J. Appl. Math. & Informatics Vol. **26***(2008), No. 1 - 2, pp. 317 - 324*
Website: http://www.kcam.biz

A STUDY OF SIMULTANEOUS APPROXIMATION BY NEURAL NETWORKS

N. HAHM[†] AND B. I. HONG[*]

ABSTRACT. This paper shows the degree of simultaneous neural network approximation for a target function in $C^r[-1,1]$ and its first derivative. We use the Jackson's theorem for differentiable functions to get a degree of approximation to a target function by algebraic polynomials and trigonometric polynomials. We also make use of the de La Vallèe Poussin sum to get an approximation order by algebraic polynomials to the derivative of a target function. By showing that the divided difference with a generalized translation network can be arbitrarily closed to algebraic polynomials on $[-1,1]$, we obtain the degree of simultaneous approximation.

AMS Mathematics Subject Classification : 41A10, 41A24, 41A29
Key words and phrases : Simultaneous Approximation, Approximation Order, Generalized Translation Network

1. Introduction

This paper is related to the problem of simultaneous approximation by a three layered feedforward neural network. A three layered feed forward network has an input layer, a hidden layer and an output layer. A feedforward network with one hidden layer is of the form

$$\sum_{i=1}^{n} c_i \psi(a_i x + b_i) \tag{1.1}$$

where the weight a_i, the threshold b_i and c_i are real numbers for $1 \le i \le n$ and ψ is an activation function. For a natural number n, $\Psi_{\psi,n}$ denotes the set of all such functions.

Received April 12, 2007. [*] Corresponding author.
[†]This research was supported by the University of Incheon Research Fund, 2006.

Most papers related to the neural network approximation investigate the density problems and complexity problem. The density problem was proved in the many papers [5, 6]. In [5], Leshno, Lin, Pinkus and Schocken proved a complete characterization of the activation function. Under some weak conditions for an activation function, any continuous functions on a compact set can be approximated arbitrarily closed by neural networks.

Recently, the complexity problem was discussed in many papers [3, 7, 11]. Those papers investigated the relation between the number of neurons and its approximation capacity.

More recently, the simultaneous approximation of a function and its derivatives was also investigated by many researchers [1, 6, 12]. In [6], Li proved that a function and its derivatives on a compact set can be simultaneously approximated by neural networks. On the other hand, Cao, Xu and Li [1] proved the degree of simultaneous approximation by neural networks in the pointwise sense, but the pointwise convergence is not so useful in the applications. In this paper, we show the degree simultaneous approximation by neural networks in the uniform sense.

2. Preliminaries

Throughout the paper, n, m, r and k are natural numbers. For each n, we denote the class of all algebraic polynomials of degree not exceeding n by Π_n and the class of all trigonometric polynomials of degree not exceeding n by Π_n^*. In each case, we define

$$E_n(f) := \inf_{P\in\Pi_n} ||f - P||_{(\infty,[-1,1])} \tag{2.1}$$

and

$$E_n^*(g) := \inf_{P^*\in\Pi_n^*} ||g - P^*||_{(\infty,[-\pi,\pi])} \tag{2.2}$$

for $f \in C[-1,1]$ and a 2π periodic function $g \in C[-\pi,\pi]$, respectively. Jackson's theorems for differentiable functions are the followings [2].

Theorem 2.1. (1) *For $f \in C^r[-1,1]$ with $||f^{(r)}||_{(\infty,[-1,1])} < \infty$,*

$$E_n(f) \le \frac{C_r}{n^r}$$

where the constant C_r is independent of n.

(2) *For a 2π periodic function $g \in C^r[-\pi,\pi]$ with $||g^{(r)}||_{(\infty,[-\pi,\pi])} < \infty$,*

$$E_n^*(g) \le \frac{C_r^*}{n^r}$$

where the constant C_r^* is independent of n.

