Research Journal of Applied Sciences 11 (10): 933-941, 2016
ISSN: 1815-932X

Approximation by Regular Neural Networks in Terms of Dunkl Transform

Eman Bhaya and Omar Al-sammak
College of Education for Pure Sciences, University of Babylon, Al Hilla, Babil, Iraq

Abstract: Dunkl operator here we introduce a modified version of and use it to prove a theorem shows that functionals and rth order modulus of smoothness in K-theorem shows thatare equivalent. We use this equivalence to introduce p<1 spaces for L_p (K) essential degree of approximation using regular neural networks and how a multivariate function in spaces for can be approximated using a p<1 spaces for L_p (K) multivariate function in forward regular neural network. So, we can have the essential approximation using regular FFN. P<1 spaces for L_p (K) ability of a multivariate function in spaces for using regular FFN.

Key words: Neural network approximation, saturation problem, spaces, direct inequality

INTRODUCTION

There are many papers introduced about the direct and inverse theorem for the approximation by neural networks called the upper and lower bounds of the rate of approximation. See for example (Konovalov *et al.*, 2009, 2008; Maiorov, 2003, Maiorov and Pinkus, 1999; Xu and Cao, 2004), we call the degree of the asymptotically identical upper and lower bounds, the essential rate of approximation.

If we have a continuous function with multivariable and compact domain subset of there exist a feed Forward Neural Networks (FNNs) as anR^d approximation for it. input units, 1 hidden and one output units can^d The FNNs with 3 layers nd written as:

$$N_n(x) = \sum_{i=1}^{m} c_i \sigma\left(\left\langle \omega_{ij} x_j \right\rangle + \theta_i\right) \quad (1)$$

$$x \in R^d, d \geq 1$$

are $(\omega_{i1}, \omega_{i2}, ..., \omega_{id})^T \in R^d$ is the thre shold, $\theta_1 \in R$ $1 \leq i \leq m$ where in the hidden layer with input neurons,^{i} connection weights of neuron with the output neuron and ^{i} are the connection strength of neuron $c_i \in R$ is the sigmoidal activation function used in the network.

In many studies of approximation theory of functions the K-functionals play an important role. The study of the relation between the modulus of smoothness and K-functionals is an important problems in the approximation theory. For many types of mduli of soothness these problems are found in Dai (2003), Ditzian and Totik (1987), Lofstrom and Peetre (1969) for example.

Recently in mathematical papers a new class of generalized translations was described and put into use, namelya generalization of Dunkl translations. The generalized Dunkl translations are constructed on the base of certain differential-difference operators (the Dunkl operators) which are widely used in mathematical physics (Dunkl, 1989; Rosler, 2003).

Here we use rank d dunkl operators for functions of d-variables. The main aim is in to introduce an equivalence of K-functionals and the modulus of and p<1, L_p (k) smoothness using rank d dunkl operators for functions. we prove direct and inverse estimation and saturation problem for the using a p<1 spaces for L_p (k) approximation of multivariate function in spaces for using a forward regular neural network .

Notations and definitions: Let R be the real line, R^d be Euclidean space of d-dimensional for natural (d≥1) and let k be any subset of R^d .

Definition: Let K be a multiple cell in d-dimension Eucliden space R^d (d≥1), the $L_p(k)$ space for p <1defined by:

$$L_p = L_p(K) = \left\{ f : K \to R; \|f\|_p = \left(\int_k |f|^p \right)^{\frac{1}{p}} < \infty \right\} \quad (2)$$

We mean byα a real value satisfies α>-1/2. The modified Dunkl operator is a differential-difference operator D which satisfies the condition.

Corresponding Author: Eman Bhaya College of Education for Pure Sciences, University of Babylon, Al Hilla, Babil (Iraq)

$$Df(x) = \frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}\dots\frac{\partial f}{\partial x_d} + \left(\alpha + \frac{1}{2}\right)\frac{f(x) - f(-x)}{x_1 x_2 \dots x_d} \quad (3)$$

The action of the operator D is defined for all functions $f \in L_p^1(k)$ where $L_p^1(k) = \{f \in L_p, f \in L_p(K)\}$ Note that: any even function $f \in L_p^2(k)$ satisfies the equality:

$$D^2 f = Bf \quad (4)$$

where $L_p^2(k) = \{f \in L_p(k)\}$ and:

$$Bf(x) = \frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2}\dots\frac{\partial^2 f}{\partial x_d^2} + \left(\frac{2\alpha+1}{x_1}\frac{\partial f}{\partial x_1}\right)\left(\frac{2\alpha+1}{x_2}\frac{\partial f}{\partial x_2}\right)\dots\left(\frac{2\alpha+1}{x_d}\frac{\partial f}{\partial_{xd}}\right) \quad (5)$$

is the modify differential Bessel operator. Let $j\alpha$ (x) be the Bessel function of first kind:

$$j_\alpha(x) = \frac{2^\alpha \Gamma(\alpha+1) j_\alpha(x)}{x^\alpha} \quad (6)$$

that defined as the solutions to theBessel differential equation (Bateman and Erdelyi, 1974):

