

# Approximation to Functions and Nonlinear Operators by One Function with Its Translation and Dilation

Chuanhai Jiang      Hongbin Guo

## Abstract

Associated with approximate capability of neural network, we discuss the possibility of approximation to function, nonlinear continuous functionals and operators by superpositions of functions in this paper.

**Key Words:** Functional, Operator, Neural network, System identification

**AMS (1991) Subject Classification:** 41A20

## 1 Introduction and Notations

Concerning with approximating function by sigma-pi neural network, Pinkus [1] discussed TDI-subspace of  $C(R^n)$  and some density problems, i.e suppose that  $g \in C(R^n)$ , set

$$\mu_g = \overline{\text{span}}\{g(\mathbf{A}\mathbf{x} - \mathbf{b})\} \quad (1.1)$$

where  $\mathbf{A}$  is nonsingular diagonal matrix and  $\mathbf{b} \in R^n$ . Then  $\mu_g$  is called the smallest TDI-subspace generated by  $g$ . With the topology of uniform convergence on compact subsets in  $C(R^n)$ , the necessary and sufficient conditions imposed on  $g$  such that  $\mu_g = C(R^n)$  have been obtained. On the other hand Chen and Chen [9] used the property of isomorphism of Fourier transform in  $L^2(R^n)$  to discuss this question in  $L^2(R^n)$ .

The materials described above associate with the approximation to the functions, but in the study of system identification, we always view a system as a nonlinear map from one space to another, then given a system defined by a map  $F$ , the objective of system identification is to build another map  $G$  which will approximate  $F$  well. Therefore it is necessary to study the problem of approximating nonlinear functionals and operators defined on some function space. In [7] Chen and Chen discussed in detail the problem of approximating nonlinear functional and nonlinear operators defined on some compact set in  $C(K)$  by Radial Bases Function(RBF) neural networks, where  $K$  is a compact subset of  $R^n$ , and obtained some explicit results, but the compactness of  $K$  plays an essential role.

Can we give a neural network model to approximate nonlinear functionals or operators defined on some compact set in  $L^p(R^n)$ ? In this paper we will answer this question and give some strong results.

Next all nonsingular diagonal matrixes of order  $n$  will be denoted by  $\mathcal{A}^n$ ,  $\Omega_n$  the volume of unit ball in  $R^n$ , and  $R^* = \{x \in R^n : \text{theresome } x_i = 0\}$ .

If  $f$  is any function on  $R^n$ , and  $a \in R^*$ ,  $b \in R^n$ , Then  $f_b^a$  is defined as following:

$$f_b^a = f(ax + b) \quad \text{for all } x \in R^n;$$

the set  $S_f$  is defined by

$$S_f = \{f_b^a : a \in R^*, b \in R^n\}.$$

If  $f \in L^p(R^n)$ , we shall denote by  $T[f]$  the closed vector subspace of  $L^p(R^n)$  generated by  $S_f$ .

## 2 Approximating Function Theorems

Harisymiv [5] obtained an elegant result and used some delicate techniques to prove it, but his proof is difficult to understand. Next we will relate Harisymiv's Theorem and give our proof.

**Theorem 2.1** *If  $f \in L^p(R^n)$ , ( $1 < p < \infty$ ) and  $f \neq 0$ , then  $T[f] \in L^p(R^n)$ .*

Showing this theorem we need some preliminary work.

Consider any function  $\phi \in \mathcal{D}(R^*)$ , then if  $x \in R^*$ , it follows that  $\phi^{x^{-1}} \in \mathcal{D}(R^*)$ . If  $s \in S'(R^n)$  we define a function  $s \nabla \phi$  on  $R^*$  by

$$s \nabla \phi(x) = s(\phi^{x^{-1}})$$

for all  $x \in R^*$ , then  $s \nabla \phi \in C^\infty(R^*)$  [5]. Thus we have

**Lemma 2.2** *If  $W$  is any relatively compact open set such that  $\bar{W} \subset R^*$ ; if  $\phi \in \mathcal{D}(R^*)$  and  $s \in D'(R^*)$ , then there exists a function  $k(x) \in S(R^n)$  such that On  $W$*

$$s \nabla \phi = \hat{k}$$

**Proof**  $\bar{W} \subset R^*$  implies that there exists open set  $\Omega \subset R^*$  such that  $\bar{W} \subset \Omega \subset R^*$ , therefore there is  $\psi \in C^\infty(R^n)$  satisfying

$$\psi(x) = \begin{cases} 1, & \text{if } x \in \bar{W} \\ 0, & \text{if } x \in R^n \setminus \Omega \end{cases}$$

Let  $k(x) = ((s \nabla \phi) \psi)^\vee$ , here and elsewhere  $^\vee$  denotes the inverse Fourier transform.  $(s \nabla \phi) \psi \in D(R^n)$ , then we have  $k(x) \in S(R^n)$ .