Note that $g' \in C^{r-1}[-\pi, \pi]$ for $g \in C^r[-\pi, \pi]$ and so we get

$$E_n^*(g') \leq \frac{C_r^{**}}{n^{r-1}}. \tag{2.3}$$

For a 2π periodic function $g \in C[-\pi, \pi]$, $S_n(g,t) := S_n(t)$ denotes the nth partial sum of its Fourier series. For each n, the de La Vallèe Poussin sum τ_n is defined by

$$\tau_n(g,t) := \tau_n(g) = \frac{1}{n} \sum_{i=n}^{2n-1} S_i(t).$$

Note that the de La Vallèe Poussin sum τ_n of a 2π periodic function g has the following properties [8, 9].

(1) $|\tau_n(g,t)| \leq \max |g(t)|$ for $t \in [-\pi, \pi]$.

(2) $\|\tau_n(g) - g\|_{(\infty,[-\pi,\pi])} \leq 4E_n^*(g)$.

(3) $\tau_n'(g,t) = \tau_n(g', t)$.

In addition, $\tau_n(T_k, t) = T_k(t)$ for any trigonometric polynomial $T_k(t)$ for $k \leq n$.

3. Main results

First of all, we get the degree of simultaneous approximation by algebraic polynomials. The basic idea comes from Theorem 2.1.

Theorem 3.1. *Assume that $f \in C^r[-1,1]$ with $\|f^{(r)}\|_{(\infty,[-1,1])} < \infty$. For each n, there is an algebraic polynomial $P_{2n-1} \in \Pi_{2n-1}$ such that*

$$\|f - P_{2n-1}\|_{(\infty,[-1,1])} \leq \frac{C^*}{n^r} \quad \text{and} \quad \|f' - P_{2n-1}'\|_{(\infty,[-1,1])} \leq \frac{C^{**}}{n^{r-1}} \tag{3.1}$$

where C^* and C^{**} are the positive constants which are independent of n.

Proof. In [9], there exists an extension $\tilde{f} \in C^r$ of f on $[-2,2]$ such that $\tilde{f}(x) = f(x)$ for $x \in [-1,1]$. For $x \in [-2,2]$, we define

$$g(t) := \tilde{f}(2\cos t) = \tilde{f}(x).$$

Then $g \in C^r[-\pi, \pi]$. From the elementary differentiation

$$\tilde{f}'(x) = \frac{d\tilde{f}(x)}{dx} = \frac{d\tilde{f}(2\cos t)}{dt} \cdot \frac{dt}{dx} = g'(t) \cdot \frac{1}{-2\sin t},$$

we obtain, for some C,

$$||\tilde{f}'||_{(\infty,[-1,1])} = \left\|f'\right\|_{(\infty,[-1,1])} = \left\|g' \cdot \frac{1}{-2\sin t}\right\|_{(\infty,[\frac{\pi}{3},\frac{2\pi}{3}])} \leq C||g'||_{(\infty,[-\pi,\pi])}. \tag{3.2}$$

By (2.3) and the properties (2) and (3) of the de La Vallèe Poussin sum, we have

$$\begin{aligned} ||g' - \tau_n'(g)||_{(\infty,[-\pi,\pi])} &= ||g' - \tau_n(g')||_{(\infty,[-\pi,\pi])} \\ &\leq C_1 E_n^*(g') \\ &\leq \frac{C^{**}}{n^{r-1}}. \end{aligned} \tag{3.3}$$

We define the algebraic polynomial P_{2n-1} by

$$P_{2m-1}(\tilde{f}, x) := P(\tilde{f}, 2\cos t) = \tau_n(g, t).$$

Then $P_{2n-1} \in \Pi_{2n-1}$. From Theorem 2.1 and the property (2) of the de La Vallèe Poussin sum,

$$\begin{aligned} ||f - P_{2n-1}||_{(\infty,[-1,1])} &\leq ||\tilde{f} - P_{2n-1}||_{(\infty,[-2,2])} \\ &= ||g - \tau_n(g)||_{(\infty,[-\pi,\pi])} \\ &\leq C_2 E_n^*(g) \\ &\leq \frac{C^*}{n^r}. \end{aligned} \tag{3.4}$$

In addition, we get the following result from (3.3).

$$\begin{aligned} ||f' - P_{2n-1}'||_{(\infty,[-1,1])} &\leq ||g' - \tau_n'(g)||_{(\infty,[-\pi,\pi])} \\ &\leq C_3 E_n^*(g') \\ &\leq \frac{C^{**}}{n^{r-1}}. \end{aligned} \tag{3.5}$$

This completes the proof. □

Now, we show that any polynomials can be simultaneously approximated by neural networks.