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + \left(x^2 - \alpha^2\right)y = 0 \quad (7)$$

Let us introduce the following exponential function:

$$e_\alpha(x) = j_\alpha(x) \quad (8)$$

The function $e_\alpha(x)$ is the generalized exponential function coincides with the usual exponential function e^x. We define a generalized exponential multivariable function as:

$$E_a(x) = e_a(x_1)e_a(x_2)\dots e_a(x_d) \quad (9)$$

The operator Df is defined for any function f∊ L_p (k). One can define the operator of the generalized modify Dunkl translation T^y f(x) in various ways. For a function f∊D one can define the operator of the generalized modify Dunkl translation u (x,y) = T^y f (x) as a solution of the Cauchy problem (Salem and Kallel, 2004)

$$D_x u(x,y) = (x,y), u(x,0) = f(x) \quad (10)$$

where D_x and D_y are the Dunkl operators applied with respect to variables x and y, correspondingly.

One can extend the operator T^y by from a subset $D \subset L_p$ onto the whole space L_p (k). The extended operator is also denoted by T^y.

Definition: Tthe r-th order difference of function f∊L_p (k)

$$\Delta_h^r f(x) = \left(1 - T^h\right)^r f(x), x = x_1, x_2, \dots, x_d \in K, R \in N \text{ and } h > 0 \quad (11)$$

ByI we mean the unit operator. If r a positive integer the rth module of smoothness can be defined by:

$$\omega_r(f,\delta)L_{p(k)} = \omega_r(f,\delta)_p = \sup_{0<h<\delta} \left\|\Delta_h^r f\right\|_p . \delta > 0, f \in L_p \quad (12)$$

Let w_p^r be the Sobolev space generated by D, i.e.:

$$W_p^r = \left\{f \in L_p : D^j f \in L_p, j = 1,2,\dots,r\right\} \quad (13)$$

Where:

$$D^j f(x) = \frac{\partial^j f(x)}{\partial x^j} = \frac{\partial^j f(x)}{\partial x_1^{j1} \partial x_2 j_2 \dots \partial x d^{jd}}, j_1 + j_2 \dots j_d = j \quad (14)$$

Definition: The K-functional on the spaces L_p and w_p^r defined by:

$$K\left(f, t; w_p^r\right) = \inf\left\{\|f - g\|_p + t\|D^r g\|_p \, g : W_p^r\right\} \quad (15)$$

where f∊L_p (K), t>0

Definition: we denoted by the Lipschitzian class Lip $(\alpha)_r$ defined by the space of all functions f in L_p (K) spaces satisfies ω_r $(f,t)_p$ = o (t^a), where 0<α≤r

Remarks:

- Here, we study the approximation using neural network with activation functions j_k of which each function σ^k R→ [0,1] has up to k+1 order continuous derivatives. K = 1,2,..., k+1
- The regular neural activation functions are the normal sigmoidal activation functions σ (x) = 1/1+e^{-ax} for positive

- Any neural network whose neural activation functions are regular (satisfies the conditions of part 1 of this remark) will be called a regular neural network.

Definition:
"Let r be an integer number and:

$$P_r(x) = a_r x^r, x \in [a,b] \subset (-\infty, \infty) \quad (16)$$

be a homogeneous univariate r degree polynomial:

MATERIALS AND METHODS

Some properties of the modified dunkl transform: The modified Dunkl transform is the integral transform:

$$F: f(x) \to f(\lambda) = \int_K fE_\alpha, \lambda \in K$$

where:

$$f(\lambda) = \langle f(x), E_\alpha(\lambda x)\rangle = \int_K f(x_1, x_2, \ldots x_d) E_\alpha(\lambda_1 x_1, \lambda_2 x_2, \ldots \lambda_d x_d) dx_1, dx_2, \ldots dx_d \quad (17)$$

The inverse of the formula of modified Dunkl transform is as follows:

$$F^{-1}: g(\lambda) \to f(x) = c\int_K gE\alpha$$

whrere:

$$f(x) = \langle g(\lambda), E_\alpha(-\lambda x)\rangle = \int_K g(\lambda_1, \lambda_2, \ldots \lambda_d) E_\alpha(\lambda_1 x_1, \lambda_2 x_2, \ldots \lambda_d x_d) d\lambda_1, d\lambda_2, \ldots d\lambda_d \quad (18)$$

where c is a constant and $k = k_1 \times k_2 \times \ldots k_d$ with $\lambda_1 \in K_1$, $\lambda_2 \in K_2, \ldots \lambda_d \in K_d$.