Comparing Lemma 2.2 with Lemma 2.1 in [5], we emphasize that  $k(x) \in S(R^n)$ , which play key role in the following lemma.

**Lemma 2.3** [5] *If  $\phi \in D(R^*)$ ,  $f \in L^p(R^n)$ ,  $g \in L^q(R^n)$ , and if*

$$f^a * g = 0$$

*for all  $a \in R^*$ , then on  $R^*$*

$$\hat{f} \nabla \phi \cdot \hat{g} = 0.$$

**Lemma 2.4** *Suppose that  $f \in L^p(R^n)$  and that  $R^* \cap \text{supp} \hat{f} \neq \emptyset$ . Then if  $g \in L^q(R^n)$  is such that  $f^a * g = 0$  for all  $a \in R^*$ , we have*

$$\text{supp} \hat{g} \subseteq R^n \setminus R^*.$$

**Proof** First we observe that

$$\text{supp} \hat{f}^a = a \cdot \text{supp} \hat{f}$$

and hence, since  $R^* \cap \text{supp} \hat{f} \neq \emptyset$ ,

$$\bigcup_{a \in R^*} \text{supp} \hat{f}^a \supseteq R^*. \quad (2.1)$$

Now suppose that  $x \in R^*$ . By (2.1) there exist  $b$  such that  $x \in \text{supp} \hat{f}^b$ . Choose a relatively compact neighbourhood  $W$  of  $x$  such that  $\bar{W} \subset R^*$ . There exists a function  $\phi \in D(W)$  such that

$$\hat{f}^b(\phi) \neq 0,$$

i.e.

$$\hat{f}^b \nabla \phi^x(x) \neq 0.$$

$\hat{f}^b \nabla \phi^x(y) \in C^\infty(R^*)$  implies that there exists a open neighbourhood  $\Omega$  of  $x$  satisfying  $\bar{\Omega} \subseteq R^*$  such that  $\hat{f}^b \nabla \phi^x \neq 0$  in  $\bar{\Omega}$ . By Lemma 2.2, there is  $k(x) \in S(R^n)$  such that

$\hat{k}(x) = \widehat{f^b} \nabla \phi^x(y)$  on  $\bar{\Omega}$ . By Lemma 2.4, then for any  $\psi \in D(\Omega)$

$$\begin{aligned} \hat{g}(\psi) &= \hat{k}(y) \cdot \hat{g}\left(\frac{1}{\hat{k}(x)}\psi\right) \\ &= \widehat{f^b} \nabla \phi^x(y) \cdot \hat{g}(\tilde{\psi}) \\ &= 0 \end{aligned} \tag{2.2}$$

this implies that  $x \notin \text{supp}\hat{g}$ , thus  $\text{supp}\hat{g} \subseteq R^n \setminus R^*$ .

Now we start to prove Theorem 2.1

**Proof of Theorem 2.1** Suppose that  $T[f]$  is a subset of  $L^p(R^n)$ , by Hahn-Banach extension theorem and Riesz representation theorem, there exists nonzero function  $g(x) \in L^q(R^n)$  such that for any  $a \in R^*$ ,  $b \in R^n$  we have

$$\int_{R^n} f^a(x+b)g(x) = 0 \tag{2.3}$$

i.e.

$$f^a * g = 0. \tag{2.4}$$

Since  $f \in L^p(R^n) \setminus 0$  and  $R^n \setminus R^*$  is  $q$ -thin set [5], if  $1 < q < \infty$ , we have

$$\text{supp}\hat{f} \cap R^* \neq \emptyset.$$

Then, by Corollary 2.5 we can follows that

$$\text{supp}\hat{g} \subset R^n \setminus R^*$$

and so  $g = 0$ . This is in contradict with  $g \neq 0$ . Then we complete the proof of Theorem 2.1.

