Neurocomputing 74 (2010) 457–460

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Approximation capability of interpolation neural networks ☆

Feilong Cao [a,]*, Shaobo Lin [b], Zongben Xu [b]

[a] Department of Mathematics, China Jiliang University, Hangzhou 310018, Zhejiang Province, PR China
[b] Institute for Information and System Sciences, Xi'an Jiaotong University, Xi'an 710049, Shannxi Province, PR China

ARTICLE INFO

Article history:
Received 31 January 2010
Received in revised form
26 August 2010
Accepted 29 August 2010
Communicated by G.-B. Huang
Available online 19 October 2010

Keywords:
Neural networks
Best approximation
Interpolation

ABSTRACT

It is well-known that single hidden layer feed-forward neural networks (SLFNs) with at most n hidden neurons can learn n distinct samples with zero error, and the weights connecting the input neurons and the hidden neurons and the hidden node thresholds can be chosen randomly. Namely, for n distinct samples, there exist SLFNs with n hidden neurons that interpolate them. These networks are called exact interpolation networks for the samples. However, for some approximated target functions (as continuous or integrable functions) not all exact interpolation networks have good approximation effect. This paper, by using a functional approach, rigorously proves that for given distinct samples there exists an SLFN which not only exactly interpolates samples but also near best approximates the target function.

1. Introduction

An important reason for the popularity of feed-forward neural networks in many applications is their universal approximation property. The single hidden layer feed-forward neural networks (SLFNs) are mathematically expressed as

$$N_{\sigma,n}(x) := \sum_{k=1}^{n} c_k \sigma(w_k \cdot x + \theta_k), \tag{1}$$

where the output weights $c_k \in \mathbb{R}$, the thresholds $\theta_k \in \mathbb{R}$, $x \in \mathbb{R}^d$ denotes the input to the network, and w_k's are the input weights. In many fundamental network models, the activation function σ of the networks satisfies the so-called sigmoidal condition, i.e.

$$\lim_{x\to\infty} \sigma(x) = 1, \quad \lim_{x\to-\infty} \sigma(x) = 0.$$

Theoretically, by a proper choice of the thresholds and weights, the SLFNs can approximate any continuous target function on any compact set with arbitrary small error. It was proved by Cybenko [5] and Funahashi [6] that any continuous function can be approximated on a compact set with uniform topology by a network of the form given in Eq. (1), using any continuous and sigmoidal activation function. Hornik et al. [8] showed that any measurable function can be approximated with such a network. Various density results on SLFN approximations were later established by many authors using various methods, for more or less general situations: [18,4,3,9,10,13].

In applications SLFNs are trained using finite input samples. It is known that n arbitrary distinct samples (x_i, f_i) $(i=1,2,\ldots,n)$ can be learned precisely by SLFNs with n hidden neurons. These networks are also called exact interpolation networks for the samples. Recently, a novel learning method for SLFNs named extreme learning machine (ELM) algorithm was proposed by Huang et al., and which has been widely studied in [12–15]. As indicated in [11–15], the SLFNs with at most n hidden neurons can learn n distinct samples with zero error, and the weights connecting the input neurons and the hidden neurons and the hidden node thresholds can be chosen randomly.

Several proofs on the existence of exact interpolation networks have been proposed in [16,17,22,23]. Recently, Llanas and Sainz [19] studied the existence of exact interpolation networks and the construction of approximate interpolation networks. Llanas and Lantarón [20] studied Hermite interpolation by SLFNs. In [1] a type of SLFNs, which could be used to approximately interpolate any set of distinct data with arbitrary precision, was presented, and the modulus of continuity of function was used as a metric to characterize the rate of convergence of the approximate interpolation networks.

Although SLFNs with n neurons can exactly interpolate n samples, neither exact interpolation networks nor approximate interpolation networks guarantees the good approximation properties of the interpolants. In some practical applications, ones usually need to find a tool that learns the given samples very well, and simultaneously approximates target function in a prescribed error. So it is natural to raise the question: Can we find an SLFN that not only exactly interpolates the given samples but also simultaneously approximates the target function very well? The main

☆ The research was supported by the Natural Science Foundation of Zhejiang Province of China (no. Y7080235).

* Corresponding author.
E-mail address: feilongcao@gmail.com (F. Cao).

0925-2312/$ - see front matter
doi:10.1016/j.neucom.2010.08.018

purpose of this paper is to give an affirmative answer to the question. Namely, we will prove that there exist SLFNs which exact interpolate any set of distinct data and simultaneously, near best approximate the target function.