Theorem 3.2. *Suppose that $\psi \in C^\infty(\mathbb{R})$ and there exists $b \in \mathbb{R}$ in some open interval $(b - \delta, b + \delta)$ in $\mathbb{R}$ such that*

$$\psi^{(k)}(b) \neq 0$$

for any k. For $f \in C^r[-1,1]$ with $\|f^{(r)}\|_{(\infty,[-1,1])} < \infty$ and a natural number k, there exists a neural network $N_{k,h} \in \Psi_{\phi,k}$ such that

$$\|x^k - N_{k,h}\|_{(\infty,[-1,1])} \le M^* \cdot h \quad \text{and} \quad \left\|(x^k)' - N'_{k,h}\right\|_{(\infty,[-1,1])} \le M^{**} \cdot h \tag{3.6}$$

where M^* and M^{**} are positive constants.

Proof. We follow the idea in [7]. For each natural number k, the neural network

$$N_{k,h}(\phi,x) := \frac{1}{\phi^{(k)}(b)} \frac{1}{h^k} \sum_{j=0}^{k} (-1)^{(k-j)} \binom{k}{j} \phi(hjx+b) \tag{3.7}$$

represents a divided difference for x^k. Thus, as it is proved in [7],

$$\|G_{k,h}(\phi,x) - x^k\|_{(\infty,[-1,1])} \le M^* \cdot h \tag{3.8}$$

for some positive constant M^*. Note that

$$\begin{aligned} &N'_{k,h}(\phi,x) \\ &= \frac{1}{\phi^{(k)}(b)} \frac{1}{h^k} \sum_{j=0}^{k} (-1)^{(k-j)} \binom{k}{j} \phi'(hjx+b)hj \\ &= \frac{1}{\phi^{(k)}(b)} \frac{k}{h^{k-1}} \sum_{j=0}^{k} (-1)^{(k-j)} \binom{k-1}{j-1} \phi'(hjx+b) \\ &= \frac{1}{\phi^{(k)}(b)} \frac{k}{h^{k-1}} \sum_{j=0}^{k-1} (-1)^{(k-j-1)} \binom{k-1}{j} \phi'(h(j+1)x+b). \end{aligned} \tag{3.9}$$

Using Taylor theorem for an integer m with $m > k-1$, the equation (3.9) can be written as

$$\begin{aligned} \phi'(h(j+1)x+b) = \phi'(hjx+b) + \sum_{i=1}^{m-1} \frac{\phi^{(i+1)}(hjx+b)}{i!}(hx)^i \\ + \frac{\phi^{(m+1)}(hjx+b+\xi)}{m!}(hx)^m \end{aligned} \tag{3.10}$$

where ξ is a value between $hjx+b$ and $h(j+1)x+b$. Thus, the equation (3.9) can be rewritten as

$$\begin{aligned} &N'_{k,h}(\phi,x) \\ &= \frac{1}{\phi^{(k)}(b)} \frac{k}{h^{k-1}} \sum_{j=0}^{k-1} (-1)^{(k-j-1)} \binom{k-1}{j} \phi'(hjx+b) \\ &= \frac{1}{\phi^{(k)}(b)} \frac{k}{h^{k-1}} \sum_{j=0}^{k-1} (-1)^{(k-j-1)} \binom{k-1}{j} \Big(\sum_{i=1}^{m-1} \frac{\phi^{(i+1)}(hjx+b)}{i!}(hx)^i \Big) \\ &= \frac{1}{\phi^{(k)}(b)} \frac{k}{h^{k-1}} \sum_{j=0}^{k-1} (-1)^{(k-j-1)} \binom{k-1}{j} \Big(\frac{\phi^{(m+1)}(hjx+b+\xi)}{m!}(hx)^m \Big). \end{aligned} \tag{3.11}$$

