International Journal of Applied Mathematics and Statistics,
Int. J. Appl. Math. Stat.; Vol. 56; Issue No. 4; Year 2017,
ISSN 0973-1377 (Print), ISSN 0973-7545 (Online)

Simultaneous Approximation of Order m by Artificial Neural Network

Hawraa Abbas Almurieb

College of Education for Pure Sciences,
Department of Mathematics, University of Babylon, Iraq
almurieb2005@yahoo.com
pure.hawraa.abbas@uobabylon.edu.iq

ABSTRACT

Estimating upper and lower bounds is a key issue in neural network approximation. Many papers conclude one or both bounds of the first and second orders in terms of modulus of smoothness in recent years. In this work, we approximate a function in the space of Lebegue-integrable multivariate functions of period 2π *with order* p*, where* $f \in L_{2\pi}^{p}([-\pi,\pi]^{m}), 1 \le p \le \infty$ *is obtained. Then we obtain two-sided estimates of* m*th order modulus of smoothness of* f*, i.e.* $\left\|D^{\beta}TN_{\gamma}(f_i) - D^{\beta}(f_i)\right\|_{p} \sim \omega_m\left(D^{\beta}f_i,\delta\right)_{p}$*, where* TN_{γ} *is the FFNs with three trigonometric hidden layer units that is defined by Suzuki [*Suzuki1998*].*

Keywords: Neural Network, Simultaneous Approximation, Modulus of Smoothness.

Mathematics Subject Classification (MSC): 41A65,46S60,

Computing Classification System (CCS) : C.I.3 ,F.4.1

1. INTRODUCTION AND MAIN RESULTS

Many scientists and researchers have used multilayered neural networks to approximate multivariate functions for several years [see Lin and Cao 2015, Li and Xu 2007, Suzuki 1998 & Wang and Xu 2010]. They have established both upper and lower bounds of simultaneous approximations for 1st and 2nd orders, spaces of function to approximate and approximators as well. That work solves many applicant issues in science and engineer. Our goal was to achieve that both bounds of modulus of smoothness of order m for a pth Lebegue integrable multivariate function that is approximated by a multi-layered feedforwrd neural network.

Given a natural number m, $t = (t_i)_{i=1}^{m} \in N^m$, a function f belongs to the space $L_{2\pi}^{p}([-\pi,\pi]^m)$ under the norm defined by

$$\|f\|_p = \begin{cases} \left((2\pi)^{-m}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}|f(x)|^p dx\right)^{\frac{1}{p}} & 1 \le p \le \infty \\ \sup\{|f(x)| : |x_i| \le \pi\} & p = \infty. \end{cases} \quad (1)$$

For $f, g \in L_{2\pi}^{p}([-\pi,\pi]^m)$, define the following from [Liflyand 2006]

$$\langle f, g\rangle = (2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(t)g(t)dt\,, \tag{2}$$

and

$$f * g(x) = (2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(t)g(x-t)dt \tag{3}$$

In this paper, we use modulus of smoothness to measure the estimates of approximation, so we need to define the kth symmetric difference by

$$\Delta_t^k f(x) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} f(x + (\tfrac{k}{2} - j)t), \tag{4}$$

and the kth modulus of smoothness by

$$\omega_k(f, \delta)_p = \sup_{\|t\| \le \delta} \left\| \Delta_t^k(t) \right\|_p \tag{5}$$

Now, let us state some important properties of the classical modulus of smoothness that will be minor in our proofs, such us [Dineva, A, Várkonyi-Kóczy, Tar and Piur 2015]

(1) $\omega_k(f, \delta)_p$ is monotone increasing about δ;
(2) $\omega_k(f, b\delta)_p \le (1+b)^k \omega_k(f, \delta)_p\,, b > 0;$
(3) $\omega_k(f, \delta)_p \le 2^j \omega_{k-j}(f, \delta)_p\,, 0 \le j \le k;$
(4) $\omega_{k+s}(f, \delta)_p \le \delta^k \omega_s\left(f^{(k)}, \delta\right)_p;$

In order to approximate $f = (f_i)_{i=1}^n$ by $g = (g_i)_{i=1}^n$ with p-norm, each g_i should approximates each f_i with p-norm.