Lemma: let $f \in L_p$ (k) then the following equality is true for any $\in R$:

$$T^y f = E_\alpha(\lambda y) \quad (19)$$

here, f→f is modified Dunkl transform:

Lemma: For real number x the following inequalities are satisfied: $|e_\alpha(x)| \leq 1$ and the equality is attained only with x = 0:

$$|1 - e_\alpha(x)| | \leq 2|x| \quad (20)$$

For a given c>0 and $|x| \geq 1$ we have $|1 - e_\alpha(x) \geq c|$ As a corollary of above Lemma 3.3 we have:

Lemma: The following inequalities are satisfied:

$$|E_\alpha(x)| \leq 1 \quad (21)$$

$$|1 - E_\alpha(x)| \leq c \quad (22)$$

$|1 - E_\alpha(x)| \geq c$ where $|x| \geq 1$ and c is a positive constant.

Proof: part 1:

$$|E_\alpha(x)| = |e_\alpha(x_1) e_\alpha(x_2) \ldots e_\alpha(x_d)| = |e_\alpha(x_1)| |e_\alpha(x_2)| \ldots |e_\alpha(x_d)| \quad (23)$$

and by the above Lemma 3.3 part (1) we have $|e_\alpha(x)| \leq 1$ then we obtain $|e_\alpha(x)| \leq 1$:

$$|1 - E_\alpha(x)| = |1 - e_\alpha(x_1) e_\alpha(x_2) \ldots e_\alpha(x_d)| \quad (24)$$

and by part (1) of this lemma $|E_\alpha(x)| \leq 1$ and part (2) of Lemma 3.3 we obtain:

$$|1 - E_\alpha(x)| \leq c \quad (25)$$

part (3) : by the part (3) of lemma (3.3) with $|x| \geq 1$:

$$\text{We get } |1 - E_\alpha(x)| \geq c, \text{since} |x| = \sup \{|x|_1, x_2, \ldots |x|_d\} \quad (26)$$

Lemma: Let $f \in L_p(K)$ then $\|\Delta_h^r f\|_p \leq c(p) \|f\|_p$
Proof:

$$\|\Delta_h^r f\|_p = \left(\int_K \left| (1 - T^h)^r f(x) \right|^p \right)^{\frac{1}{p}} \quad (27)$$

$$= \left(\int_K \left| (1 - E_\alpha(\lambda h))^r f(x) \right|^p \right)^{\frac{1}{p}} \quad (28)$$

Lemma 3.4 part (2) we get:

$$\|\Delta_h^r\|_p \leq c(\alpha, r, p) \left(\int_K |f(x)|^p \right)^{\frac{1}{p}} \leq c(p) \|f\|_p \quad (29)$$

Lemma: If a function belong to the sobolev space W_p^r then

$$D^r f = \lambda^r f(x) \quad (30)$$

Lemma: Let, $\in w_p^r, t > 0$ then $\omega_r(f, \delta)_p \leq c(p) t^r \|D^r\| f$.

Proof: Assume that $0<h\leq t$.

$$\Delta_h^r(x)=\left(1-T^h\right)^r f \tag{31}$$

the difference with h as a step. Properties (3.2) yield:

$$\|\Delta_h^r F\|_p=\left(\int_k\left|\left(1-E_\alpha(\lambda h)\right)^r f(x)\right|^p\right)^{\frac{1}{p}} \tag{32}$$

and Lemma yield:

$$\|D^r f\|_p=\left(\int_K\left|\lambda^r f(x)\right|^p\right)^{\frac{1}{p}} \tag{33}$$

implies the equality:

$$\|\Delta_h^r f\|_p=h^r\left(\int_K\left|\frac{1-E_\alpha(\lambda h))^r}{(\lambda h)^r}\right|^p\right)^{\frac{1}{p}} \tag{34}$$

$$\|\Delta_h^r f\|_p=h^r\left(\int_K\left|\left(1-E_\alpha(s)\right)^r\right)s^{-r}\lambda^r f(x)\right|^p\right)^{\frac{1}{p}} \tag{35}$$

using Lemma 3.4 for any s we have:

$$\left|\left(1-E_\alpha(s)\right)^r s^{-r}\right|\leq c(p) \tag{36}$$

then Relations give:

$$\|\Delta_h^r f\|_p=\leq c(p)h^r\left(\int_K\left|\lambda^r f(x)\right|^p\right)^{\frac{1}{p}} \tag{37}$$

and Relations give:

$$\|\Delta_h^r f\|_p=\leq c(p)h^r\|D^r f\|_p \tag{38}$$

taking the supremum on h∈ (0,t]. we get. we obtain $\omega_r(f,\delta)_p\leq c(p)t^r\|D^r f\|$

Remark: For any function $f\in L_p(K)$ and v>0 define the map:

$$p_v(f)(x)=\int_{-v}^{v} f(x)E_\alpha(\lambda x)d\lambda=\chi_v(\lambda)f(x) \tag{39}$$

where $\chi_v(\lambda)$ is the characteristic function of [-v,v] .

