Research Article

The Estimate for Approximation Error of Neural Network with Two Weights

Fanzi Zeng and Yuting Tang

Key Laboratory for Embedded and Network Computing of Hunan Province, Hunan University, Changsha 410082, China

Correspondence should be addressed to Fanzi Zeng; zengfanzhi@126.com

Received 29 October 2013; Accepted 8 December 2013

Academic Editors: J. Shu and F. Yu

Copyright © 2013 F. Zeng and Y. Tang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The neural network with two weights is constructed and its approximation ability to any continuous functions is proved. For this neural network, the activation function is not confined to the odd functions. We prove that it can limitlessly approach any continuous function from limited close subset of R^m to R^n and any continuous function, which has limit at infinite place, from limitless close subset of R^m to R^n . This extends the nonlinear approximation ability of traditional BP neural network and RBF neural network.

1. Introduction

So far, many neural networks models have been presented and applied in pattern recognition, automatic control, signal processing, and aided decision. Among these models, BP (feedforward) neural networks and RBF (radial basis function) neural networks are widely used because of their nonlinear ability of approximation to any continuous function. Up to now, these two classes of neural networks have successfully been applied to approach any nonlinear continuous function defined on bounded close subsets [1–7]. However, the approximation result to any continuous function defined on unbounded set has been very rare [8]. This inspires us to look for a new network which is of more wide approximation ability and can approach any continuous function both on a bounded set and on an unbounded set.

Wang and Zhao [9] presented a new neural network called neural networks with two weights on the base of combining the advantage of BP neural network with that of RBF neural network. This model can not only simulate BP neural network and RBF neural network, but also simulate neural networks with higher order. This neural network not only contains direction weight value with respect to BP network but also contains core weight value with respect to

RBF network. The function of its neurons is of the following form:

$$y = \phi \left[\sum_{j=1}^m \left(\frac{\omega_j (x_j - z_j)}{|\omega_j (x_j - z_j)|} \right)^s \left| \omega_j (x_j - z_j) \right|^p + \theta \right], \quad (1)$$

where y is the output of neurons, ϕ is the activation function, θ is the threshold value, ω_j is the direction weight value, z_j is the core weight value, x_j is the input, and p and s are two parameters.

In (1), when $z_j = 0$, $p = 0$, and $s = 1$, then (1) reduces to mathematical model BP network neurons; when ω_j takes a fixed value, $p = 2$ and $s = 0$, then (1) becomes mathematical model RBF network neurons.

In [10], the approximation operators with logarithmic sigmoidal function of a class of neural networks with two weights and a class of quasi-interpolation operators were investigated. Using these operators as approximation tools, the upper bounds of estimate errors were estimated.

In [2], the authors considered single hidden layer neural networks:

$$N_n(x) = c_0 + \sum_{i=1}^n c_i \phi(\omega_i x + \theta_i), \quad N_n : R \rightarrow R, \quad (2)$$

where c_i represents output weights, ϕ is the activation function, θ_i is threshold value, and ω_j is direction weight value.

The main result in [2] is as follows.

Theorem 1. Suppose ϕ is a bounded, strictly monotonously increasing, and odd function defined on R , $f \in C[a, b]$, and $n \in N$. Then there exists a feedforward neural network (BP) with one hidden layer $N_n \rightarrow R$ defined by

$$N_n(x) = c_0 + \sum_{i=1}^n c_i \phi(\omega_i x + \theta_i), \quad N_n : R \rightarrow R, \quad (3)$$

such that

$$\sup_{x \in [a, b]} |N_n(x) - f(x)| \leq \frac{5}{2} \omega \left(f, \frac{b-a}{n} \right), \quad (4)$$

where for $i = 1, 2, \dots, n$

$$\begin{aligned} x_i &= a + i \frac{b-a}{n}, & x_0 &= a, & m &= \sup_{x \in R} \phi(x), \\ c_i &= \frac{1}{2m} (f(x_i) - f(x_{i-1})), & d_n &= \phi^{-1} \left(m - \frac{m}{2n} \right), \\ c_0 &= f(a) - \sum_{i=1}^n c_i \phi(\omega_i x + \theta_i), \\ \theta_i &= -\frac{d_n n}{b-a} \left[2a + (2i-1) \frac{b-a}{n} \right], & \omega_i &= \frac{2d_n n}{b-a}. \end{aligned} \quad (5)$$