It can be easily seen that Theorem 2.1 is equivalent to the following theorem:

**Theorem 2.5** Suppose that  $g \in L^p(R^n)$  and  $g \neq 0$ , then for any  $f(x) \in L^p(R^n)$  and any  $\epsilon > 0$ , there exist positive integer  $N$ , non-singular diagonal matrix  $A_i$ ,  $X_i \in$

$R^n$ ,  $i = 1, \dots, N$ , and constants  $c_i, i = 1, \dots, N$ , which all depend on  $f$ , such that

$$\|f(x) - \sum_{i=1}^N c_i(f)g(A_i x + X_i)\|_p < \epsilon. \quad (2.5)$$

But for the functions in a compact set of  $L^p(R^n)$ , we have the following more strong theorem.

**Theorem 2.6** *Suppose that  $V$  is a compact set of  $L^p(R^n)$ ,  $g \in L^p(R^n)$  and  $g \neq 0$ . then for any  $\epsilon > 0$ , there exist positive integer  $N$ , non-singular diagonal matrix  $A_i$ ,  $X_i \in R^n$ ,  $i = 1, \dots, N$ , which are independent of  $f \in V$ , and constants  $c_i(f), i = 1, \dots, N$ , which depend on  $f \in V$ , such that*

$$\|f(x) - \sum_{i=1}^N c_i(f)g(A_i x + X_i)\|_p < \epsilon \quad (2.6)$$

holds for all  $f \in V$ . Moreover, every  $c_i(f)$  is a continuous functional defined on  $V$ .

In order to prove Theorem 2.6, we need some Lemmas.

**Lemma 2.7** *[11]  $V$  is a compact set in  $L^p(R^n)$ , then we have:*

1. *There is a constant  $A > 0$ , such that for all  $f \in V$ ,  $\|f\|_p < A$ ;*
2. *For any  $\epsilon > 0$ , there is a constant  $B > 0$ , such that for all  $f \in V$*

$$\int_{\|x\|>B} |f(x)|^p dx < \epsilon;$$

3. *) For any  $\epsilon > 0$ , there is  $h_0 > 0$ , such that for all  $f \in V$ , provided that  $h < h_0$*

$$\|f(x+h) - f(x)\|_p < \epsilon.$$

Suppose  $f \in L^p(\mathbb{R}^n)$ . For any fixed  $h > 0$ , we let

$$f_h(x) = \frac{1}{m(B(x, h))} \int_{B(x, h)} f(t) dt.$$

where  $B(x, h)$  is the ball in  $\mathbb{R}^n$  with center at  $x$  and radius  $h$ ,  $m(B(x, h))$  is the volume of the ball in  $\mathbb{R}^n$ . Then by the definition of  $f_h(x)$  and Minkowski's inequality for integrals, we can easily prove

**Lemma 2.8**

1.  $\|f_h\|_p \leq \|f\|_p$ .
2. For any  $\epsilon > 0$ , there is  $h_0 > 0$ , if  $h < h_0$ , we have  $\|f_h - f\|_p < \epsilon$ .
3. Suppose that  $V$  is a compact set in  $L^p(\mathbb{R}^n)$ , for any fixed  $h > 0$ , let  $V_h = \{f_h : f \in V\}$ . Then there exists constant  $A_h > 0$  such that for all  $f_h \in V_h$ ,  $|f_h(x)| \leq A_h$ .
4. For any  $\epsilon > 0$  there exists  $\delta(\epsilon)$ , if  $\|x - y\| < \delta(\epsilon)$ , then for all  $f_h \in V_h$ ,  $|f_h(x) - f_h(y)| < \epsilon$  uniformly holds, where  $A_h, \delta(\epsilon)$  are independent of  $f_h$ .

With these lemmas in hand, next we will prove Theorem 2.6.

**Proof of Theorem 2.6** For any fixed  $\epsilon > 0$ , by (2) of Lemma 2.8, there exists  $h_0 > 0$ , such that

$$\|f_h - f\|_p < \frac{\epsilon}{8}. \tag{2.7}$$

Combining this inequality with (2) of Lemma 2.7, we conclude that for previous  $\epsilon$  and some fixed  $h < h_0$ , there is  $B > 0$  such that

$$\left( \int_{\|x\| \geq B} |f_h(x)|^p dx \right)^{\frac{1}{p}} < \frac{\epsilon}{4} \tag{2.8}$$

holds for all  $f_h \in V$ .

