left(\frac{\varepsilon}{2} - (\tilde{f} - f^j)_{C(R_+)}\right) \leq \sum_{j\in F_{-}(\tilde{f})} \left(\frac{\varepsilon}{2} - \|\tilde{f} - f^j\|_{C(\mathbb{R}_+)}\right) \tag{3.25}$$

$$\leq \sum_{j\in F_{-(f)}} \left(\frac{\varepsilon}{2} - (\tilde{f} - f^j)_{C(R_+)}\right) + \sum_{j\in F_0(f)} \left(\frac{\varepsilon}{2} - (\tilde{f} - f^j)_{C(R_+)}\right). \tag{3.26}$$

Suppose $\tilde{f} \to f$ in (3.26), we derive

$$\sum_{j\in F_{-(\tilde{f})}} \left(\frac{\varepsilon}{2} - (\tilde{f} - f^j)_{C(R_+)}\right) \to \sum_{j\in F_{-(f)}} \left(\frac{\varepsilon}{2} - (f - f^j)_{C(R_+)}\right) \tag{3.26}$$

This obviously shows $d(\tilde{f}) \to d(f)$. Hence, $\tilde{f} \to f$ results

$$d(\tilde{f})\left(\frac{\varepsilon}{2} - (\tilde{f} - f^j)_{C(R_+)}\right)\lambda_i^j \to d(f)\left(\frac{\varepsilon}{2} - (f - f^j)_{C(R_+)}\right)\lambda_i^j. \tag{3.28}$$

Suppose $N = \sum_{j\in F_{-(f)}} N_j$ and define c_k in terms of

$$f_h = \sum_{j\in F_{-}(f)} \sum_{i=1}^{N_j} d(f)\left(\frac{\varepsilon}{2} - \|f - f^j\|_{C(\mathbb{R}_+)}\right)\lambda_i{}^j \varphi_i{}^j(x) \tag{3.29}$$

$$\equiv \sum_{k=1}^{N} c_k\, \varphi_k(x) \tag{3.30}$$

From (30), c_k is continuous.

4. Solving Universal Approximation by a Class of Neural Networks based on Hankel Approximate Identity in The Lebesgue Spaces on Real Positive Numbers

In order to prove Theorem 4.2, Theorem 4.1 is provided.

Theorem 4.1: Assume $1 \leq p \leq +\infty$. If $f \in L_P (R_+)$then

$$lim_{x \to y} ||f(x) - f(y)||_{L_P(\mathbb{R}_+)} = 0 \tag{4.1}$$

Now, we can prove Theorem 4.2 that demonstrates that Hankel convolution linear operators based on Hankel approximate identity have uniform convergence property in the Lebesgue spaces on real positive numbers.

Theorem 4.2: Suppose $L_P(R_+)$ be the spaces of Lebesgue integrable functions with a compact support on R_+ . Suppose $\{ \varphi_n (x)\} n \in N$, $\varphi_n : R_+ \to R$ be a Hankel approximate identity. Suppose f be a function in $L_P (R_+)$ .Then $\varphi_n *_H f$ converges uniformly to f on $L_P (R_+)$.

Proof. Using generalized Minkowski inequality, we get the following relation:

$$||\varphi_n *_H f - f||_{L_P(\mathbb{R}_+)} \leq \int_{\mathbb{R}_+} ||f(xy) - f(x)||_{L_P(\mathbb{R}_+)} |\varphi_n(y)| dy. \tag{4.2}$$

$$||\varphi_n *_H f - f||_{L_P(\mathbb{R}_+)} \leq (\int_0^{\delta} + \int_{\delta}^{+\infty}) ||f(xy) - f(x)||_{L_P(\mathbb{R}_+)} |\varphi_n(y)| dy \tag{4.3}$$

Using Theorem 4.1, for any $\varepsilon > 0$, there is a $\eta > 0$ such that if $|xy - x| < \delta$,

$$||\varphi_n *_H f - f||_{L_P(\mathbb{R}_+)} \leq (\int_0^{\delta} + \int_{\delta}^{+\infty}) ||f(xy) - f(x)||_{L_P(\mathbb{R}_+)} |\varphi_n(y)| dy \quad . \tag{4.4}$$

