Lp Direct Theorem for Exponential Neural Networks

Eman S. Bhaya
University of Babylon College of Education for pure sciences Mathematics Department.
emanbhaya@itnet.uobabylon.edu.iq
Sarah A. Alameedee
University of Babylon College of Education for pure sciences Mathematics Department.
Sarahabd54@yahoo.com

Abstract

In this paper we introduce a Jackson type theorem for the approximation of function f In $L_p(R^d)$ by exponential neural networks.

Keywords : Neural network approximation , Modulus, L_p space, beast approximation.

الخلاصة

قدمنا في هذا البحث نوعا من مبرهنة جاكسون للتقريب باستخدام الشبكات العصبيه الصناعيه للدوال في الفضاءات $L_p(R^d)$.

الكلمات المفتاحية: التقريب باستخدام الشبكه العصبيه, مقياس النعومه, فضاءات L_p , التقريب الافضل.

1-Notations and definitions

Let $\mathbb{R}$be the set of reals, $\mathbb{R}^d$ be the d-dimensional Euclidean space $(d \geq 1)$.

Let

$$x = (x_1, x_2, \dots, x_3) \in \mathbb{R}^d,$$

$$e^x = (e^{x_1}, e^{x_2} \dots e^{x_d}), i = 1,2, \dots, d$$

And let, $P_n(d)$ be the space of all algebraic polynomials of d variables, $P_n^E(d)$the set of all real exponential Polynomials of d variables taking the form

$$\sum_{\lambda \in l}(N \cup \{0\})^d . a_\lambda e^{-\lambda x} \text{for positive } l$$

And let , $R_n^\sigma(d)$the set of all polynomials of the form

$$\sum_{\lambda \in l(N \cup \{0\})^d} a_\lambda f(-\lambda x + b_\lambda)(l > 0)$$

And $f: \mathbb{R} \to \mathbb{R}$.

let $|k|th$ order partial derivatives of $f$as

$$D^{|k|} f(x) := \frac{\partial^{|k|} f}{\partial x^{|k|}}(x) = \frac{\partial^{|k|} f}{\partial x_1^{k_1} \partial x_2^{k_1} \dots \partial x_s^{k_1}}(x) \quad \text{(Chui\& Li, 1993)}$$

Definition 1.1 (He& xu,1995)The mathematical expression of the feed forward neural networks (FNNs) with d-input layers and one hidden and output layer is of the form:

$$N(x) = \sum_{i=1}^{m} c_i \rho(\sum_{j=1}^{d} w_{ij} x_j + \theta_i) \,, x \epsilon \mathcal{R}^d \,, d \geq 1$$

where $1 \leq i \leq m, \theta_i \in \mathbb{R}, c_i \in \mathbb{R}$ are the connection strength of component$i$ with the output components,$w_i = (w_{i1}, w_{i2}, \ldots, w_{id})^T$ are weights of the components i in the hidden layer for the input components and ρ is the activation function used in the network. In this paper we mean by the components the neuron in the (FNN).

Definition 1.2 The p_quasi norm on $\mathbb{R}^d$ is denoted by $\|\cdot\|_p$ and

defined as

$$\|g\|_p = (\int_{\mathbb{R}^d} |g|^p)^{\frac{1}{p}} \infty > p > 0$$

And let $L_p(\mathbb{R}^d)$ be the space of all function g on $\mathbb{R}^d$ satisfying $\|g\|_p < \infty$.

Definition1.3 Let $f \in L_p(\mathbb{R}^d)$ and E is a set of real or complex

Functions , the distance from f to E defined by :

$$d_p(f, E) = \underset{g \in E}{Sup} \|f - g\|_p$$

Definition1.4 Let (Q, d) be ametric space then if $g \in L_p(Q)$, the 1st order modulus of smoothness of a function$g$has the form:

$$\omega(g, t)_p = \underset{\|x_1 - x_2\| \leq t}{\text{Sup}} \|g(x_1) - g(x_2)\|_p$$

Definition1.5 (Wang& Xu,2010) let Q be metric space with metric d then if $f \in L_p(Q)$given a direction $\in \mathbb{R}^d$, the rth order Symmetric difference of f defined by

$$\Delta_h^r f(x) = \sum_{i=0}^{r} (-1)^{r-i} \binom{r}{i} f(x + \left(\frac{r}{2} - i\right) he)$$

and ,the rth modulus of smoothness of a function$f$have the form

$$\omega_r(f, t)_p = \underset{x \pm \frac{he}{2} \in Q \cdot \|h\| \leq t}{\text{Sup}} \|\Delta_h^r f(x)\|_p$$

Definition1.6 (Ritter, 1999)A real function f is called nearly exponential if it satisfies , for any$\varepsilon > 0$,we can find real $\beta, \gamma, \rho, \tau$ such that

$$|\gamma \sigma(\beta t + \tau) + \rho - e^t| < \varepsilon \text{for all} t \leq 0$$

Example 1.7 (Wang& Xu, 2010)

Let $(\beta = 1, \sigma = 0, \gamma = 1)$ the sigmoid function $f(t) = \frac{1}{1+e^{-t}}$

Can also putting $\beta = 1, \rho = 0$ and $\gamma = \frac{1}{\sigma(\tau)}$

Then

$$\left|\frac{f(t+\tau)}{f(\tau)}-e^{t}\right|=\frac{e^{t+\tau}}{e^{t+\tau}+1}|1-e^{t}|\leq e^{\tau}|e^{t}-e^{2t}|$$

This converges to 0 for $t \leq 0$ and $\tau \to -\infty$

In Ritter , Ritter obtained the following result for the exponential neural network .

