Chapter 23
On the Universal Approximation Capability of Flexible Approximate Identity Neural Networks

Saeed Panahian Fard and Zarita Zainuddin

Abstract This study presents some class of feedforward neural networks to investigate the universal approximation capability of continuous flexible functions. Based on the flexible approximate identity, some theorems are constructed. The results are provided to demonstrate the universal approximation capability of flexible approximate identity neural networks to any continuous flexible function.

Keywords Flexible approximate identity · Flexible approximate identity activation functions · Flexible approximate identity neural networks · Uniform convergence · Universal approximation

23.1 Introduction

One of the most important issues in theoretical studies for neural networks is concerned with the universal approximation capability of feedforward neural networks. There have been many papers related to this topic over the past 30 years [1].

A few authors [2–5] recently deal with the concept of approximation of non-linear functions by approximate identity neural networks (AINNs). These networks are based on the widely-known sequences of functions named approximate identities [6].

Flexible approximate identity neural networks (FAINNs) are the generalization of AINNs. These networks use flexible approximate identity as activation functions with a traditional multilayer architecture. Lately, new model of feedforward neural networks called the generalized Gaussian radial basis function neural networks has been proposed which is shown in [7]. These neural networks are special case of the FAINNs.

S. Panahian Fard (✉) · Z. Zainuddin
School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Georgetown, Pulau Penang, Malaysia
e-mail: saeedpanahian@yahoo.com

W. E. Wong and T. Ma (eds.), *Emerging Technologies for Information Systems, Computing, and Management*, Lecture Notes in Electrical Engineering 236, DOI: 10.1007/978-1-4614-7010-6_23,

The main goal of this study is to investigate the universal approximation capability of FAINNs to any continuous flexible function. Based on a convolution linear operator in the real linear space of all continuous flexible functions, some theorems are presented. These theorems verify the approximation capability of FAINNs.

This paper is organized as follows. In Sect. 23.2, as the main technical tool, the definition of flexible approximate identity is given. And basic definitions and theorems are introduced. The main result is presented in Sect. 23.3. Conclusions are drawn in Sect. 23.4.

23.2 Preliminaries

The definition of flexible approximate identity which will be used in Theorem 1 is presented as follows.

Definition 1 Let $A = A(a_1, \ldots, a_m), a_i \in \mathbb{R}, i = 1, \ldots, m$ be any parameters. $\{\varphi_n(x, A)\}_{n \in N}, \varphi_n(x, A) : R \to R$ is said to be a flexible approximate identity if the following properties hold:

(1) $\int\limits_{R} \varphi_n(x, A)dx = 1;$

(2) Given ε and $\delta > 0$, there exists N such that if $n \geq N$ then

$$\int\limits_{|x| > \delta} |\varphi_n(x, A)|dx \leq \varepsilon.$$

Now, we will be able to give the following theorem in order to construct the hypothesis of Theorem 3 in the next section.

Theorem 1 *Let* $\{\varphi_n(x, A)\}_{n \in N}, \varphi_n(x, A) : \mathrm{R} \to \mathrm{R}$ *be a flexible approximate identity. Let* f *be a function on* $C[a, b]$. *Then* $\varphi_n * f$ *uniformly converges to* f *on* $C[a, b]$.

Proof Let $x \in [a, b]$ and $\varepsilon > 0$. There exists a $\delta > 0$ such that $|f(x) - f(y)| < \frac{\varepsilon}{2\|\varphi\|_1}$ for all $y, |x - y| < \delta$. Let us define $\{\varphi_n * f\}_{n \in N}$ by $\varphi_n(x, A) = n\varphi(nx, A)$. Then,

$$\begin{aligned} \varphi_n * f(x) - f(x) &= \int\limits_{R} n\varphi(ny, A)\{ f(x - y) - f(x)\} \, dy \\ &= \left(\int\limits_{|y| < \delta} + \int\limits_{|y| < \delta} \right) n\varphi(ny, A)\{f(x - y) - f(x)\}dy \\ &= I_1 + I_2, \end{aligned}$$

where $I_1 + I_2$ are as follows:

$$\begin{aligned}|I_1| &\leq \int\limits_{|y|<\delta} n\varphi(ny, A)\{f(x-y) - f(x)\}dy \\ &< \frac{\varepsilon}{2||\phi||_1} \int\limits_{|y|<\delta} n\varphi(ny, A)dy \\ &= \frac{\varepsilon}{2||\phi||_1} \int\limits_{|t|<n\delta} \varphi(t, A)dt \\ &\leq \frac{\varepsilon}{2||\phi||_1} \int\limits_{R} \varphi(t, A)dt = \frac{\varepsilon}{2}.\end{aligned}$$

For I_2, we have

$$\begin{aligned}|I_2| &\leq 2||f||_{C[a,b]} \int\limits_{|y|\geq\delta} n|\varphi(ny, A)|dy \\ &= 2||f||_{C[a,b]} \int\limits_{|t|\geq n\delta} |\varphi(t, A)|dt.\end{aligned}$$