Lemma: If $f\in L_p(K)$ then:

$$\|f-p_v(f)\|_p\leq c(p)\|\Delta_{j/v}^r f\|_p, v>0 \tag{40}$$

Proof:

$$\text{let}\left|1-E_\alpha(t)\right|\geq c(p)\text{ with }|t|\geq 1 \tag{41}$$

$$\|f-p_v(f)\|_p=(\int_K\left|1-\chi_v(\lambda)f(x)^p\right|^{\frac{1}{p}} \tag{42}$$

$$=\left(\int_K\left|\frac{1-\chi_v(\lambda)}{\left(\left(1-E_\alpha(\lambda 1/v)\right)^r\right)\left(1-E_\alpha(\lambda 1/v)\right)^r f(r)}\right|^p\right)^{\frac{1}{p}} \tag{43}$$

$$=\sup_{\lambda\in K}\left(\int_K\left|\frac{1-\chi_v(\lambda)}{\left(\left(1-E_\alpha(\lambda 1/v)\right)^r\right)\left(1-E_\alpha(\lambda 1/v)\right)^r f(x)}\right|^p\right)^{\frac{1}{p}} \tag{44}$$

$$\leq c(p)\|\Delta_{1/v}^r f\|_p \tag{45}$$

Corollary:

$$\|f-p_v(f)\|_p\leq c(p)\omega_r(f,1/v)_p \tag{46}$$

Lemma: The inequality

$$\|D^r\left(p_v(f)\right)\|_p\leq c(p)v^r\|\Delta_{1/v}^r f\|_p \tag{47}$$

Is true Proof : using correlation we obtain:

$$\|D^r\left(p_v(f)\right)\|_p=\left(\int_K\left|\lambda^r\chi_v(\lambda)f(x)\right|^p\right)^{\frac{1}{p}} \tag{48}$$

$$=\left(\int_K\left|\frac{\lambda^r-\chi_v(\lambda)}{\left(\left(1-E_\alpha(\lambda 1/v)\right)^r\right)\left(1-E_\alpha(\lambda 1/v)\right)^r f(x)}\right|^p\right)^{\frac{1}{p}} \tag{49}$$

Not that:

$$\frac{\sup_{\lambda\in K}\left(\lambda^r\chi_v(\lambda)\right)}{\left|1-E_\alpha(\lambda 1/v)\right|^r}=\frac{v^r\sup_{|\lambda|\leq v}(\lambda 1/v)^r}{\left|1-E_\alpha(\lambda 1/v)\right|^r} \tag{50}$$

Then, formula yields:

$$= \frac{v^r Supt^r \quad |t| \le 1}{|1 - E_\alpha(t)|^r} \tag{51}$$

$$Let^{c(p)} = \frac{v^r Supt^r \quad |t| \le 1}{|1 - e_\alpha(t)|^r} \tag{52}$$

Corollary:

$$\|D^r(p_v(f))\|_p \le C(p) V^r \omega r(f, 1/v)_p \tag{53}$$

Equivalence between k-functionals and r-th modulus of smoothness in terms of dunkl transform for neural networks: Our main result of this section is:

Theorem: For any function $f \in L_p(K), \delta > 0$ then

$$c(p)\omega_r(f,\delta)_p \le K_r(f,\delta^r)_p \le c(p)\omega_r(f,\delta)_p \tag{54}$$

Proof : proof of the inequality:

$$c(p)\omega_r(f,\delta)_p \le K_r(f,\delta^r)_p \quad \text{Let} h \in (0,\delta], g \in W_p^r \tag{55}$$

$$\|\Delta_h^r f\|_p = \left(\int_K |\Delta_h^r (f - g) + g|^p \right)^{\frac{1}{p}} \tag{56}$$

$$\le \left(\int_K |\Delta_h^r (f - g)|^p \right)^{\frac{1}{p}} + \left(\int_K |\Delta_h^r g|^p \right)^{\frac{1}{p}} \tag{57}$$

using Lemma and Lemma we obtain:

$$\|\Delta_h^r f\|_p \le c(p)\left(\int_K |f - g|^p \right)^{\frac{1}{p}} + c(p)h^r \left(\int_K |D^r g|^p \right)^{\frac{1}{p}} \tag{58}$$

$$\|f - g\|_p + \delta^r \|D^r g\|_p \tag{59}$$

take the supremum on 0<h≤δ and the infimum on any function $g \in w_p^r$

We obtain:

$$\omega_r(f,\delta)_p \le c(p)\omega_r(f,\delta)_p \tag{60}$$

proof of the inequality:

$$K_r(f,\delta^r)_p \le c(p)\omega_r(f,\delta)_p \tag{61}$$

We have, $p_v(f) \in W_p^r$ so using K-functional definition to get:

$$K_r(f,\delta^r)_p \le \|f - p_v(f)\|_p + t\|D^r(p_v(f))\|_p \tag{62}$$

Using Corollaries and. Let us proceed with inequality:

$$K_r(f,\delta^r)_p \le c(p)\omega_r(f,1/v)_p + c(p)(\delta v)^r \omega_r(f,1/v)_p \tag{63}$$

since, v be any positive value, let us, $v = \frac{1}{\delta}$.