The second part of (3.11) is

$$\frac{1}{\phi^{(k)}(b)}\frac{k}{h^{k-1}}\sum_{j=0}^{k-1}(-1)^{(k-j-1)}\binom{k-1}{j}\phi'(hjx+b)=k\cdot N_{k-1,h}(\phi',x). \quad (3.12)$$

The third part of (3.11) is

$$\begin{aligned}&\frac{1}{\phi^{(k)}(b)}\frac{k}{h^{k-1}}\sum_{j=0}^{k-1}(-1)^{(k-j-1)}\binom{k-1}{j}\left(\sum_{i=1}^{m-1}\frac{\phi^{(i+1)}(hjx+b)}{i!}(hx)^i\right)\\&=\sum_{i=1}^{m-1}h^i\Big(\frac{1}{\phi^{(k)}(b)}\cdot\frac{k}{h^{k-1}}\sum_{j=0}^{k-1}\binom{k-1}{j}(-1)^{(k-j-1)}\phi^{(i+1)}(hjx+b)\Big)\\&=\sum_{i=1}^{m-1}h^i(kx^{k-1}+M_1\cdot h)\\&\le M_2\cdot h\end{aligned} \quad (3.13)$$

where M_1 and M_2 are positive constants.

The last part of (3.11) is

$$\begin{aligned}&\frac{1}{\phi^{(k)}(b)}\frac{k}{h^{k-1}}\sum_{j=0}^{k-1}(-1)^{(k-j-1)}\binom{k-1}{j}\left(\frac{\phi^{(m+1)}(hjx+b+\xi)}{m!}(hx)^m\right)\\&\le M_3\cdot h\end{aligned} \quad (3.14)$$

since $m-k+1>0$. From (3.12), (3.13) and (3.14), we have

$$||(x^k)'-N'_{k,h}||_{(\infty,[-1,1])}\le M^{**}\cdot h. \quad (3.15)$$

By (3.8) and (3.15), we completes the proof. □

The following is the main theorem of this paper.

Theorem 3.3. *Suppose that $\psi\in C^\infty(\mathbb{R})$ and there exists $b\in\mathbb{R}$ in some open interval $(b-\delta,b+\delta)$ in $\mathbb{R}$ such that*

$$\psi^{(k)}(b)\neq 0$$

for any k. For $f\in C^r[-1,1]$ with $||f^{(r)}||_{(\infty,[-1,1])}<\infty$ and a natural number n, there exists a neural network $N_{n,h}\in\Psi_{\phi,n}$ such that

$$||f-N_{n,h}||_{(\infty,[-1,1])}\le\frac{D_1}{n^r}\quad\text{and}\quad||f'-N'_{k,h}||_{(\infty,[-1,1])}\le\frac{D_2}{n^{r-1}} \quad (3.16)$$

where D_1 and D_2 are positive constants which are independent of n.

Proof. Let $n \geq 2$ be given. We choose the largest integer m such that $2m \leq n$. By Theorem 3.1, there exists an algebraic polynomial $P_{2m-1} \in \Pi_{2m-1}$ such that

$$||f - P_{2m-1}||_{(\infty,[-1,1])} \leq \frac{C^*}{m^r} \quad \text{and} \quad ||f' - P'_{2m-1}||_{(\infty,[-1,1])} \leq \frac{C^{**}}{m^{r-1}}$$

where C^* and C^{**} are constants which are independent of n. We set $P_{2m-1}(x) = \sum_{j=0}^{2m-1} a_j x^j$. For each j with $0 \leq j \leq 2m-1$, we choose h_j so small that

$$||x^j - N_{j,h_j}||_{(\infty,[-1,1])} \leq \frac{E_1}{\sum |a_j| m^r} \quad \text{and} \quad ||(x^j)' - N'_{j,h_j}||_{(\infty,[-1,1])} \leq \frac{E_2}{\sum |a_j| m^{r-1}}$$