Using triangular inequality, we conclude that

$$||f(xy) - f(x)||_{L_P(\mathbb{R}_+)} \leq 2||f||_{L_P(\mathbb{R}_+)}. \quad . \tag{4.5}$$

Applying the last inequalities (4.5) and (4.4) in inequality (4.3), the following relations obtained:

$$||\varphi_n *_H f - f||_{L_P(\mathbb{R}_+)} \leq \int_0^{\delta} \frac{\varepsilon}{2} |\varphi_n(y)| dy + \int_{\delta}^{+\infty} 2||f||_{L_P(\mathbb{R}_+)} |\varphi_n(y)| dy \tag{4.6}$$

$$\leq \frac{\varepsilon}{2} \int_0^{\delta} |\varphi_n(y)| dy + 2||f||_{L_P(\mathbb{R}_+)} \int_{\delta}^{+\infty} |\varphi_n(y)| dy \tag{4.7}$$

$$\leq \frac{\varepsilon}{2} \int_0^{+\infty} |\varphi_n(y)| dy + 2||f||_{L_P(\mathbb{R}_+)} \int_{\delta}^{+\infty} |\varphi_n(y)| dy \tag{4.8}$$

According to Definition1, there is an N provided that for $n \geq N$

$$\int_{\delta}^{+\infty} |\varphi_n(y)|dy \le \frac{\varepsilon}{4||f||_{L_P(\mathbb{R}_+)}}$$

(4.9)

Using inequality (4.9) in inequality (4.10), we get for $n \ge N$

$$||\varphi_n *_H f - f||_{L_P(\mathbb{R}_+)} \le \frac{\varepsilon}{2}\int_0^{+\infty} |\varphi_n(y)|dy + 2||f||_{L_P(\mathbb{R}_+)} \frac{\varepsilon}{4||f||_{L_P(\mathbb{R}_+)}}$$

(4.10)

$$\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(4.11)

We are now in a position to introduce Theorem 4.3 that illustrates that a class of feedforward neural networks based on Hankel approximate identity have universal approximation property in the Lebesgue spaces on real positive numbers.

Theorem 4.3: Suppose $L_P(R_{+})$ be the spaces of Lebesgue integrable functions on with a compact support on $R_{+,}$ and a $V \subset L_P(\mathbb{R}_+)$ compact set. Suppose $\{\varphi_n(x)\}$n∈N , $\varphi_n : R_{+\to} R$ be a Hankel approximate identity. Suppose the family of functions $\{\sum_{j=1}^{M} \lambda_j \varphi_j(x) | \lambda_j \in R, \quad x \in R^d$, M ∈ N be dense in $L_P(R_{+})$, and given $\varepsilon > 0$. Then there is an $N \in N,$ which depends on V and ε but not on f, such that for any $f \in V$, there are weights $c_k = c_k(f, V, \varepsilon)$ satisfying

$$\left(f(x) - \sum_{i=1}^{N} c_k \varphi_k(x) \right)_{L_P(R_+)} < \varepsilon$$

(4.12)

Moreover, every c_k is a continuous function of $f \in V$.

Proof. For any given ε>0, there is a finite $\frac{\varepsilon}{2}$-net $(f^1, \dots f^M)$ for a compact set V. This implies that for any $f \in V,$ there is an f^j such that $(f - f^j)_{L_P(R_+)} < \frac{\varepsilon}{2}$. For any f^j, by assumption of the theorem, there are $\lambda_i^j \in R, \quad N_j \in N$, and $\varphi_i^j(x)$ that

$$\left(f^j(x) - \sum_{i=1}^{N_j} \lambda_i^j \varphi_i^j(x) \right)_{L_P(R_+)} < \frac{\varepsilon}{2}. \quad (4.13)$$