$$\text{Weget} K_r(f,\delta^r)_p \le c(p)\omega_r(f,\delta)_p + c(p)\delta^r \omega_r(f,\delta)_p \tag{64}$$

$$K_r(f,\delta^r)_p \le c(p)\omega_r(f,\delta)_p \tag{65}$$

The essential order of approximation using regular neural networks: XU Zongben and CAO Feilong proved in [13] the following result "Let [a, b] be a compact interval, $\sigma \epsilon j_r$ be a regular neural activation function and $p_r(x)$a homogeneous univariate polynomial of the form . Then for any,ϵ>0, there is a neural network of the form (1.1) the number of whose hidden units is:

$$\text{not less than } (r + 1) \text{ such that} \quad |N_n(x) - P_r(x)| < \epsilon \tag{66}$$

As a direct consequence of above lemma we introduce the following theorem ".

Theorem: "For any regular neural activation function $\sigma \epsilon j_r$ and a homogeneous univariate polynomial p_r (x) and a given ϵ>0 we can find a neural network with form :

$$\text{with not less than } r+1 \text{hidden layers} \quad \text{satisfies} \|N_n(x) - p_r(x)\|_p \tag{67}$$

Proof: Since, $\sigma \epsilon j_r$ be a regular neural activation function and p_r (x) a homogeneous univariate polynomial then by above lemma we have for any ϵ>0 there is a neural network of the from the number of whose hidden units is not less than (r+1) such that :

$$|N_n(x) - p_r(x)| < \frac{\in}{(\mu(k))^{1/p}} \quad (68)$$

$$\Rightarrow \int_K |N_n(x) - p_r(x)|^p < \left(\frac{\in}{\mu(k)^{1/p}}\right)^p \quad (69)$$

$$\Rightarrow \int_K |N_n(x) - p_r(x)|^p)^{1/p} < \left(\int_K (\frac{\in}{\mu(k)^{1/p}})^p\right)^{1/p} \quad (70)$$

$$\Rightarrow \int_K |N_n(x) - p_r(x)|^p)^{1/p} < \left(\int_K (\frac{\in}{\mu(k)^{1/p}})^p\right)^{1/p} \quad (71)$$

$$\Rightarrow \|N \downarrow n(x) - p \downarrow r(x)\| \downarrow p < (\in (\mu(k)))^{\uparrow} (1/p))/((\mu(k))^{\uparrow}(1/p)) \quad (72)$$

$$\Rightarrow \|N_n(x) - p_r(x)\|_p < \in \quad (73)$$

In this section we construct an FNN to realize universal approximation to any integral multivariate functions in L_p (K),p<1 we will use the Bernstein-Durrmeger operation as base tools .

Definition: Let K be any subset of R^d the Bernstein-Durremger operater B_n in $L_1(T)$ defined by :

$$\begin{gathered}\sum p_{n.k}(x)\phi_{n.k}(f) \\ (B_{\downarrow}nf)(\alpha) = |k| \le n\end{gathered} \quad (74)$$

where, xϵK, f϶ L_1 (K) and:

$$\phi_{n.k}(f) = (n+d)_{!/n!} \int_k P_{n.k}(u) f(u) du \quad (75)$$

Lemma: If $f \in L_p(k)$ for p<1 then:

$$\|B_n f - f\|_p \le c(p)\omega_r(f, 1/n)_p \quad (76)$$

Proof :

$$\|B_n f - f\|_p \le c(p) K_r (f, 1/n)^r)_p \quad (77)$$

$$\le c(p)\omega_r\left(f, \frac{1}{n}\right)_p$$

Lemma: "If for If $f \in L_p(k)$ for p<1 then then:

$$\omega_r\left(f, \frac{1}{n}\right) \le c(p)\sum_{i=1}^{n} \|B_i f\|_p \quad (78)$$

$$\begin{gathered}\omega_r(f,\delta)_p = \omega_r(f - B_n f + B_n f, \delta)_p \le \\ c(p)\omega_r(f - B_n f, \delta) + \omega_r(B_n f, \delta) \le \\ c(p)\|f - B_n f\| + j\end{gathered} \quad (79)$$

$$\begin{gathered}B_n(f) - B_0(f) - B_n(f) - B_2(f) + \\ (B_{2^1} - 1(f)) + ... + (B_1(f) - B_0(f))\end{gathered} \quad (80)$$

$$J = \omega_r(B_n f_y \delta) = \omega_r\left(\sum_{i=1}^{1} B_2 i - B_2 i - 1, 2^{-1}\right)_p \quad (81)$$

$$= \omega_r\left(\sum_{i=1}^{1} B_2 i - f + f - B_{2i-1}, 2^{-1}\right)_p \quad (82)$$

$$\le c(p)\sum_{(i=1)^1} \left\|f - B_{(2^i)} f\right\|_p \quad (83)$$

$$\le c(p)\sum_{(i=1)^n} \|f - B_i f\|_p \quad (84)$$

$$\therefore \omega_r(f,\delta) \le c(p)\sum_{(i=1)^n} \|B_i f - f\|_p \quad (85)$$