Let  $\phi(x) = ce^{-\|x\|_{R^n}}$ , satisfying  $\int_{R^n} \phi(x)dx = 1$ , and  $\phi_\delta(x) = \delta^{-n}\phi(\delta^{-1}x)$ . It is well known that there is  $\delta_0 > 0$ , such that for  $\delta < \delta_0$  and all  $f \in V$

$$\|f * \phi_\delta - f\|_p < \frac{\epsilon}{8}. \quad (2.9)$$

By Young inequality and Lemma 2.8, when  $h < h_0$  and for all  $f \in V$

$$\|f * \phi_\delta - f_h * \phi_\delta\|_p \leq \|f - f_h\|_p \leq \frac{\epsilon}{8}. \quad (2.10)$$

From (2.9) (2.10), for all  $f \in V$  we have

$$\|f - f_h * \phi_\delta\| \leq \frac{\epsilon}{4}. \quad (2.11)$$

Combining (2.8) with (2.11) and the fact that  $\int_{R^n} \phi_\delta(x)dx = 1$ , for all  $f \in V$  we have

$$\|f(x) - \int_{\|t\| < B} f_h(t)\phi_\delta(x-t)dt\|_p < \frac{\epsilon}{2}. \quad (2.12)$$

For the fixed  $h$  and  $\delta$ , we can approximate the integral  $\int_{\|t\| \leq B} f_h(t)\phi_\delta(x-t)dt$  by Riemann sum  $\sum_{i=1}^M f(t_j)\phi_\delta(x-t_j)m(\Delta t_j)$ , where  $\bigcup_{i=1}^M \Delta t_j$  is a division of the ball  $\|t\|_{R^n} \leq B$ ,  $t_j \in \Delta t_j$ , then we have

$$\begin{aligned} & \int_{\|t\| \leq B} f_h(t)\phi_\delta(x-t)dt - \sum_{i=1}^M f_h(t_j)\phi_\delta(x-t_j)m(\Delta t_j) \\ &= \sum_{i=1}^M \int_{\Delta t_j} [f_h(t)\phi_\delta(x-t) - f_h(t_j)\phi_\delta(x-t_j)]dt \\ &= \sum_{i=1}^M \int_{\Delta t_j} [f_h(t) - f_h(t_j)]\phi_\delta(x-t)dt \\ &+ \sum_{i=1}^M f_h(t_j) \int_{\Delta t_j} [\phi_\delta(x-t) - \phi_\delta(x-t_j)]dt \\ &= I_1 + I_2. \end{aligned}$$

Now we are to estimate  $\|I_1\|_p$  and  $\|I_2\|_p$ , by Minkowski inequality , we have

$$\begin{aligned}\|I_1\|_p &\leq \sum_{i=1}^M \left( \int_{R^n} \left( \int_{\Delta t_j} [f_h(t) - f_h(t_j)] \phi_\delta(x-t) dt \right)^p dx \right)^{\frac{1}{p}} \\ &\leq \sum_{i=1}^M \int_{\Delta t_j} \left( \int_{R^n} |[f_h(t) - f_h(t_j)] \phi_\delta(x-t)|^p dx \right)^{\frac{1}{p}} dt.\end{aligned}$$

For small enough fixed  $h$  and  $\delta$ , by Lemma 2.8, there is  $\eta > 0$  if  $\text{diam}(\Delta t_j) < \eta$ , from the previous inequality we can conclude  $\|I_1\|_p < \epsilon/8$ . Similarly, we can also obtain  $\|I_2\|_p < \epsilon/8$ .

Combining with (2.12), for all  $f \in V$  we have

$$\|f(x) - \sum_{i=1}^M f_h(t_j) m(\Delta t_j) \phi_\delta(x-t_j)\|_p < \frac{3\epsilon}{4}. \quad (2.13)$$

By Theorem 2.5 for each fixed  $j$ ,  $j = 1, \dots, M$ , there are positive integer  $N_j$ , constants  $d_{ij}$ , non-singular diagonal matrixes  $A_{ij}$ , vectors  $X_{ij} \in R^n$ ,  $i = 1, \dots, N_j$  such that

$$\|\phi_\delta(x-t_j) - \sum_{i=1}^{N_j} d_{ij} (A_{ij}x - X_{ij})\|_p < \frac{\epsilon}{4L} \quad (2.14)$$

where  $L = \sup_{f \in V} \sum_{j=1}^M |f_h(t_j)| m(\Delta t_j)$ .