For any given $f \in V,$ we set

$$F-(f) = \{ j \,|\, (f - f^j)_{L_P(R_+)} < \frac{\varepsilon}{2} \}, \quad (4.14)$$

$$F_0(f) = \{ j \,|\, (f - f^j)_{L_P(R_+)} = \frac{\varepsilon}{2} \}, \quad (4.15)$$

$$F_{+(f)} = \{ j \,|\, (f - f^j)_{L_P(R_+)} > \frac{\varepsilon}{2} \}. \quad (4.16)$$

As the result, based on the definition of $\frac{\varepsilon}{2}$ –net, $F_{-}(f)$ is not empty set. If $\tilde{f} \in V$ limits f provided that $(\tilde{f} - f)_{L_P(R_+)}$ is small, then we derive $F_{-(f)} \subset F_{-(\tilde{f})}$ and $F_{+(f)} \subset F_{+(\tilde{f})}$.

Hence, $F_-(\tilde{f}) \cap F_+(f) \subset F_-(\tilde{f}) \cap F_+(\tilde{f}) = \emptyset,$

which indicates $F_{-(\tilde{f})} \subset F_{-(f)} \cup F_0(f)$. The ending part is the following relation.

$$F_{-(f)} \subset F_{-(\tilde{f})} \subset F_{-(f)} \cup F_0(f). \tag{4.17}$$

Define

$$d(f) = \left(\sum_{j \in F_{-(f)}} \left(\frac{\varepsilon}{2} - (f - f^j)_{L_p(R_+)} \right) \right)^{-1} \tag{4.18}$$

and

$$f_h = \sum_{j \in F_{-(f)}} \sum_{i=1}^{N}{}_j \, d(f) \left(\frac{\varepsilon}{2} - (f - f^j)_{L_p(R_+)} \right) \lambda_i^j \varphi_i^j(x) \tag{4.19}$$

Then $f_h \in \{ \sum_{j=1}^{M} \lambda_j \varphi_j(x) \}$ approximates f: $(f - f_h)_{L_p(R_+)}$

$$= \left(\sum_{j \in F_{-(f)}} d(f) \left(\frac{\varepsilon}{2} - (f - f^j)_{L_p(R_+)} \right) \left(f - \sum_{i=1}^{N}{}_j \lambda_i^j \varphi_i^j(x) \right) \right)_{L_p(R_+)} \tag{4.20}$$

$$= \left\| \sum_{j \in F_{-}(f)} d(f) \left(\frac{\varepsilon}{2} - \| f - f^j \|_{L_P(\mathbb{R}_+)} \right) \left(f - f^j + f^j - \sum_{i=1}^{N_j} \lambda_i{}^j \varphi_i{}^j(x) \right) \right\|_{L_P(\mathbb{R}_+)} \tag{4.21}$$

$$\leq \sum_{j \in F_{-}(f)} d(f) \left(\frac{\varepsilon}{2} - \| f - f^j \|_{L_P(\mathbb{R}_+)} \right) \left(\| f - f^j \|_{L_P(\mathbb{R}_+)} + \left\| f^j - \sum_{i=1}^{N_j} \lambda_i{}^j \varphi_i{}^j(x) \right\|_{L_P(\mathbb{R}_+)} \right) \tag{4.22}$$

$$< \sum_{j \in F_{-(f)}} d(f) \left(\frac{\varepsilon}{2} - (f - f^j)_{L_p(R_+)} \right) \left(\frac{\varepsilon}{2} + \frac{\varepsilon}{2} \right) \tag{4.23}$$

$$= \varepsilon. \tag{4.24}$$

Next, the following part is dedicated to the proving the continuity of c_k. We use (4.17) to derive $\sum_{j \in F_{-(f)}} \left(\frac{\varepsilon}{2} - (\tilde{f} - f^j)_{L_p(R_+)} \right)$

$$\sum_{j \in F_{-}(f)} \left(\frac{\varepsilon}{2} - \| \tilde{f} - f^j \|_{L_P(\mathbb{R}_+)} \right)$$

$$\leq \sum_{j \in F_{-}(\tilde{f})} \left(\frac{\varepsilon}{2} - \| \tilde{f} - f^j \|_{L_P(\mathbb{R}_+)} \right) \tag{4.25}$$

$$\leq \sum_{j\in F_-(f)}\left(\frac{\varepsilon}{2}-\|\tilde{f}-f^j\|_{L_P(\mathbb{R}_+)}\right)+$$

$$\sum_{j\in F_0(f)}\left(\frac{\varepsilon}{2}-\|\tilde{f}-f^j\|_{L_P(\mathbb{R}_+)}\right). \tag{4.27}$$