where E_1 and E_2 are constants which are independent of m. Now, we set

$$N_n = \sum_{j=0}^{2m-1} N_{j,h_j} \in \Psi_{\psi,2m} \subset \Psi_{\psi,n}.$$

Since $\frac{1}{n} \leq \frac{1}{2} \cdot \frac{1}{m} \leq \frac{1}{n-1} \leq \frac{2}{n}$ for $n \geq 2$, $\frac{1}{m^r} \leq \frac{4^r}{n^r}$. Therefore, from (3.4) and (3.8), we have

$$\begin{aligned} ||f - N_n||_{(\infty,[-1,1])} &\leq ||f - P_{2m-1}||_{(\infty,[-1,1])} + ||P_{2m-1} - N_n||_{(\infty,[-1,1])} \\ &\leq \frac{C^*}{m^r} + \frac{E_1}{m^r} \leq \frac{D_1}{n^r} \end{aligned}$$

and also have

$$\begin{aligned} ||f' - N'_n||_{(\infty,[-1,1])} &\leq ||f' - P'_{2m-1}||_{(\infty,[-1,1])} + ||P'_{2m-1} - N'_n||_{(\infty,[-1,1])} \\ &\leq \frac{C^{**}}{m^{r-1}} + \frac{E_2}{m^{r-1}} \leq \frac{D_2}{n^{r-1}} \end{aligned}$$

from (3.5) and (3.15). This completes the proof. □

References

1. F. Cao, Z. Xu and Y. Li, *Pointwise Approximation for Neural Networks*, Lecture Note in Computer Science 3496(2005), 39-44.
2. E. W. Cheney, *Approximation Theory*, AMS Chelsea (2000).
3. B. I. Hong and N. Hahm, *Approximation Order to a Function in $\bar{C}(\mathbb{R})$ by Superposition of a Sigmoidal Function*, Applied Mathematics Letters 15(2002), 591 - 597.

4. N. Hahm, M. Yang and B. I. Hong, *Degree of Approximation to a Smooth Function by Generalized Translation Networks*, Honam Mathematical Journal 27(2005), 225-232.
5. M. Leshno, V. Y. Lin, A. Pinkus and S. Schoken, *Multilayer Feedforward Networks with a Nonpolynomial Activation Function Can Approximate Any Function*, Neural Networks 7(1995), 861-867.
6. X. Li, *Simultaneous Approximation of Multivariate Functions and Their Derivatives by Neural Networks with One Hidden Layers*, Neurocomputing 12(1998), 327-343.
7. H. N. Mhaskar, *Neural Networks for Optimal Approximation of Smooth and Analytic Function*, Neural Computation 8(1996), 164-177.
8. I. P. Natanson, *Constructive Function Theory*, Frederick Ungar Publ. (1964).
9. E. M. Stein, *Singular Integrals and Differentiability properties of Functions*, Princeton Univ. Press (1970).
10. A. F. Timan, *Theory of Approximation of Functions of a Real Variable*, Dover (1994).
11. Jian Jun Wang, ZongBen Xu and Weijun Xu, *Approximation Bounds by Neural Networks in L^p_ω*, Lecture Note in Computer Science 3173(2004), 1-6.
12. Zong-Ben Xu and Fei-Long Cao, *Simultaneous L^p-approximation Order for Neural Networks*, Neural Networks 18(2005), 914-923.

Nahmwoo Hahm received his Ph.D. from University of Texas at Austin under the direction of Professor E. W. Cheney. Since 2000, he has been a professor of University of Incheon. His research interests are numerical analysis and neural networks.

Department of Mathematics, University of Incheon, Incheon 402-749, Korea
e-mail: nhahm@incheon.ac.kr

Bum Il Hong received his Ph.D. in mathematics at Purdue in 1990. Since 1996, he has been in the department of mathematics at Kyung Hee University as a professor. His research interests are numerical analysis, numerical p.d.e. and neural networks.

Department of Mathematics, Kyung Hee University, Yongin 449-701, Korea
e-mail: bihong@khu.ac.kr