Substituting (2.14) into (2.13) we conclude that there are  $N = \sum_{j=1}^M N_j$ , positive numbers  $\lambda_j$ , non-singular diagonal matrixes  $A_i$ , vectors  $X_i$ ,  $i = 1, \dots, N$ , which are independent of  $f \in V$  and constants  $c_i(f)$ ,  $i = 1, \dots, N$  depend on  $f \in V$ , such that

$$\|f(x) - \sum_{i=1}^N \lambda_i g(A_i x - X_i)\|_p \leq \epsilon \quad (2.15)$$

for all  $f \in V$ . From the process of the above proof, we see that all  $c_i(f)$  are linear combination of  $f_h(t_j)$ , i.e, linear continuous functionals of  $f$ . Theorem 2.6 is completely proved.

In the following, we can see that Theorem 2.6 play an important role in the application of approximation of operators.

### 3 Approximating Functionals and Operators

Taking advantage of Stone-Weierstrass theorem we discuss problems of approximating continuous operators from  $L^{p_1}(R^{n_1})$  to  $L^{p_2}(R^{n_2})$  in this section.

Suppose that  $V$  is a compact set of  $L^p(R^n)$ , ( $1 < p < \infty$ ), non-zero  $g(x) \in L^q(R^n)$ , ( $1/p + 1/q = 1$ ). From Theorem 2.5 we know that the family of all finite linear combinations  $\sum_{i=1}^N c_i g(A_i x + b_i)$  is dense in  $L^q(R^n)$ , where  $A_i \in \mathcal{A}_n$ ,  $c_i \in R$  and  $b_i \in R^n$ . We denote by

$$S := \{G(x) : G(x) = \sum_{i=1}^N c_i g(A_i x + b_i), A_i \in \mathcal{A}_n, c_i \in R^1, b_i \in R^n\}, \quad (3.1)$$

then we have

**Lemma 3.1** *Suppose that  $f_1$  and  $f_2$  are distinct elements of  $L^p(R^n)$ ,  $r_1$  and  $r_2$  are distinct real numbers, (let  $r_1 = r_2$  if  $f_1 = f_2$ ). Then there exist real number  $\rho$  and  $G(x) \in S$  such that*

$$\rho + \int_{R^n} G(x) f_1(x) dx = r_1, \quad (3.2)$$

$$\rho + \int_{R^n} G(x) f_2(x) dx = r_2. \quad (3.3)$$

**Proof** From the hypotheses we know  $f_1 - f_2 \neq 0$ , since  $S$  is dense in  $L^q(R^n)$ , there exists  $Q(x) \in S$  such that

$$\int_{R^n} Q(x) (f_1(x) - f_2(x)) dx \neq 0. \quad (3.4)$$

Let

$$G(x) = \frac{r_1 - r_2}{\int_{R^n} Q(x)(f_1(x) - f_2(x))dx} Q(x), \quad (3.5)$$

then  $G(x) \in S$  and

$$\int_{R^n} G(x)(f_1(x) - f_2(x))dx = r_1 - r_2. \quad (3.6)$$

Set  $\rho = r_1 - \int_{R^n} G(x)f_1(x)dx$ , clearly  $\rho + \int_{R^n} G(x)f_2(x)dx = r_2$ . Lemma 3.1 is proved.

Let  $A$  stands for the set of all  $R$ -valued continuous functional defined on  $V$ , and

$$A_0 := \{h : V \mapsto R^1, h(f) = \rho + \int_{R^n} G(x)f(x)dx, G(x) \in S, \rho \in R^1\}, \quad (3.7)$$

then  $A_0$  is a subset of  $A$ . We want to take advantage of Stone-Weierstrass theorem to construct approximations to continuous functionals defined on  $V$ .

The lattice operations  $a \vee b$  and  $a \wedge b$  on pairs of elements  $a$  and  $b$  of  $A$  are defined by  $(a \vee b)(x) = \max[a(x), b(x)]$  and  $(a \wedge b)(x) = \min[a(x), b(x)]$ . By Stone-weierstrass theorem, we can easily conclude the following lemma.