To introduce Lemma 5.6 we need the flowing notations: (Xu and Cao, 2004). Le z_+^d t be the set of all non-negative multi-integers in R^d:

$$\begin{gathered}\text{For any } x = (x_1, x_2, ..., x_d) \in R^d \\ \text{and } k = (k_1, k_2, ..., k_d) \in Z_+^d, \text{Let}\end{gathered} \quad (86)$$

$$\begin{gathered}|x| = \sum_{i=1}^{d} x_i, |k| = \sum_{i=1}^{d} x_i, x^k = x_1^k ... x_d^{kd} \\ \text{and } k! = k_1! k_2! ... k_d!\end{gathered} \quad (87)$$

We say that x≤y, for any yϵR^s, iff xi≤yi for any 1≤i≤s. For any fixed point p. Let:

$$N_p = \frac{p + d - 1}{d - 1}$$

Be the number of multi-integers $i = (i_1, i_2, ..., i_t)$ in Z d/+ that satisfy:

$$j_1 + j_2 + ... + j_{d-1} = p \quad (88)$$

Let:

$$l_p = \frac{p+d-2}{d-2}$$

be the number of multi-integers $j=(j1, j2,...jd) in z_+^{d-1}$ in that satisfy:

$$j_1 + j_2 + ... + j_{d-1} = p \quad (89)$$

Denote by j_{Np-1}+l,1 ≤l≤l_p a generic multi-integers j in Z^{d-1}_+ satisfing $j_1 + j_2 + ... + j_{d-1} = p$. The f1,1 ≤$N_p$ a generic multi-integers j in Z^{s-1}_+,satisfing:

$$p_1 = (1, j_1), 1 \le 1 \text{and} p_1^{(p)} = \frac{1}{2(1+p)} p_1, 1 \le 1 \le N_p \quad (90)$$

For any $1 \le l \le N_p$ Le $i_l^{(p)} = (p - |jl|, jl)$ then each $i_l^{(p)} = p$ is multi-integer in z_+^d that satisfies $|i_l^{(p)}| = p$. Define:

$$B_n f(x) = \sum_{p=0}^{N} \sum_{l=1}^{N_p} \left\langle x, p_l^{(n)} \right\rangle^p \quad (91)$$

we then have $|p_{l.}^{(p)}| \le \frac{1}{2}$ for any 1 ≤l≤N_p" The flowing lemma provides an equivalent expression of Bernstein-durrmeyer operator B_n.

Lemma: "For any, ϵ L_p(T) the Berustin-Durrmeyer operator B_nf in can be expressed as:

$$B_n f(x) = \sum_{p=0}^{n} \sum_{1}^{N_p} d_l^{(p)}(x, p_l^{(n)})^p \quad (92)$$

where, $\langle x, p_l^{(n)} \rangle$ is inner product x and $p_l^{(n)}, d_l^{(p)}$ are uniquely determined by:

$$\begin{pmatrix} (j_1)^{j1} & (j_2)^{j1} & \cdots & (j_{NP})^{j1} \\ (j_1)^{j2} & (j_2)^{j2} & \cdots & (j_{NP})^{j2} \\ (j_1)^{jNp} & (j_2)^{jNP} & \cdots & (j_{NP})^{NP} \end{pmatrix} \begin{pmatrix} d_1^{(p)} \\ d_2^{(p)} \\ \vdots \\ d_{NP}^{(p)} \end{pmatrix} = \left(\frac{2(1+n)^p}{p!} \right) \begin{pmatrix} i_1^{(p)}! & c_1^{(p)}(f) \\ i_2^{(p)}! & c_2^{(p)}(f) \\ i_{NP}^{(p)}! & c_{NP}^{(p)}(f) \end{pmatrix} \quad (93)$$

$$c_l^{(p)}(f) = \frac{n!}{(n-p)!} \sum_{q \le i_l^{(p)}} \Phi_{n,q}(f) \frac{1}{q!\left(i_l^{(p)} - q\right)_l} (-1) \left| l_l^{(p)} - q \right| \quad (94)$$

$$N_{p+1}(x) = \sum_{i=1}^{Np} c_{i,p} \sigma\left(\omega_{i,p} \left\langle x, p_l^{(n)} \right\rangle + 0\right), K_p \ge p+1 \quad (95)$$

Remark: We observe that in expression (5.7) each term $\langle x, p_l^{(n)} \rangle$ can be viewed as homogeneous uniriate polynomial of $\langle x, p_l^{(n)} \rangle$ with order p and so by Theorem 5.1, it can be approximated arbitrarily by a network of the form.