Our results show that such exact interpolation SLFN exists, and the number of neurons depends on the separation radius of the interpolation nodes. Concretely, if the interpolation nodes are near equally distribute on [0,1], i.e. the separation radius q of the interpolation nodes satisfy $q \sim 1/n$, where n is the number of the interpolation sample, then the number of neurons of the exact interpolation networks, N, satisfies $N \sim n$.

This paper is organized as follows. In Section 2 we will state the main results of this paper. In Section 3 we will draws a conclusions of this paper. The proofs of main results will be given in the Appendix.

2. Main results

Let $\mathbb{N}$ and $\mathbb{R}$ denote the set of natural numbers and real numbers, respectively. Let $X = \{x_0, x_1, \ldots, x_n\} \subset \mathbb{R}$ denote a set of distinct data, and $\{f_0, f_1, \ldots, f_n\} \subset \mathbb{R}$ denote a set of real numbers. Then

$$\{(x_0,f_0),(x_1,f_1),\ldots,(x_n,f_n)\} \tag{2}$$

is called the set of interpolation samples, and $\{x_i\}_{i=0}^n$ is called the node system of interpolation.

If there exists an SLFN, $N_{\sigma,n}$, formed as (1) such that

$$N_{\sigma,n}(x_i) = f_i, \quad i = 0,1,\ldots,n, \tag{3}$$

then we say that $N_{\sigma,n}$ is an exact interpolation network for the sample set (2). If there exists an SLFN, $N_{\sigma,n}$, formed as (1) such that

$$|N_{\sigma,n}(x_i) - f_i| < \varepsilon, \quad i = 0,1,\ldots,n$$

for positive real number $\varepsilon > 0$, then we call $N_{\sigma,n}$ an approximate interpolation network for the sample set (2).

Suppose that function f is continuous on [0,1], and $\{x_j\}_{j=0}^n \subseteq [0,1]$. Let $\Phi_{\sigma,n}$ be the set of all SLFNs formed as (1). We say that an SLFN, $N_{\sigma,n}$, near best approximates f if there exists an absolute constant C such that

$$\|f(\cdot) - N_{\sigma,n}(\cdot)\| \le C E_{\sigma,n}(f), \tag{4}$$

where $\|\cdot\|$ denotes the uniform norm on [0,1] defined by $\|f\| = \max_{x \in [0,1]} |f(x)|$, and

$$E_{\sigma,n}(f) = \inf_{g \in \Phi_{\sigma,n}} \|f(\cdot) - g(\cdot)\|$$

denotes the best approximation of f by SLFNs in $\Phi_{\sigma,n}$.

Our aim is to prove that there exist SLFNs satisfying (3) and (4). The main result is as follows.

Theorem 1. *Let $X = \{x_0, x_1, \ldots, x_n\} \subseteq [0,1]$ be a set of distinct points, and σ be a bounded sigmoidal function on $\mathbb{R}$. Denote by $q_X = \frac{1}{2}\min_{i \ne j} |x_i - x_j|$ the separation radius of X. Then for every continuous function f defined on [0,1] there exists an SLFN, $N_{\sigma,L} \in \Phi_{\sigma,L}$, satisfies (3) and*

$$\|f - N_{\sigma,L}\| \le 5 E_{\sigma,L}(f), \tag{5}$$

where $L = \lceil (6(1+\|\sigma\|))/q_X \rceil$, $\lceil t \rceil$ denotes the smallest integer number not smaller than t, and $\|\sigma\| = \max_{t \in \mathbb{R}} |\sigma(t)|$.

We will give the proof of Theorem 1 in the appendix.

To characterize the degree of approximation by SLFNs, we need to introduce the modulus of smoothness of continuous function f as

$$\omega(f,t) = \sup_{0 \le h \le t} \max_{x, x+h \in [0,1]} |f(x+h) - f(x)|.$$

This modulus is usually used as a tool for measuring approximation error. It is also used to measure the smoothness of a function and its accuracy in approximation theory and Fourier analysis (see [7]). The function f is called (M,α)- Lipschitz continuous $(0 < \alpha \le 1)$, which can be written as $f \in \mathrm{Lip}(M,\alpha)$, if and only if there exists a constant $M > 0$ such that

$$\omega(f,\delta) \le M\delta^\alpha.$$

The following conclusion was established by Chen in [2]: if σ is a bounded sigmoidal function, then for $f \in C[0,1]$ there exists an SLFN, $N_{\sigma,n} \in \Phi_{\sigma,n}$, such that

$$\|f - N_{\sigma,n}\| \le (1+\|\sigma\|)\,\omega\left(f, \frac{1}{n}\right), \tag{6}$$

where $C[0,1]$ is the space of continuous functions defined on [0,1].