Since

$$\lim_{n\to+\infty} \int\limits_{|t|>n\delta} |\varphi(t, A)|dt = 0,$$

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

$$\int\limits_{|t|\geq n\delta} |\varphi(t, A)|dt < \frac{\varepsilon}{4||f||_{C[a,b]}}.$$

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

$$||\varphi_n * f(x) - f(x)||_{C[a,b]} < \varepsilon.$$

We use the following (cf. [8]) in order to prove Theorem 3 which is given as the main result in the Sect. 23.3:

Definition 2 Let $\varepsilon > 0$. A set $V_\varepsilon \subset C[a, b]$ is called ε-net of a set V, if $\tilde{f} \in V_\varepsilon$ can be found for $\forall f \in V$ such that $\left\|f - \tilde{f}\right\|_{C[a,b]} < \varepsilon$.

Definition 3 The ε-net is said to be finite if it is a finite set of elements.

Theorem 2 *A set V in $C[a, b]$ is compact iff $\forall$ $\varepsilon > 0$ in R there is a finite ε-net.*

Now, we present the universal approximation capability of FAINNs in the next section.

23.3 Main Result

The main aim of this section is to investigate the conditions for the universal approximation capability of FAINNs to any continuous flexible function. Now, the following theorem is proposed to show the universal approximation capability of FAINNs.

Theorem 3 *Let* $C[a,b]$ *be linear space of all continuous functions on the real interval* $[a,b]$, and $V \subset C[a,b]$ *a compact set. Let* $A = A(a_1,\ldots,a_m), a_i > 0, i = 1,\ldots,m$ *be any parameters,* $\{\varphi_n(x,A)\}_{n\in N}, \varphi_n(x,A) : \mathrm{R} \to \mathrm{R}$ *be a flexible approximate identity. Let the family of functions* $\left\{\sum_{\mathrm{j}=1}^{\mathrm{M}} \lambda_{\mathrm{j}}\varphi_{\mathrm{j}}(\mathrm{x},\mathrm{A}) | \lambda_{\mathrm{j}} \in \mathrm{R}, \mathrm{x} \in \mathrm{R}, \mathrm{M} \in \mathrm{N}\right\}$, *be dense in* $C[a,b]$, *and given* $\varepsilon > 0$. *Then there exists* $\mathrm{N} \in \mathrm{N}$ *which depends on* V *and* ε *but not on* f, *such that for any* $f \in V$, *there exist weights* $c_k = c_k(f, V, \varepsilon)$ *satisfying*

$$\left\| f(x) - \sum_{i=1}^{N} c_k\varphi_k(x,A) \right\|_{C[a,b]} < \varepsilon$$

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

Proof The method of proof is analogous to that of Theorem 1 in [9]. Because V is compact, for any $\varepsilon > 0$, there is a finite $\frac{\varepsilon}{2}$-net $\{f^1,\ldots,f^M\}$ for V. This implies that for any $f \in V$, there is an f^j such that $\|f - f^j\|_{C[a,b]} < \frac{\varepsilon}{2}$. For any f^j, by assumption of the theorem, there are $\lambda_i^j \in R, N_j \in \mathrm{N}$, and $\varphi_i^j(x,A)$ such that

$$\left\| f^j(x) - \sum_{i=1}^{N_j} \lambda_i^j\varphi_i^j(x,A) \right\|_{C[a,b]} < \frac{\varepsilon}{2}. \tag{23.1}$$

For any $f \in V$, we define

$$F_-(f) = \left\{ j \middle| \|f - f^j\|_{C[a,b]} < \frac{\varepsilon}{2} \right\},$$

$$F_0(f) = \left\{ j \middle| \|f - f^j\|_{C[a,b]} = \frac{\varepsilon}{2} \right\},$$

$$F_+(f) = \left\{ j \middle| \|f - f^j\|_{C[a,b]} > \frac{\varepsilon}{2} \right\}.$$

Therefore, $F_-(f)$ is not empty according to the definition of $\frac{\varepsilon}{2}$ -net. If $\tilde{f} \in V$ approaches f such that $\|f - f^j\|_{C[a,b]}$ is small enough, then we have $F_-(f) \subset F_-(\tilde{f})$ and $F_+(f) \subset F_+(\tilde{f})$. Thus $F_-(\tilde{f}) \cap F_+(f) \subset F_-(\tilde{f}) \cap F_+(\tilde{f}) = \emptyset$, which implies $F_-(\tilde{f}) \subset F_-(f) \cup F_0(f)$. We finish with the following.

$$F_-(f) \subset F_-(\tilde{f}) \subset F_-(f) \cup F_0(f). \tag{23.2}$$

Define

$$\begin{aligned} d(f) &= \left[\sum\nolimits_{j \in F_-(f)} \left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]}\right)\right]^{-1} \quad \text{and} \\ f_h &= \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f)\left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]}\right) \lambda_i^j \varphi_i^j(x, A) \end{aligned} \tag{23.3}$$