Finally, we need to define the three hidden trigonometric layer feedforward neural network defined by [Suzuki 1998]

$$TN_\gamma[f] = \left(TN_\gamma[f_i]\right)_{i=1}^n = \left(TN_\gamma[f_1], \ldots, TN_\gamma[f_n]\right)^T, \tag{6}$$

Where

$$TN_\gamma[f_i](x) = \theta_\gamma[f_i] + \sum_{\substack{\text{combinations}\\ \text{of } p \neq q \in N_0^m}}^{0 \le p_\mu, q_\nu \le \gamma} \left\{\alpha_{\gamma,p,q}[f_i]\cos(p-q)x + \beta_{\gamma,p,q}[f_i]\sin(p-q)x\right\}$$

We will prove the following equivalence of simultaneous approximation by Suzuki's three hidden layer neural network with modulus of smoothness of order m. It is summarized by:

$$\left\|D^{\beta}TN_{\gamma}(f_i) - D^{\beta}(f_i)\right\|_p \sim \omega_m\left(D^{\beta}f_i, \delta\right)_p$$

Our main result is separated in terms of two theorems as follow:

Theorem 1. For $D^{\beta}f = D^{\beta}(f_i)_{i=1}^{n} \in L_{2\pi}^{p}([-\pi,\pi]^m)$, we have

$$\left\|D^{\beta}TN_{\gamma}(f_i) - D^{\beta}(f_i)\right\|_p \le \left[\left(\frac{2}{\pi}\right)^{2m} + \frac{m\log(\pi m)}{2^m}\right](1+\delta^{-1})^m \omega_m\left(D^{\beta}f_i, \delta\right)_p$$

Theorem 2. For $D^{\beta}f = D^{\beta}(f_i)_{i=1}^{n} \in L_{2\pi}^{p}([-\pi,\pi]^m)$, we have

$$\left\|D^{\beta}TN_{\gamma}(f_i) - D^{\beta}(f_i)\right\|_p \ge \omega_m\left(D^{\beta}f_i, \delta\right)_p$$

2. AUXILIARY LEMMAS

In order to prove our theorems, we need to define Dirichlet kernel of m-dimension al space N_0^m from [Liflyand 2006] as follow:

$$K_n(t) = \sum_{0<\sum c_i<m} \prod_{i=1}^{m} \frac{\sin\left(2^{c_i} - \frac{1}{2}\right)t_i - \sin\left(2^{c_{i+1}} - \frac{1}{2}\right)t_i}{2\sin\left(\frac{t_i}{2}\right)} \quad (7)$$

One dimension, Dirichlet Kernel has very useful properties, we have to verify them for m dimension to be used in our proofs. They will be summarized in the next lemma.

Lemma 1 (Properties of Dirichlet Kernel of m Dimension)

1. K_n is even.
2. $|K_n(t)| \le \left(n + \frac{1}{2}\right)^n$
3. $\frac{\left|\sin\left(n+\frac{1}{2}\right)t\right|}{t} \le |K_n(t)| \le \frac{\pi}{2t}$ for $0 < t < \pi$,
4. $2\left(\frac{4}{\pi}\right)^m \log m \le \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \left|K_{\gamma}(t)\right| dt \le 2^m \left[\left(\frac{2^m}{\pi}\right)^m + \left(\pi\frac{\log(\pi m)}{2}\right)^m\right]$

Proof:

The fact that

$$K_n(t) = \sum_{0<\sum t_j \le n} \prod_{i=1}^{m} \left(\frac{1}{2} + \sum_{j=1}^{n} \cos i t_j\right)$$

solves easily properties 1 and 2. While 3 comes directly from the properties of 1-dimensional Dirichlet kernel $\frac{\left|\sin\left(n+\frac{1}{2}\right)t_i\right|}{t_i} \le |K_n(t_i)| \le \frac{\pi}{2t_i}$ for $0 < t_i < \pi$. For more information see [Carothers,2006].