Remark: As $B_n f(x)$ can be approximate f, the following neural networks:

$$N_n(x) = \sum_{p=0}^{n} \sum_{l=1}^{Np} c_{i,p} \sigma\left(\omega_{i,p} \left\langle x, p_l^{(n)} \right\rangle + \theta\right) \quad (96)$$

Then, can approximate f to any accuracy. The network will be the FNN models we propose use in this paper. The network are clearly of from (Eq. 1) and contain hidden units where:

$$m = N_0, k_1 N_1 + ... + k_n N_n \quad (97)$$

$$\ge N_0 + 2N_1 + ... + (n+1) N_n \quad (98)$$

$$= \sum_{k=0}^{n} (k+1) \begin{pmatrix} k+d-1 \\ d-1 \end{pmatrix} = m_0(n) \quad (99)$$

RESULTS AND DISCUSSION

Theorem 5.11 (The Direct Theorem for Approximation of multivariate function in L_p(k) spaces for p<1 by forward Regular Neural Networks): For any $f \in L_p(k), p < 1$ there is a regular, one hidden layer FNN, N_n(x) of the form (1.1) with σϵj_n and the hidden unit: number and satisfies:

$$m \ge \sum_{k=0}^{n} (k+1) \begin{pmatrix} k+d-1 \\ d-1 \end{pmatrix} = m_0(n) \quad (100)$$

$$\|N_n - f\|_p \le c(p) \omega_r \left(f, \frac{1}{n} \right)_p \quad (101)$$

Proof: we assume $f \in L_p(k)$ Then, by Lemma 5.6, the Bernstein durrmeyer operator $B_n$fcan be defined and expressed as:

$$B_n f(x) = \sum_{p=0}^{n} \sum_{l=1}^{N_p} d_l^{(p)} \left\langle x, p_l^{(n)} \right\rangle^p \quad (102)$$

and furthermore, it approximates f in the following sense: Lemma:

$$\|B_n f - f\|_p \le c(p) \omega_r \left(f, \frac{1}{n} \right)_p \quad (103)$$

and by remark (5.9) part 1, we have:

$$\|N_n - f\|_p \le c(p)\|B_n f - f\|_p\, \omega_r\left(f, \frac{1}{n}\right)_p \quad (104)$$

Theorem 5.12 (The inverse estimation for the approximation of multivariate function in $L_p(k)$ spaces for 1using a forward regular neural network): For any, $f \in L_p(k), p<1$ there is a regular, one hidden layer FNN, $N_n(x)$ of the form (1.1) with $\sigma\in j_n$ and the hidden unit:

$$m \ge \sum_{k=0}^{n}(k+1)\begin{pmatrix} k+d-1 \\ d-1 \end{pmatrix} \quad (105)$$

number such that:

$$c(p)\omega_r\left(f, \frac{1}{n}\right)_p \le \|N_n - f\|_p \quad (106)$$

Proof: we assume $f\in L_p(k)$ Then, by Lemma 5.6, the Bernstein-durrmeyer foperator B_n can be defined and expressed as:

$$B_n f(k) = \sum_{p=0}^{n}\sum_{l=1}^{N_p} d_l^{(p)} \left\langle x, p_l^{(n)} \right\rangle^p \quad (107)$$

and since $\left|p_l^{(n)}\right| \le \frac{1}{2}$ we have $\left\langle x, p_p^{(n)} \right\rangle \le 1$:

$$\left(-1 \le \left\langle x, p_l^{(n)} \right\rangle \le 1\right) \quad (108)$$

Each term $\left\langle x, p_l^{(n)} \right\rangle^p$ in (5.7) is univariate homogeneous polynomial of $\left\langle x, p_l^{(n)} \right\rangle$ with order define on. Now by Theorem 5.1 we have $\left\langle x, p_l^{(n)} \right\rangle^p$ can be approximated by neural network:

$$N_{K_p} = \sum_{i=1}^{N_p} c_p \sigma\left(\omega_{p,i}\left\langle x, p_l^{(n)} \right\rangle + \theta\right), \quad c_{p,i}, \omega_{p,i} \in R, K_p \ge p+1 \quad (1.09)$$

With accuracy :

$$\left\| N_{K_p} - \left\langle x, p_l^{(n)} \right\rangle^p \right\|_p \le\epsilon \quad (110)$$

Then for constructed FNN:

$$N_{K_p} = \sum_{p=0}^{n}\sum_{i=1}^{N_p} d_l^{(p)} \sum_{i=1}^{kp} c_p \sigma\left(\omega_{p,i}\left\langle x, p_l^{(n)} \right\rangle + \theta\right), \quad c_{p,i}, \omega_{p,i} \in R, K_p \ge p+1 \quad (111)$$

and we have:

$$\|N_n - f\|_p = \|N_n f - B_n f - f\|_p \quad (112)$$

$$\le \|N_n f - B_n f\|_p + \|B_n f - f\|_p \quad (113)$$

$$\le c(p)\omega_r\left(f, \frac{1}{n}\right)_p + \|N_n f - B_n f\|_p \quad (114)$$