**Lemma 3.2** *Let  $A$  be the family of continuous  $R$ -valued functions defined on compact set  $V$ ,  $A_0$  an arbitrary subfamily of  $A$ , and  $U(A_0)$  the family of all functions generated from  $A_0$  by the finite lattice operations. Suppose that  $f \in A$  has the property that for any  $v$  and  $w$  in  $V$  there exists a function  $h \in A_0$  such that  $h(v) = f(v)$  and  $h(w) = f(w)$ , then  $f \in \overline{U(A_0)}$ .*

**Lemma 3.3** *[8][1] Suppose that  $K$  is a compact set in  $R^n$ .  $g(x) \in C(R^1)$  and is not a polynomial, then for any  $\epsilon > 0$ , there exist a positive integer  $N$ , real numbers  $\theta_i$ , vectors  $\omega_i \in R^n$  and constants  $c_i$ ,  $i = 1, 2, \dots, N$ , such that*

$$|f(x) - \sum_{i=1}^N c_i g(\omega_i \cdot x + \theta_i)| < \epsilon \quad (3.8)$$

holds for  $x \in K$  and  $f(x) \in C(K)$ .

**Theorem 3.4** *Suppose that  $V$  is a compact set of  $L^p(\mathbb{R}^n)$ , nonzero  $g_1(x) \in L^q(\mathbb{R}^n)$ , ( $1/p + 1/q = 1$ ),  $g_2(x) \in C(\mathbb{R}^1)$  and is not a polynomial,  $S$  and  $A_0$  are defined as above.  $F$  is a continuous functional defined on  $V$ , then for any  $\epsilon > 0$ , there exist positive integers  $N, m, h_j \in A_0, \lambda_i \in \mathbb{R}^m, c_i, \theta_i \in \mathbb{R}^1, i = 1, \dots, N, j = 1, \dots, m$ , such that*

$$|F(f) - \sum_{i=1}^N c_i g_2(\lambda_i \cdot x^m + \theta_i)| < \epsilon, \quad (3.9)$$

where  $x^m = (h_1(f), \dots, h_m(f))'$ , and  $h_j(f) = \rho_j + \int_{\mathbb{R}^n} G_j(x) f(x) dx, G_j(x) \in S$ .

**Proof** For any  $\epsilon > 0$ , by Lemma 3.1 and Lemma 3.2 there exist positive  $m$ , elements  $h_j \in A_0, j = 1, \dots, m$ , and a lattice map  $L$ , such that

$$|F(f) - L(h_1, \dots, h_m)| < \frac{\epsilon}{2} \quad (3.10)$$

holds for  $f \in V$ . By the definition of  $h_j$ , we know that  $h_j$  is a continuous map defined on  $V$ , so  $K_j := h_j(V)$  is a compact set of  $\mathbb{R}^1$ . Let  $K := K_1 \times \dots \times K_m$ , then  $K$  is a compact set of  $\mathbb{R}^m$ . Therefore,  $L$  can be viewed as a continuous function defined on  $K$ , by Lemma 3.3, for above  $\epsilon > 0$ , there exist positive integer  $N$ , real numbers  $c_i, \theta_i, \lambda_i \in \mathbb{R}^m, i = 1, \dots, N$ , such that

$$|L(h_1, \dots, h_m) - \sum_{i=1}^N c_i g_2(\lambda_i \cdot x^m + \theta_i)| < \frac{\epsilon}{2} \quad (3.11)$$

holds for  $f \in V$ . Combining with (3.10), we see that

$$|F(f) - \sum_{i=1}^N c_i g_2(\lambda_i \cdot x^m + \theta_i)| < \epsilon \quad (3.12)$$

holds for  $f \in V$ . Theorem 3.4 is proved.

**Lemma 3.5** [7] [1] *Suppose that  $g \in C(R_+)$  and is not an even polynomial,  $K$  is a compact set in  $R^n$ , then for any  $\epsilon > 0$ , there exist a positive integer  $N$ , real number  $c_i, \lambda_i, \theta_i \in R^n, i = 1, \dots, N$ , such that*

$$|f(x) - \sum_{i=1}^N c_i g(\lambda_i \|x - y_i\|)| < \epsilon \quad (3.13)$$

*holds for all  $x \in K$  and  $f \in C(K)$ .*

Similar to Theorem 3.4, we can prove

**Theorem 3.6** *Suppose that  $V$  is a compact set of  $L^p(R^n)$ ,  $g_1(x) \in L^q(R^n)$  and  $g_1(x) \neq 0$ , ( $1/p + 1/q = 1$ ),  $g_2(x) \in C(R^1)$  and is not an even polynomial,  $S$  and  $A_0$  are defined as above.  $F$  is a continuous functional defined on  $V$ , then for any  $\epsilon > 0$ , there exist positive integers  $N, m, h_j \in A_0, c_i, \lambda_i \in R^1, \theta_i \in R^n, i = 1, \dots, N, j = 1, \dots, m$ , such that*

$$|F(f) - \sum_{i=1}^N c_i g_2(\lambda_i \|x^m + \theta_i\|)| < \epsilon, \quad (3.14)$$

*where  $x^m = (h_1(f), \dots, h_m(f))'$ , and  $h_j(f) = \rho_j + \int_{R^n} G_j(x) f(x) dx, G_j(x) \in S$ .*

From now on we start to discuss approximation to nonlinear continuous operators from  $L^{p_1}(R^{n_1})$  to  $L^{p_2}(R^{n_2})$ .