Let $\tilde{f}\to f$ in (4.27), we get

$$\sum_{j\in F_{-(\tilde{f})}}\left(\frac{\varepsilon}{2}-\left(\tilde{f}-f^j\right)_{L_p(R_+)}\right)\to \sum_{j\in F_{-(f)}}\left(\frac{\varepsilon}{2}-\left(f-f^j\right)_{L_p(R_+)}\right) \tag{4.28}$$

Apparently, this ascertain that $d\left(\tilde{f}\right)\to d\left(f\right)$. Thus, $\tilde{f}\to f$ results

$$d\left(\tilde{f}\right)\left(\frac{\varepsilon}{2}-\left(\tilde{f}-f^j\right)_{L_p(R_+)}\right)\lambda_i^j\to d\left(f\right)\left(\frac{\varepsilon}{2}-\left(f-f^j\right)_{L_p(R_+)}\right)\lambda_i^j. \tag{4.29}$$

Let $N=\sum_{j\in F_{-(f)}} N_j$ and define c_k in terms of

$$f_h=\sum_{j\in F_-(f)}\sum_{i=1}^{N_j} d(f)\left(\frac{\varepsilon}{2}-\|f-f^j\|_{L_P(\mathbb{R}_+)}\right)\lambda_i{}^j\varphi_i{}^j(x) \tag{4.30}$$

$$\equiv \sum_{k=1}^{N} c_k\,\varphi_k(x) \tag{4.31}$$

From (4.29), c_k is continuous.

5. Conclusion

Hankel approximate identity notion has been constructed. Then, the theoretical framework has been discussed in two directions. First, we have shown that how Hankel convolution linear operators can be used to driven analysis of the universal approximation by a class of feedforward neural networks based on Hankel approximate identity in the space of continuous functions on real positive numbers. Second, we have indicated that how to obtain universal approximation by thenetworks in the Lebesgue spaces on real positive numbers. The results may be highlighted, if we provide the experimental evaluation about the proposed theorems in our further work.

References

[1] Principe J C, Chen B (2015) *Universal approximation with convex optimization:gimmic or reality?*. IEEE Computational Intelligence Magazine vol. 10, pp. 68-77

[2] Zemanian A H (1966) *A distributional Hankel transformation*, SIAM J. Appl. Math, vol. 14, 561-576

[3]Arteaga C, Marrero I (2014) *Direct from seminorms arising in the theory of interpolation by Hankel translates of a basis function*, Advances in computational Mathematics, vol. 183, 167-183

[4]Baddour N (2014) *Application of the generalized shift operator to the Hankel transform*, Springer Plus, vol. 3, 246

[5]Baddour N, Chouinard, U (2015) *Theory and operational rules for the discrete Hankel Transform*. Optical Society of America, vol. 32, pp. 611-622

[6]Pandy, R K, Singh, V K, Singh, O P (2009) *An improved method for computing Hankel transform*. Journal of the Franklin Institute, vol. 364, pp. 102-111

[7]Fernandes J C L, Vilhena M T, Bodman B E J (2013) *On a comparative analysis of the solution of kinetic neutron diffusion equation by the Hankel transform formalism and the spectral method*. Progress in Nuclear energy, vol. 69, pp. 71-76

[8]Lebedev V I (1997) *An introduction to functional analysis in computational mathematics*. Brikh¨auser Boston

[9]Wu W, Nan D, Li Z, Long J, and Wang J (2007) *Approximation to compact set of functions by feedforward neural networks*, Proceedings 20th International Joint Conference on Neural Networks, pp. 1222-1225,