In this paper, one is concerned with the neural networks with two weights and a single hidden layer:

$$N_n(x) = \sum_{i=1}^n c_i \phi \left(l_i^s \frac{(x-z)^s}{|(x-z)|^s} |x-z|^p - \theta_i \right). \quad (6)$$

One objective is to construct a new BP neural network which is different from the BP neural network in [2] and prove this network has the approximation ability to any nonlinear continuous function. Another objective is to prove that, by adjusting the values of two parameters s and p , the neural network with two weights can not only approximate any continuous functions defined on a bounded close subset, namely, it has the same approximation ability as BP neural network in [2], but also approximate any continuous function defined on an unbounded set. Namely, the neural network with two weights has a better approximation ability than BP neural network and RBF neural network.

2. Theoretical Result

We use the following notations: the symbols N , N_0 , and R represent the set of positive integers, nonnegative integers, and real number, respectively. Using $C[a, b]$ to denote the space of continuous real-valued functions defined on $[a, b]$, it is equalled by the supremum norm $\|\cdot\|$. Let f be a real-valued function defined on $[a, b]$. $\eta(f, \delta)$, the modulus of continuity of f , is defined as $\eta(f, \delta) = \sup_{0 < h \leq \delta} \|f(\cdot + h) - f(\cdot)\|$. Then

we have $\lim_{\delta \rightarrow 0} \eta(f, \delta) = 0$, and for any real number $\lambda \geq 0$, $\eta(f, \lambda\delta) \leq (1 + \lambda)\eta(f, \delta)$.

The modulus of continuity is usually considered as the measure of the smoothness of function and the approximation error in approximation theory. The function f is called Lipschitz α ($0 < \alpha \leq 1$) continuous and is written as $f \in \text{Lip}_{C(f)} \alpha$, if there exists a constant $C(f)$ such that $\eta(f, \delta) \leq C(f)\delta^\alpha$. $C(h)$ denotes the positive constant depends only on h and its value may be different at different occurrence.

Our main result is as follows.

Theorem 2. Suppose ϕ is a bounded, strictly monotonously increasing, and odd function defined on R , $f \in C[a, b]$, and $n \in N$. Then there exists a feedforward neural network (BP) with one hidden layer $N_n \rightarrow R$ defined by

$$N_n(x) = \sum_{i=1}^n c_i \phi(\omega_i x + \theta_i), \quad (7)$$

such that

$$\sup_{x \in [a, b]} |N_n(x) - f(x)| \leq \left(1 + \frac{n+1}{4n} \right) \omega \left(f, \frac{b-a}{n} \right), \quad (8)$$

where, for $i = 1, 2, \dots, n$,

$$\begin{aligned} x_i &= a + i \frac{b-a}{n}, & d_n &= \phi^{-1} \left(\frac{m}{2n} - m \right), & x_0 &= a, \\ m &= \sup_{x \in R} |\phi(x)|, & c_i &= \frac{1}{2m} (f(x_i) - f(x_{i-1})), \\ \omega_i &= \frac{(d_n - d_n^*)n}{b-a}, \\ \theta_i &= d_n - \frac{(d_n - d_n^*)n}{b-a} \left[a + (i-1) \frac{b-a}{n} \right], \\ d_n^* &= \phi^{-1} \left(m - \frac{m}{2n} \right). \end{aligned} \quad (9)$$

Proof. Divide $[a, b]$ into n equal intervals; each has length of $(b-a)/n$ and let $a = x_0 < x_1 < \dots < x_n = b$. For $i = 1, 2, \dots, n$, we set $c_i = (1/2m)(f(x_i) - f(x_{i-1}))$, $d_n = \phi^{-1}((m/2n) - m)$, $m = \sup_{x \in R} |\phi(x)|$, $\theta_i = d_n - ((d_n - d_n^*)n/(b-a))[a + (i-1)((b-a)/n)]$, $d_n^* = \phi^{-1}(m - (m/2n))$, $\omega_i = (d_n - d_n^*)n/(b-a)$.