**Theorem 3.7** *Suppose that  $V$  is a compact set of  $L^{p_1}(R^{n_1})$ ,  $T$  is a continuous operator mapping  $V$  into  $L^{p_2}(R^{n_2})$ ,  $g_1(x) \in C(R^1)$  and is not a polynomial, nonzero  $g_2 \in L^{p_2}(R^{n_2})$ , nonzero  $g_3(x) \in L^{q_2}(R^{n_2}), (1/p_2 + 1/q_2 = 1)$ ,  $S = \{G(x) : G(x) = \sum_{i=1}^N c_i g_3(A_i x + b_i), A_i \in \mathcal{A}^{n_2}, b_i \in R^{n_2}\}$ ,  $A_0 = \{h : T(V) \mapsto R^1, h(f) = \rho + \int_{R^{n_2}} G(x) f(x) dx, G(x) \in S, f(x) \in T(V), \rho \in R^1\}$ . Then for any  $\epsilon > 0$ , there exist positive integers  $N, N_i, m_i,$*

$A_i \in \mathcal{A}^{n_2}$ ,  $b_i \in R^{n_2}$ , real numbers  $c_{i,k}$ ,  $\theta_{i,k}$ ,  $\lambda_{i,k} \in R^{m_i}$ ,  $k = 1, \dots, N_i$ ,  $j = 1, \dots, m_i$ , such that

$$\|T(u)(y) - \sum_{i=1}^N \sum_{k=1}^{N_i} c_{i,k} g_1(\lambda_{i,k} \cdot x^{m_i} + \theta_{i,k}) g_2(A_i x + b_i)\|_{L^{p_2}(R^{n_2})} < \epsilon \quad (3.15)$$

holds for all  $u \in V$ .

**Proof**  $T$  is a continuous operator mapping  $V$  into  $L^{p_2}(R^{n_2})$ , then  $T(V) := \{T(u) : u \in V\}$  is a compact set of  $L^{p_2}(R^{n_2})$ . By Theorem 2.6, for any  $\epsilon > 0$ , there exist positive integer  $N$ ,  $A_i \in \mathcal{A}^{n_2}$ ,  $b_i \in R^{n_2}$ ,  $i = 1, \dots, N$ , which all are independent of  $T(u) \in T(V)$ , and constants  $\sigma_i(T(u))$  depending on  $T(u)$ , such that

$$\|T(u)(y) - \sum_{i=1}^N \sigma_i(T(u)) g_2(A_i x + b_i)\|_{L^{p_2}} < \frac{\epsilon}{2} \quad (3.16)$$

holds for all  $u \in V$ . Moreover, every  $\sigma_i(T(u))$  is a continuous functional defined on  $V$ . For each  $\sigma_i(T(u))$ , by Theorem 3.4, there exist positive integers  $N_i$ ,  $m_i$ ,  $h_{i,j} \in A_0$ , real numbers  $c_{i,k}$ ,  $\theta_{i,k}$ ,  $\lambda_{i,k} \in R^{m_i}$ ,  $k = 1, \dots, N_i$ ,  $j = 1, \dots, m_i$ , such that

$$|\sigma_i(T(u)) - \sum_{k=1}^{N_i} c_{i,k} g_1(\lambda_{i,k} \cdot x^{m_i} + \theta_{i,k})| < \frac{\epsilon}{2L}, \quad (3.17)$$

where  $x^{m_i} = (h_{i,1}(T(u)), \dots, h_{i,m_i}(T(u)))'$ ,  $h_{i,j} \in A_0$ . Because  $N, A_i, b_i$  are independent of  $T(u) \in T(V)$ , we can set

$$L = \sum_{i=1}^N \|g_2(A_i x + b_i)\|_{L^{p_2}(R^{n_2})}$$

. Combining (3.17) with (3.16) we have

$$\|T(u)(y) - \sum_{i=1}^N \sum_{k=1}^{N_i} c_{i,k} g_1(\lambda_{i,k} \cdot x^{m_i} + \theta_{i,k}) g_2(A_i x + b_i)\|_{L^{p_2}(R^{n_2})} < \epsilon \quad (3.18)$$

holds for all  $u \in V$ . Theorem 3.7 is proved.