The upper estimate in (4) is proved as follow:

$$\int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |K_\gamma(t)| dt = \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \left| \sum_{0<\sum c_i<m} \prod_{i=1}^{m} \frac{\sin\left(2^{c_i} - \frac{1}{2}\right)t_i - \sin\left(2^{c_{i+1}} - \frac{1}{2}\right)t_i}{2\sin\left(\frac{t_i}{2}\right)} \right| dt_i$$

$$\le 2^m \sum_{0<\sum c_i<m} \left[\int_0^{\pi} \dots \int_0^{\pi} \left| \prod_{i=1}^{m} \frac{\sin\left(2^{c_i} - \frac{1}{2}\right)t_i}{2\sin\left(\frac{t_i}{2}\right)} \right| dt_i + \int_0^{\pi} \dots \int_0^{\pi} \left| \prod_{i=1}^{m} \frac{\sin\left(2^{c_{i+1}} - \frac{1}{2}\right)t_i}{2\sin\left(\frac{t_i}{2}\right)} \right| dt_i \right]$$

$$\le 2^m \sum_{0<\sum c_i<m} \left[\int_0^{\rho} \dots \int_0^{\rho} \left| \prod_{i=1}^{m} \left(2^{c_i} - \frac{1}{2}\right) \right| dt_i + \int_{\rho}^{\pi} \dots \int_{\rho}^{\pi} \left| \prod_{i=1}^{m} \frac{\pi}{2t_i} \right| dt_i + \int_0^{\rho} \dots \int_0^{\rho} \left| \prod_{i=1}^{m} \frac{\sin\left(2^{c_{i+1}} - \frac{1}{2}\right)t_i}{2\sin\left(\frac{t_i}{2}\right)} \right| dt_i \right.$$

$$\left. + \int_0^{\rho} \dots \int_0^{\rho} \left| \prod_{i=1}^{m} \frac{\pi}{2t_i} \right| dt_i \right], \qquad where\ \rho = \frac{1}{m}$$

$$\le 2^m \sum_{0<\sum c_i<m} \left[\prod_{i=1}^{m} \frac{\left(2^{c_i} - \frac{1}{2}\right)}{m} + \prod_{i=1}^{m} \frac{\left(2^{c_{i+1}} - \frac{1}{2}\right)}{m} + 2\prod_{i=1}^{m} \frac{\pi}{2}(\log\pi + \log m) \right]$$

$$\le 2^m \left[2(2^{m^2}) + 2\left(\frac{\pi}{2}(\log(\pi m))\right)^m \right] \le 2^m \left[\left(\frac{2^m}{\pi}\right)^m + \left(\pi \frac{\log(\pi m)}{2}\right)^m \right].$$

For the lower bound, we have

$$\int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} |K_\gamma(t)| dt = \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \left| \sum_{0<\sum c_i<m} \prod_{i=1}^{m} \frac{\sin\left(2^{c_i} - \frac{1}{2}\right)t_i - \sin\left(2^{c_{i+1}} - \frac{1}{2}\right)t_i}{2\sin\left(\frac{t_i}{2}\right)} \right| dt_i$$

$$\ge 2^m \sum_{0<\sum c_i<m} \left[\int_0^{\pi} \dots \int_0^{\pi} \left| \prod_{i=1}^{m} \frac{\sin\left(2^{c_i} - \frac{1}{2}\right)t_i - \sin\left(2^{c_{i+1}} - \frac{1}{2}\right)t_i}{t_i} \right| dt_i \right]$$