Theorem 1 together with (6) yields the following Corollary 1.

Corollary 1. *Under the conditions of Theorem 1, there exists an SLFN, $N_{\sigma,L} \in \Phi_{\sigma,L}$, which satisfies (3) and*

$$\|f - N_{\sigma,L}\| \le 5(1+\|\sigma\|)\,\omega\left(f, \frac{1}{L}\right),$$

where L is the same as that in (5).

Theorem 1 gives the interpolation and approximation properties of the SLFN with bounded sigmoidal activation function. Here we introduce another type of SLFNs with the activation functions of analytic and non-polynomial function, which has the same interpolation and approximation properties as the SLFNs stated in Theorem 1.

The following result has been proved in [24].

Proposition 1. *Suppose that σ has up to $n+1$ orders of bounded derivatives on [0,1], and is not a polynomial with degree at most $n+1$. Then for $f \in C[0,1]$, there exists an SLFN, $N_{\sigma,n}$, such that*

$$E_{\sigma,n}(f) \le \omega\left(f, \frac{1}{n}\right). \tag{7}$$

Indeed, from the proof of Theorem 1 in the appendix, we know that for any collection of functions, if the Jackson-type inequality (7) holds, then we can get the following interpolation properties of such collection of functions by using the same method in proving Theorem 1.

Theorem 2. *Let $X = \{x_0, x_1, \ldots, x_n\} \subseteq [0,1]$ be a set of distinct points. Suppose that σ has up to $n+1$ orders of bounded derivatives on [0,1], and is not a polynomial with degree at most $n+1$. Then for $f \in C[0,1]$ there exists an SLFN, $N_{\sigma,L'} \in \Phi_{\sigma,L'}$, satisfies (3) and*

$$\|f - N_{\sigma,L'}\| \le 5 E_{\sigma,L'}(f),$$

where $L' = \lceil 6/q_X \rceil$, and $q_X = \frac{1}{2}\min_{i \ne j} |x_i - x_j|$.

If the points in X satisfy the distribution: $q_X \sim 1/n$, where $A \sim B$ means there exists an absolute constant C such that $C^{-1}A \le B \le CA$, then the number of neurons of the interpolation network, L, satisfies $L \sim n$. By the definition of q_X, we know that the condition $q_X \sim 1/n$ shows that the nodes of interpolation arrange neither too dense nor too sparse. The best case of such arrangement is $x_j = j/n$, $j = 0,\ldots,n$.

The following Corollary 2 is a special case of Corollary 1.

Corollary 2. *Let $X = \{x_0, x_1, \ldots, x_n\} \subseteq [0,1]$ be a finite set of distinct points with $q_X \sim 1/n$ and σ be a bounded sigmoidal function. Then for $f \in \mathrm{Lip}(M,1)$ there exists an SLFN $N_{\sigma,N}$ such that $N_{\sigma,N}$ interpolates f on X and satisfies*

$$\|f - N_{\sigma,N}\| \le \frac{5M(1+\|\sigma\|)}{N},$$

where N is a positive integral number and $N \sim n$.

3. Conclusions

Theoretically, SLFNs with at most $n+1$ hidden neurons can learn $n+1$ distinct samples with zero error, and the weights connecting the input neurons and the hidden neurons can be chosen "almost" arbitrarily. That is, there are exact interpolation networks for the given distinct samples. However, for some approximated target functions, such as continuous or integrable functions, not all exact interpolation networks have good approximation effect.In this paper, by using functional approach, it has been rigorously proved that for $n+1$ arbitrary distinct samples (x_i, f_i) $(x_i \in [0,1], f_i \in \mathbb{R}, i = 0,1,2,\ldots n)$, there is an SLFN with at most $L = \lceil (6(1+\|\sigma\|))/q_X \rceil$ hidden neurons and with any bounded sigmoidal activation function can learn these distinct samples with zero error. Simultaneously, this SLFN is the near best approximant for continuous target functions. For analytic non-polynomial activation function, the similar results have also been obtained. The obtained results show that such exact interpolation SLFN with good approximation effect not only exists, but also the number of neurons of the SLFN depends on the separation radius of the interpolation nodes.