2-Auxilary results

Lemma2.1 (Xu& wang, 2006)Let f be a continuous function on $[0,1]^d$ and, $n \in N$, then we can find a nearly exponential type of forword neural network, $R_n^\sigma(d)$ have the form (1,1) , its number of hidden layer components is $M_n \geq min_{C<\varepsilon}(n+1)^d$ (where $C=\left(\frac{1}{2}+\frac{\pi^2}{4}\sqrt{d}\right)\omega\left(f,\frac{1}{n+2}\right)$.$n$ is any integer, and satisfies

$$d_\infty(f,R_n^\sigma(d))\leq\left(\frac{1}{2}+\frac{\pi^2}{4}\sqrt{d}\right)\omega(f,\frac{1}{n+2})$$

Lemma2.2 (Wang& xu, 2010)Let V is compact subset of R^d and $f \in C(V)$. Then there is a nearly exponential forward neural network, hidden layer components

$M_n \geq min_{B<\epsilon}\ (n+1)^d$

(here $B=\frac{1}{2}\left(\frac{\sqrt{d}\pi^2}{2}+1\right)^2\omega_2\left(f,\frac{1}{n+2}\right)$, n is any integer satisfying

$$d_\infty\left(f,R_n^\sigma(d)\right)\leq\frac{1}{2}\left(\frac{\sqrt{d}\pi^2}{2}+1\right)^2\omega_2\left(f,\frac{1}{n+2}\right);$$

The following lemma from[2] enables us to prove our theorems

Lemma2.3 (M.K.Kareem) If $f \in L_p[a,b]^d, 0<p<\infty$,then

$$E_{m-1}(f)_p \leq c(p,m,d)\omega_m(f,h,[a,b]^a)_p$$

where $h=(h_1,h_2,\dots,h_d)$.

3-The main results

We can strength the Lemma2.3 by proving it in terms of the rth order modulus of smoothness .

Theorem 3.1 Let $f \in L_p([0,1]^d)$ and $n \in N$ then there is a nearly exponential type of forward neural networks , and let $R_n^\sigma(d)$ have the form(1,1) , its number of hidden layer components is :

$$M_n \geq min_{C<\varepsilon}(n+1)^d$$

(where $C=c(p,d)\omega(f,\frac{1}{n})_p$, n is any integer satisfy

$$d_p(f,R_n^\sigma(d))\leq c(p,d)\omega_r(f,\frac{1}{n})_p$$

Proof :

Using lemma1.1 we get

$$\|p-f\|_p \le c(p,d)\omega_r(f,\frac{1}{n})_{p+\in} \qquad (1)$$

Given , $\alpha > 0$, define the function

$$F_\alpha(x) := \frac{1-e^{-\alpha x}}{1-e^{-\alpha}} x \in [0,1]$$

Its d-dimensional extension , $x \to F_\alpha(x)$ is topological isomorphism of $[0,1]^d$ and the family $(F_\alpha)_{\alpha>0}$ converges to the identity function F_0 on $[0,1]^d$ as $\alpha \to 0$, then we can choose α satisfying

$$\|p(F_\alpha)-p(F_0)\|_p \le \varepsilon \qquad (2)$$

We have $p(F_\alpha)$is an exponential polynomial in $P_n^E(d)$. Using (1) and (2) we obtain

$$\|p(F_\alpha)-f\|_p \le \|p(F_\alpha)-p(F_0)\|_p + \|p(F_0)-f\|_p$$

$$\le c(p,d)\omega_r(f,\frac{1}{n})_p + 2\varepsilon$$

Which is true for any $\varepsilon > 0$ therefore,

$$d_p\big(f,R_n^\sigma(d)\big) \le c(p,d)\omega_r(f,\tfrac{1}{n})_p \qquad \square$$

Theorem3.2 Let V be a compact subset of $\mathbb{R}^d$ and $f \in L_p(V)$. Then there is a nearly exponential forward neural networks with
hidden layer number of components

$$M_n \ge min_{B<\varepsilon}(n+1)^d$$

(where $B = c(p,d)\omega_r(f,\frac{1}{n})_p$, n is an integer, such that

$$d_p\big(f,R_n^\sigma(d)\big) \le c(p,d)\omega_r\left(f,\frac{1}{n}\right)_p \qquad (3)$$

Proof: Let V be a compact subset of $\mathbb{R}^d$, assume$T$is the Euclideanmap from$[0,1]^d$, to the compact set V , we have$f \in L_p(V)$, then $f(T)$in $L_p$is an extension of f on V. Hence Using Therom3.1 to get (3) directly.

References

Chui, C. K., &Li, x. (1993). Neural networks with one hidden layar. In K. Jetter, & F. I. Utreras (Eds), Multivariate approximation: World Scientific Press.

He., xingui, & Xu, Shaohua. (1995). Process neural Networks Theory and Applications. China: Springer.

M.A.Kareem (2011):On the multiapproximation in suitable spaces, M. SC .thesis, University of Babylon.

Ritter, G. (1999). Efficient estimation of neural weights by polynomial approximation. IEEE Transactions on Information Theory, 45, 1541-1550.

Wang, Jianjun, &Xu, Zongben. (2010). New study on neural networks: The essential order of approximation, Neural Networks . Science in china, Series F, 23, 618-624.

Xu, Z. B., & Wang , J. J. (2006). The essential order of approximation for neural networks. Science in China, Series F. Information Sciences, 49(4), 446-460.