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

$$\begin{aligned} &\|f - f_h\|_{C[a,b]} \\ &= \left\| \sum_{j \in F_-(f)} d(f)\left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]}\right)\left(f - \sum_{i=1}^{N_j} \lambda_i^j \varphi_i^j(x, A)\right) \right\|_{C[a,b]} \\ &= \left\| \sum_{j \in F_-(f)} d(f)\left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]}\right)\left(f - f^j + f^j - \sum_{i=1}^{N_j} \lambda_i^j \varphi_i^j(x, A)\right) \right\|_{C[a,b]} \\ &\leq \sum_{j \in F_-(f)} d(f)\left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]}\right)\left(\|f - f^j\|_{C[a,b]} + \left\| f^j - \sum_{i=1}^{N_j} \lambda_i^j \varphi_i^j(x, A) \right\|_{C[a,b]}\right) \\ &< \sum_{j \in F_-(f)} d(f)\left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]}\right)\left(\frac{\varepsilon}{2} + \frac{\varepsilon}{2}\right) = \varepsilon. \end{aligned} \tag{23.4}$$

In the following step, We prove the continuity of c_k. For the proof, we use (23.2) to obtain

$$\begin{aligned} &\sum_{j \in F_-(f)} \left(\frac{\varepsilon}{2} - \|\tilde{f} - f^j\|_{C[a,b]}\right) \\ &\leq \sum_{j \in F_-(\tilde{f})} \left(\frac{\varepsilon}{2} - \|\tilde{f} - f\|_{C[a,b]}\right) \\ &\leq \sum_{j \in F_-(\tilde{f})} \left(\frac{\varepsilon}{2} - \|\tilde{f} - f^j\|_{C[a,b]}\right) \\ &+ \sum_{j \in F_0(f)} \left(\frac{\varepsilon}{2} - \|\tilde{f} - f^j\|_{C[a,b]}\right). \end{aligned} \tag{23.5}$$

Let $f \to \tilde{f}$ in (23.5), then we have

$$\sum_{j \in F_-(\tilde{f})} \left(\frac{\varepsilon}{2} - \|\tilde{f} - f^j\|_{C[a,b]} \right) \to \sum_{j \in F_-(f)} \left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]} \right) \tag{23.6}$$

This obviously demonstrates $d(\tilde{f}) \to d(f)$. Thus, $\tilde{f} \to f$ results

$$d(\tilde{f}) \left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]} \right) \lambda_i^j \to d(f) \left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]} \right) \lambda_i^j. \tag{23.7}$$

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

$$\begin{aligned} f_h &= \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f) \left(\frac{\varepsilon}{2} - \|f - f^j\|_{C[a,b]} \right) \lambda_i^j \varphi_i^j(x, A) \\ &\equiv \sum_{k=1}^{N} c_k \varphi_k(x, A) \end{aligned}$$

From (23.7), c_k is a continuous functional of f. This completes the proof.

23.4 Conclusion

Some class of feedforward neural networks with a traditional multilayer architecture has been constructed to obtain an approximation of any flexible continuous function. By employing the flexible approximate identity, Theorem 1 is established. This theorem constructs the hypothesis for Theorem 3. In Theorem 3, it has been proved that if a flexible approximate identity neural networks with a hidden layer is dense in $C\,[a,b]$, then for a given compact set $V \subset C\,[a,b]$ and an error bound ε, one can approximate any continuous flexible function $f \in V$ with the accuracy ε.

References

1. Ismailov, V.E.: Approximation by neural networks with weights varying on a finite set of directions. J. Math. Anal. Appl. **398**, 72–83 (2012)
2. Hahm, N., Hong, B.I.: An approximation by neural networks with a fixed weight. Comput. Math. Appl. **47**, 1897–1903 (2004)
3. Li, F.: Function approximation by neural networks. In: Proceedings 5th International Symposium on Neural Networks, pp. 348–390 (2008)
4. Turchetti, C., Conti, M., Crippa, P., Orcioni, S.: On the approximation of stochastic processes by approximate identity neural networks. IEEE Trans. Neural Netw. **9**, 1069–1085 (1998)
5. Zainuddin, Z., PanahianFard, S.: Double approximate identity neural networks universal approximation in real Lebesgue spaces. Lect. Notes Comput. Sci. **7663**, 409–415 (2012)

6. Wheeden, R.L., Zygmund, A.: Measure and Integral. Marcel Dekker, New York (1977)
7. Navaro, F.F., Martínez, C.H., Monederi, J.S., Gutiérrez, P.A.: MELM-GRBF: a modified version of the extreme learning machine for generalized radial basis function neural networks. Neurocomputing **74**, 2502–2510 (2011)
8. Lebedev, V.: An Introduction to Functional Analysis and Computational Mathematics. Brikhäuser, Boston (1997)
9. Wu, W., Nan, D., Li, Z., Long, J., Wang, J.: Approximation to compact set of functions by feedforward neural networks. In: Proceedings 20th International Joint Conference on Neural Networks, pp. 1222–1225 (2007)