From the view point of engineering, the training data set may be the input-output data pairs of a system. Then the trained SLFN can be used as the system model. So a main problem is if the SLFN system model can retain most important dynamics of the system, such as robustness with respect to different input disturbances and so on.

To solve the question, an approach used in applications is that, in addition to the input-output training data pairs, one still needs to add some constraints, reflecting some important dynamic behaviors of the system, to the optimization process. In such a way, the trained SLFN can sufficiently represent the system. In fact, the performance of SLFNs is not only related to the output weight matrix, also related to the selection of the input weight matrix as well. How to properly select the input weights is also important for the training of SLFNs. It is believed that some further work involving the selection of the input weights may get the better optimization results for applications.

Acknowledgement

The authors wish to thank the referees for their helpful suggestions.

Appendix A

The proof of the main result is based on the following lemma, which was stated in [21].

Lemma 1. *Let $\mathcal{Y}$ be a (possibly complex) Banach space, $\mathcal{V}$ a subspace of $\mathcal{Y}$, and Z^* a finite-dimensional subspace of $\mathcal{Y}^*$, the dual of $\mathcal{Y}$. If for every $z^* \in Z^*$ and some $\gamma > 1$, γ independent of z^*,*

$$\|z^*\|_{\mathcal{Y}^*} \le \gamma \|z^*|_{\mathcal{V}}\|_{\mathcal{V}^*},$$

then for $y \in \mathcal{Y}$ there exists $v \in \mathcal{V}$ such that v interpolates y on Z^*; that is, $z^*(y) = z^*(v)$ for all $z^* \in Z^*$. In addition, v approximates y in the sense that $\|y-v\|_{\mathcal{Y}} \le (1+2\gamma)\mathrm{dist}_{\mathcal{Y}}(y,\mathcal{V})$.

Now, let us prove Theorem 1. At first, we prove that there exists a continuous function g satisfying the following properties:

(i) $\|g\| = 1$, (ii) $z^*(g) = \|z^*\|$, and (iii) $\omega(g,\varepsilon) \le 2\varepsilon/q_X$.

The details of proof will be given later. If g satisfies (i), (ii) and (iii), then from (6) and (iii) it follows that there exists $N^g_{\sigma,L} \in \Phi_{\sigma,L}$ such that

$$\|N^g_{\sigma,L} - g\| \le (\|\sigma\|+1)\omega\left(g, \frac{1}{L}\right) \le \frac{2(\|\sigma\|+1)}{q_X L}.$$

For every $\beta > 1$ we assume that $L = \lceil (2(\beta+1)(1+\|\sigma\|))/((\beta-1)q_X) \rceil$. Then we have

$$\|N^g_{\sigma,L} - g\| \le \frac{\beta-1}{\beta+1}.$$

Furthermore, by (i) we also have

$$\|N^g_{\sigma,L}\| \le \frac{\beta-1}{\beta+1} + 1 = \frac{2\beta}{\beta+1}.$$

Without loss of generality, suppose that $z^* \in Z^*$ satisfies $\|z^*\| = 1$. Since z^* is a linear operator and (ii) holds, there holds

$$1 = \|z^*\| = z^*(g) = z^*(g - N^g_{\sigma,L}) + z^*(N^g_{\sigma,L}).$$

Hence, from

$$\|z^*(g - N^g_{\sigma,L})\| \le \|z^*\| \|g - N^g_{\sigma,L}\| \le \frac{\beta-1}{\beta+1},$$

we have

$$z^*(N^g_{\sigma,L}) \ge 1 - |z^*(g - N^g_{\sigma,L})| \ge 1 - \frac{\beta-1}{\beta+1} = \frac{2}{\beta+1}.$$

Consequently,

$$\|z^*\| = 1 \le \frac{\beta+1}{2} z^*(N^g_{\sigma,L}) \le \frac{\beta+1}{2} \|z^*|_{\Phi_{\sigma,L}}\| \|N^g_{\sigma,L}\| \le \frac{\beta+1}{2} \cdot \frac{2\beta}{\beta+1} \| \|z^*|_{\Phi_{\sigma,L}}\| = \beta \|z^*|_{\Phi_{\sigma,L}}\|.$$

We apply Lemma 1 to the case in which the underlying space is the space $\mathcal{Y} = C[0,1]$. Then we let $Z^* = \mathrm{span}\{\delta_{x_i}\}_{i=0}^n$, where δ_{x_i} is the point evaluations operator. Finally, we take $\mathcal{V} = \Phi_{\sigma,L}$. Thus by setting $\beta = 2$, from Lemma 1 we deduce that there exists an SLFN, $N_{\sigma,L} \in \Phi_{\sigma,L}$, satisfying (3) and (5).