The term $\|N_n f - B_n f\|_p$ in above can be arbitrarily small, because (5.13) and (5.14) imply:

$$\|N_n f - B_n f\|_p = \left\| \sum_{p=0}^{n}\sum_{l=p}^{N_p} d_l^{(p)} \begin{Bmatrix} \left\langle x, p_l^{(n)} \right\rangle \\ p - \sum_{i=1}^{kp} c_{i,p}\sigma\left(\omega_{i,p}\left\langle x, p_l^{(n)} \right\rangle + \theta\right) \end{Bmatrix} \right\|_p \quad (115)$$

$$= \sum_{p=0}^{n}\sum_{l=1}^{N_p} \left|d_l^{(p)}\right| \max\left|\left\langle x, p_l^{(n)} \right\rangle p - N_{K_p}\right| \le\epsilon \sum_{p=0}^{n}\sum_{l=1}^{N_p}\left|d_l^{(p)}\right| \quad (116)$$

$$\|B_n f - f\|_p \le c(p)\omega r\left(f, \frac{1}{n}\right)_p \quad (117)$$

Then inequality (5.15) imply:

$$c(p)\omega r\left(f, \frac{1}{n}\right)_p \le c(p)\sum_{i=1}^{n}\|B_n f - f\|_p \quad (118)$$

and by (Lemma:

$$c(p)\omega r\left(f, \frac{1}{n}\right)_p \le \sum_{i=1}^{n}\|N_n f - f\|_p \quad (119)$$

Theorem 5.16 (The saturation problem for the approximation of multivariate function in $L_p(k)$ spaces for p<1 using a forward regular neural network): For any $f \in L_p(k), p<1$ there is a regular, one hidden layer FNN, $N_n(x)$of the form (1.1) with $\sigma\epsilon j_n$ and the hidden unit number such that:

$$m \ge \sum_{k=0}^{n}(k+1)\begin{pmatrix} k+d-1 \\ d-1 \end{pmatrix} = m_0(n) \quad (120)$$

$\|N_n f - B_n f\|_p = \left(n_2^{-\alpha}\right)$ if and only if $f\in Lip(a)_r$ if:

Proof: let $f\in Lip(a)_r$:

$$\|N_n f - f\|_p \le c(p)\omega r\left(f, \frac{1}{n}\right)_p \quad (121)$$

$$\Rightarrow \|N_n - f\|_p = o\left(\frac{1}{n}\right)^{(a)}$$

$$c(p)\omega r\left(f, \frac{1}{n}\right)_p \le \sum_{i=1}^{n} \|N_i f - f\|_p \quad (122)$$

We obtain:

$$\omega r\left(f, \frac{1}{n}\right)_p = o\left(\frac{1}{n^\alpha}\right) \quad (123)$$

CONCLUSION

We define a type of K-functional and a modulus of a smoothness in terms of Dunkl operator, then we introduced a relationship of equivalence between them and we prove direct and inverse estimation and saturation problem for the approximation of multivariate function $L_p(k)$in spaces for p <1 using a forward regular neural network .

REFRENCES

Bateman, H. and A. Erdelyi, 1974. Higher Transcendental Functions. McGraw-Hill, New York, USA.,.

Dai, F., 2003. Some equivalence theorems with K-functionals. J. Approximation Theory, 121: 143-157.

Ditzian, Z. and V. Totik, 1987. Moduli of Smoothness. Springer, Berlin, Germany,.

Dunkl, C.F., 1989. Differential-difference operators associated to reflection groups. Trans. Am. Math. Soc., 311: 167-183.

Konovalov, V.N., D. Leviatan and V.E. Maiorov, 2008. Approximation by polynomials and ridge functions of classes of s-monotone radial functions. J. Approximation Theory, 152: 20-51.

Konovalov, V.N., D. Leviatan and V.E. Maiorov, 2009. Approximation of Sobolev classes by polynomials and ridge functions. J. Approximation Theory, 159: 97-108.

Lofstrom, J. and J. Peetre, 1969. Approximation theorems connected with generalized translations. Math. Ann., 181: 255-268.

Maiorov, V. and A. Pinkus, 1999. Lower bounds for approximation by MLP neural networks. Neurocomputing, 25: 81-91.

Maiorov, V., 2003. On best approximation of classes by radial functions. J. Approximation Theory, 120: 36-70.

Rosler, M., 2003. Dunkl Operators: Theory and Applications. In: Orthogonal Polynomials and Special Functions, Koelink, E. and W.V. Assche (Eds.). Springer, Berlin, Germany, ISBN: 978-3-540-44945-4, pp: 93-135.

Salem, N.B. and S. Kallel, 2004. Mean-periodic functions associated with the Dunkl operators. Integral Transforms Spec. Functions, 15: 155-179.

Xu, Z. and F. Cao, 2004. The essential order of approximation for neural networks. Sci. China Ser. F. Inf. Sci., 47: 97-112.