$$\ge 2^m \sum_{0<\sum c_i<m} \left[\int_0^{\left(2^{c_i}-\frac{1}{2}\right)\pi} \dots \int_0^{\left(2^{c_i}-\frac{1}{2}\right)\pi} \left| \prod_{i=1}^{m} \frac{\sin x_i}{x_i} \right| dx_i + \int_0^{\left(2^{c_{i+1}}-\frac{1}{2}\right)\pi} \dots \int_0^{\left(2^{c_{i+1}}-\frac{1}{2}\right)\pi} \left| \prod_{i=1}^{m} \frac{\sin x_{i+1}}{x_i} \right| dx_i \right]$$

$$\ge 2^m \sum_{0<\sum r_i<m} \left[\int_0^{\left(2^{c_i}-\frac{1}{2}\right)\pi} \dots \int_0^{\left(2^{c_i}-\frac{1}{2}\right)\pi} \left| \prod_{i=1}^{m} \frac{\sin x_i}{x_i} \right| dx_i + \int_0^{\left(2^{c_{i+1}}-\frac{1}{2}\right)\pi} \dots \int_0^{\left(2^{c_{i+1}}-\frac{1}{2}\right)\pi} \left| \prod_{i=1}^{m} \frac{\sin x_{i+1}}{x_i} \right| dx_i \right]$$

$$\ge 2^m \sum_{0<\sum c_i<m} \left[\int_0^{m\pi} \dots \int_0^{m\pi} \left| \prod_{i=1}^{m} \frac{\sin x_i}{x_i} \right| dx_i + \int_0^{m\pi} \dots \int_0^{m\pi} \left| \prod_{i=1}^{m} \frac{\sin x_{i+1}}{x_i} \right| dx_i \right]$$

$$\geq 2^{2m} \sum_{0<\sum c_i<m} \left[\prod_{i=1}^{m} \sum_{k=1}^{m} \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} \dots \int_{(k-1)\pi}^{k\pi} |\sin x_i| dx_i + \prod_{i=1}^{m} \sum_{k=1}^{m} \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} \dots \int_{(k-1)\pi}^{k\pi} |\sin x_{i+1}| dx_i \right]$$

$$\geq 2 \left(\frac{4}{\pi}\right)^m \log m. \blacksquare$$

5. Proof of Theorem 1.

Let $r = (r_i)_{i=1}^m \in N_0^m$ and $\gamma \in N$.

Applying (1),(3), , we have

$$TN_\gamma f = \left(TN_\gamma f_i\right)_{i=1}^n = \left(K_\gamma * f_i\right)_{i=1}^n$$

where

$$\left(K_\gamma * f_i\right)(x) = (2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \sum_{i=1}^{m} f_i(x+t) K_\gamma(t) dt$$

Then, we have by (5),(7) and Lemma 1

$$\left\|D^\beta TN_\gamma(f_i) - D^\beta(f_i)\right\|_p = \left\|(2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} K_\gamma(t) \sum_{i=1}^{m} \left(D^\beta f_i(x+t) - D^\beta f_i(x)\right) dt\right\|_p$$

$$\leq (2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} K_\gamma(t) \left\|\Delta_t^m D^\beta f_i(x)\right\| dt$$

$$\leq (2\pi)^{-m} (1+\delta^{-1})^m \omega_m\left(D^\beta f_i, \delta\right)_p \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} K_\gamma(t)\, dt$$

Finally, by lemma 1, we have

$$\left\|D^\beta TN_\gamma(f_i) - D^\beta(f_i)\right\|_p \leq \left[\left(\frac{2}{\pi}\right)^{2m} + \frac{m \log(\pi m)}{2^m}\right] (1+\delta^{-1})^m \omega_m\left(D^\beta f_i, \delta\right)_p. \blacksquare$$