Similar to the proof of Theorem 3.7, we can prove

**Theorem 3.8** Suppose that  $V$  is a compact set of  $L^{p_1}(R^{n_1})$ ,  $T$  is a continuous operator mapping  $V$  into  $L^{p_2}(R^{n_2})$ ,  $g_1(x) \in C(R^1)$  and is not an even polynomial, nonzero  $g_2 \in L^{p_2}(R^{n_2})$ , nonzero  $g_3(x) \in L^{q_2}(R^{n_2})$ ,  $(1/p_2 + 1/q_2 = 1)$ ,  $S = \{G(x) : G(x) = \sum_{i=1}^N c_i g_3(A_i x + b_i), A_i \in \mathcal{A}^{n_2}, b_i \in R^{n_2}\}$ ,  $A_0 = \{h : T(V) \mapsto R^1, h(f) = \rho + \int_{R^{n_2}} G(x)f(x)dx, G(x) \in S, f(x) \in T(V), \rho \in R^1\}$ . Then for any  $\epsilon > 0$ , there exist positive integers  $N, N_i, m_i, A_i \in \mathcal{A}^{n_2}, b_i \in R^{n_2}$ , real numbers  $c_{i,k}, \theta_{i,k}, \lambda_{i,k}$ ,  $k = 1, \dots, N_i, j = 1, \dots, m_i$ , such that

$$\|T(u)(y) - \sum_{i=1}^N \sum_{k=1}^{N_i} c_{i,k} g_1(\lambda_{i,k} \|x^{m_i} + \theta_{i,k}\|) g_2(A_i x + b_i)\|_{L^{p_2}(R^{n_2})} < \epsilon \quad (3.19)$$

holds for all  $u \in V$ .

## References

- [1] A. Pinkus, "TDI-subspace of  $C(R^d)$  and some density problems from neural networks", *J. Approx. Theory*, 85, 1996, 269-287.
- [2] G. Cybenko, "Approximation by superposition of a sigmoidal function", *mathematics of Control, Signals and Systems* vol.2, no.4, 1989, pp 304-314.
- [3] H. N. Mhaskar and C. A. Micchelli, "Approximation by superposition of sigmoidal and radial basis functions", *Advance Applied Mathematics*, vol. 13, 1992, 350-373
- [4] I. W. Sandberg, "Structure theorems for nonlinear systems", *Multidimensional Systems and Signal Processing*, vol. 2, no. 3, 1991, 267-286.
- [5] S. R. Harasymiv, "A note of dilations in  $L^p$ ", *Pacific J. of Math.*, 21 1967, 493-501.

- [6] T. Chen, H. Chen, and R-W. Liu, "Approximation capability in  $C(R^n)$  by multilayer feedforward networks and related problems", *IEEE Trans. Neural Networks*, vol.6 no. 1 Jan. 1995.
- [7] T. Chen and H. Chen, "Approximation capability to functions of several variables, nonlinear functionals and operators by radial basis function neural networks", *IEEE Trans. on Neural Networks*, vol. 6, no. 4, 1995, 904-910.
- [8] T. Chen and H. Chen, "Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems", *IEEE Trans. on Neural Networks*, vol. 6, no. 4, 1995, 911-917.
- [9] T. Chen and H. Chen, "Denseness of radial basis function in  $L^2(R^n)$  and its application in neural networks", *Chin. Ann.of Math.*, 17B: 2, 1996, 219-226.
- [10] T. Chen and H. Chen, "Universal approximation capability of EBF neural networks with arbitrary activation functions", *Circuits Systems, Signal Processing*, vol.15, no. 5, 1996, 671-683.
- [11] M. Cheng, *et*, Real Analysis, High Education Press. 1993.
- [12] Stein, E. M., and Weiss, G., Introduction to Fourier analysis on Euclidean space, Princeton University Press. 1971.

Chuanhai Jiang

Dept.of Information, Shanghai University of Finance and Economics,  
Shanghai,P.R. of China 200433

Hongbin Guo

Institute of Mathematics, Fudan University, Shanghai, P.R. of China 200433  
e-mail: 970008@fudan.edu.cn