The only thing reminder is to construct the continuous function g such that g satisfies (i), (ii) and (iii). In fact, for any $z^* = \sum_{j=0}^n c_j \delta_{x_j} \in Z^*$, let

$$g(x) = \sum_{j=0}^n \mathrm{sgn}(c_j)\left(1 - \frac{(x-x_j)^2}{q_X^2}\right)_+, \tag{8}$$

where $(t)_+ = \max\{t,0\}$, and $\mathrm{sgn}(t)$ is the symbol function satisfying $\mathrm{sgn}(t) = 1$ for $t \ge 0$ and sgn $(t) = 0$ for $t < 0$.

From the definition, the continuity of g is obvious. For (8) we also have if $|x-x_j| \ge q_X$, $j = 0,1,\ldots,n$ then $g(x) = 0$. This means $g(x) = 0$ unless $|x-x_j| < q_X$ for some $x_j \in X$. Moreover, on $|x-x_j| \le q_X$, we have

$$g(x) = \mathrm{sgn}(c_j)\left(1 - \frac{(x-x_j)^2}{q_X^2}\right).$$

So

$$|g(x)| = 1 - \frac{(x-x_j)^2}{q_X^2} \le |g(x_j)| = 1.$$

All above yield $|g(x)| \le 1$ for all $x \in [0,1]$. Since $|g(x_j)| = 1$ for all $x_j \in X$, we get $\|g\| = 1$, which means g satisfies (i).

By using the definition of z^* again, we obtain

$$z^*(g) = \sum_{j=0}^n c_j \delta_{x_j} g(x) = \sum_{j=0}^n c_j g(x_j) = \sum_{j=0}^n c_j \mathrm{sgn}(c_j) = \sum_{j=0}^n |c_j| = \|z^*\|.$$

Thus (ii) holds.

Now, we prove that g satisfies (iii). We first deduce the following estimate for $g'(x)$

$$|g'(x)| \le \frac{2}{q_X}. \tag{9}$$

Indeed, if $|x-x_j| > q_X$ for all $x_j \in X$, then $g(x) = 0$ in a neighborhood of x, and $g'(x) = 0$. If there exists an x_j such that $|x-x_j| \le q_X$ (by the definition of g, we know that there is at most one x_j satisfying the above condition), then

$$g(x)=\mathrm{sgn}(c_j)\left(1-\frac{(x-x_j)^2}{q_X^2}\right),\quad |x-x_j|\le q_X.$$

Thus

$$g'(x)=-\mathrm{sgn}(c_j)\frac{2}{q_X}\frac{x-x_j}{q_X},\quad |x-x_j|\le q_X.$$

So

$$|g'(x)|\le\frac{2}{q_X}.$$

This finishes the proof of (9).

Because of

$$|g(x+\varepsilon)-g(x)|=\left|\int_x^{x+\varepsilon}g'(t)\,dt\right|\le\int_x^{x+\varepsilon}\frac{2}{q_X}dt=\frac{2\varepsilon}{q_X},$$

we know from the definition of $\omega(f,t)$ that $\omega(g,\varepsilon)\le 2\varepsilon/q_X$, which yields that g satisfies (iii).

This complete the proof of Theorem 1.