6. Proof of Theorem 2.

Using the properties of the classical modulus of smoothness mentioned above and applying (1) and (3) again, we have

$$\omega_m\left(D^\beta f_i, \delta\right)_p \leq \delta^m \left\|D^\beta(f_i)\right\|_p \leq \delta^m \left[\left\|D^\beta TN_\gamma(f_i) - D^\beta(f_i)\right\|_p + \left\|D^\beta TN_\gamma(f_i)\right\|_p\right]$$

$$\leq \delta^m \left[\left\| D^\beta TN_\gamma(f_i) - D^\beta(f_i) \right\|_p + \left\| (2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} K_\gamma(t) \sum_{i=1}^{m} D^\beta f_i(x+t) dt \right\|_p \right]$$

$$\leq \delta^m \left[\left\| D^\beta TN_\gamma(f_i) - D^\beta(f_i) \right\|_p + (2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} K_\gamma(t) \left\| \Delta_t^m D^\beta f_i(x) \right\|_p dt \right]$$

$$\leq \delta^m \left[\left\| D^\beta TN_\gamma(f_i) - D^\beta(f_i) \right\|_p + (2\pi)^{-m} (1+\delta^{-1})^m \omega_m\left(D^\beta f_i, \delta\right) \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} K_\gamma(t)\, dt \right]$$

Thus, Lemma 1. finishes the proof

$$\left\| D^\beta TN_\gamma(f_i) - D^\beta(f_i) \right\|_p \geq \left[\delta^{-m} - (1+\delta^{-1})^m (2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} K_\gamma(t)\, dt \right] \omega_m\left(D^\beta f_i, \delta\right)$$

$$\geq \left[\delta^{-m} - 2(1+\delta^{-1})^m \left(\frac{2}{\pi^m}\right)^m \log m \right] \omega_m\left(D^\beta f_i, \delta\right). \blacksquare$$

7. CONCLUSIONS AND FUTURE WORK

In this paper, the essential problem of simultaneous approximation of functions and its derivatives using neural networks is studied in terms of modulus of smoothness of order m. We found a reduced error of approximation by raising the order of modulus of smoothness to m, the order of the multivariate function. The function itself belongs to a Lebegue-integrable periodic multivariate functions, which is useful in various fields of approximation. Also, using FNNs to approximate multivariate functions is closer to touch the nonlinearity of some complicated functions. That clears the way to study simultaneous neural networks approximation in terms of Dunkl operator in the future.

8. REFERENCES

Carothers, N L., 2006. “A Short Course on Approximation Theory”, Bowling Green State University.

Dineva, A, Várkonyi-Kóczy, A R, Tar J and Piur, V.,2015. “Intelligent Neural Network Design for Nonlinear Control using Simultaneous Perturbation Stochastic Approximation (SPSA) Optimization”.

Ditizian, Z and Totik, V., 1987. “Modulus of smoothness”, Springer-Verlag, New York.

Liflyand, E R., 2006. “Lebesgue Constants of Multiple Fourier series”, Online Journal of Analytic Combinatorics, Issue 1 , # 5.

Lin, S Cao, F., 2015. “Simultaneous Approximation by Spherical Neural Networks”, Neurocomputing, Volume 175, Part A,348–354.

Li, F and Xu, Z., 2007. “The Essential Order of Simultaneous Approximation for Neural Networks”, Applied Mathematics and Computations 194 , 120-127.

Jain, S.K., Mahajan, B.K., Soni, B., 2009. “Approximation of Signal in the Holder Metric”, International Journal of Applied Mathematics and Statistics 14, Number 109, 93-96.

Padhy, B. P., Banitamani Mallik, Misra U.K., Mehendra Misra, 2013. “ On Degree of Approximation of Fourier Series by Product Means”, International Journal of Mathematics and Computation 19, Issue Number 2, 34-41.

Suzuki, S., 1998. “Constructive Function Approximation by Three Layer Artificial Neural Networks”, Neural Networks 11, 1049-1058.

Wang J and Xu Z., 2010. “New study on neural networks: The essential order of approximation”, Neural Networks 23, 618_624.