References

[1] F.L. Cao, Y.Q. Zhang, Z.R. He, Interpolation and rates of convergence for a class of neural networks, Appl. Math. Model 33 (2009) 1441–1456.
[2] D.B. Chen, Degree of approximation by superpositions of a sigmoidal function, Approx. Theory Appl. 9 (1993) 17–28.
[3] T.P. Chen, H. Chen, Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical system, IEEE Trans. Neural Network 6 (1995) 911–917.
[4] C.K. Chui, X. Li, Approximation by ridge functions and neural networks with one hidden layer, J. Approx. Theory 70 (1992) 131–141.
[5] G. Cybenko, Approximation by superpositions of sigmoidal function, Math. Control Signals Syst. 2 (1989) 303–314.
[6] K.I. Funahashi, On the approximate realization of continuous mapping by neural networks, Neural Networks 2 (1989) 183–192.
[7] G.G. Lorentz, Approximation of Functions, Rinehart and Winston, New York, 1966.
[8] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximation, Neural Networks 2 (1989) 359–366.
[9] G.B. Huang, H.A. Babri, General approximation theorem on feedforward networks, in: 1997 IEEE International Conference Neural Networks (ICNN97), Houston, TX, 1997.
[10] G.B. Huang, H.A. Babri, Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions, IEEE Trans. Neural Network 9 (1) (1998) 224–229.
[11] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: a new learning scheme of feedforward neural networks, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN04), Budapest, Hungary, July 25–29, 2004.
[12] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (2006) 489–501.
[13] G.B. Huang, L. Chen, C.K. Siew, Universal approximation using incremental constructive feedforward networks with random hidden nodes, IEEE Trans. Neural Networks 17 (4) (2006) 879–892.
[14] G.B. Huang, L. Chen, Convex incremental extreme learning machine, Neurocomputing 70 (2007) 3056–3062.
[15] G.B. Huang, L. Chen, Enhanced random search based incremental extreme learning machine, Neurocomputing 71 (2008) 3460–3468.
[16] Y. Ito, Independence of unscaled basis functions and finite mappings by neural networks, Math. Sci. 26 (2001) 117–126.
[17] Y. Ito, K. Saito, Superposition of linearly independent functions and finite mappings by neural networks, Math. Sci. 21 (1996) 27–33.
[18] M. Leshno, V.Y. Lin, A. Pinks, S. Schocken, Multilayer feedforward networks with a nonpolynomial activation function can approximant any function, Neural Networks 6 (1993) 861–867.
[19] B. Llanas, F.J. Sainz, Constructive approximate interpolation by nerual networks, J. Comput. Appl. Math. 188 (2006) 283–308.
[20] B. Llanas, S. Lantarón, Hermite interpolation by neural networks, Appl. Math. Comput. 191 (2007) 429–439.
[21] F.J. Narcowich, J.D. Ward, Scattered-data interpolation on R^n: error estimates for radial basis and band-limited functions, SIAM J. Math. Anal. 36 (2004) 284–300.
[22] A. Pinkus, Approximation theory of the MLP model in neural networks, Acta Numer. (1999) 143–195.
[23] E.D. Sontag, Feedforward nets for interpolation and classification, J. Comput. Syst. Sci. 45 (1992) 20–48.
[24] T.F. Xie, F.L. Cao, The ridge function representation of polynomials and an application to neural networks, Acta Math. Sinica, English Series, in press, doi:10.1007/s10114-011-9407-1.

Feilong Cao, male, was born in Zhejiang Province, China, on August, 1965. He received the B. S. degree in Mathematics and the M.S. degree in Applied Mathematics from Ningxia University, China in 1987 and 1998, respectively. In 2003, he received the Ph.D. degree in Institute for Information and System Science, Xi'an Jiaotong University, China. From 1987 to 1992, he was an Assistant Professor. During 1992 to 2002, he was an Associate Professor. He now is a Professor in China Jiliang University. His current research interests include neural networks and approximation theory. He is the author or coauthor of more than 100 scientific papers.

Shaobo Lin received the BS degree in Mathematics in 2006 and the MS degree in Basic Mathematics in 2009 from Hangzhou Normal University, China. Now he is a doctor student in College of Science of Xi'an Jiaotong University, China. His main interest is approximation theory, and spherical harmonics analysis.

Zongben Xu, received the MS degree in mathematics in 1981 and the PhD degree in applied mathematics in 1987 from Xi'an Jiaotong University, China. In 1998, he was a postdoctoral researcher in the Department of Mathematics, The University of Strathclyde, United Kingdom. He worked as a research fellow in the Information Engineering Department from February 1992 to March 1994, the Center for Environmental Studies from April 1995 to August 1995, and the Mechanical Engineering and Automation Department from September 1996 to October 1996, at The Chinese University of Hong Kong. From January 1995 to April 1995, he was a research fellow in the Department of Computing in The Hong Kong Polytechnic University. He has been with the Faculty of Science and Institute for Information and System Sciences at Xi'an Jiaotong University since 1982, where he was promoted to associate professor in 1987 and full professor in 1991, and now serves as an authorized PhD supervisor in mathematics and computer science, vice president of Xi'an Jiaotong University, and director of the Institute for Information and System Sciences.

Professor Xu has published 116 academic papers. His current interest is in computational Intelligence (particularly, neural networks, evolutionary computation, data mining theory and algorithms), Nonlinear Functional Analysis and Geometry of Banach Spaces.