For any $x \in [a, b]$, given n , we assume that $x \in [x_{i-1}, x_i]$ and note that

$$N_n(x) = \sum_{i=1}^n \frac{1}{2m} (f(x_i) - f(x_{i-1})) \phi(\omega_i x + \theta_i). \quad (10)$$

Fix $x^* \in [x_{i-1}, x_i]$ and choose f , so that $N_n(x_i^*) = f(x_i^*)$. Let $G_i(x_i^*) = \phi(\omega_i x + \theta_i)$; it follows that

$$\begin{aligned} N_n(x) - f(x_i^*) &= N_n(x) - \sum_{i=1}^n \frac{1}{2m} (f(x_i) - f(x_{i-1})) G_i(x_i^*) \\ &= \sum_{i=0}^n \frac{1}{2m} (f(x_i) - f(x_{i-1})) [G_i(x) - G_i(x_i^*)]. \end{aligned} \quad (11)$$

There are two possible cases to consider: (i) $x > x_i^*$; (ii) $x \leq x_i^*$.

Case (i). When $x > x_i^*$, we have, for $i = 1, 2, \dots, n$,

$$\begin{aligned} 0 &< G_i(x) - G_i(x_i^*) \\ &< G_i(x_i) - G_i(x_i^*) \\ &< \phi\left(\omega_i\left(a + i\frac{b-a}{n}\right) + \theta_i\right) + m \\ &= \phi(d_n) + m \\ &= \frac{m}{2n}. \end{aligned} \quad (12)$$

Case (ii). When $x \leq x_i^*$, we have, for $i = 1, 2, \dots, n$,

$$\begin{aligned} 0 &> G_i(x) - G_i(x_i^*) \\ &> G_i(x_{i-1}) - m \\ &= \phi\left(\omega_i\left(a + (i-1)\frac{b-a}{n}\right) + \theta_i\right) - m \\ &= \phi(d_n^*) - m \\ &= -\frac{m}{2n}. \end{aligned} \quad (13)$$

From (12) and (13), it follows that

$$|G_i(x) - G_i(x_i^*)| \leq \frac{m}{2n}, \quad i = 1, 2, \dots, n. \quad (14)$$

Thus from (14), we have

$$\begin{aligned} N_n(x) - f(x_i^*) &\leq \sum_{i=0}^n \frac{1}{2m} |f(x_i) - f(x_{i-1})| |G_i(x) - G_i(x_i^*)| \\ &\leq \frac{n+1}{2m} \eta\left(f, \frac{b-a}{n}\right) \times \frac{m}{2n} \\ &= \frac{n+1}{4n} \eta\left(f, \frac{b-a}{n}\right). \end{aligned} \quad (15)$$

Therefore, we have

$$\begin{aligned} |N_n(x) - f(x)| &\leq |N_n(x) - f(x_i^*)| + |f(x) - f(x_i^*)| \\ &= \left(1 + \frac{n+1}{4n}\right) \eta\left(f, \frac{b-a}{n}\right). \end{aligned} \quad (16)$$

This ends the proof of Theorem 2. $\square$

Remark 3. The activation functions are assumed to be odd functions in [2], while in this paper, this assumption is removed. On the other hand, in the construction of neural networks, threshold values and direction weight values are different from those in [2].

Theorem 4. Suppose ϕ is a bounded, strictly monotonously increasing, and odd function defined on R , $f \in C[a, b]$, and

$n \in N$, $z \notin [a, b]$. Then there exists a neural network with two weights and one hidden layer $N_n : [a, b] \rightarrow R$ defined by

$$N_n(x) = \sum_{i=1}^n c_i \phi\left(l_i^s \frac{(x-z)^s}{|(x-z)|^s} |x-z|^p - \theta_i\right) \quad (17)$$

with $p < s$, $p < 0$ such that

$$\sup_{x \in [a, b]} |N_n(x) - f(x)| \leq \left(1 + \frac{n+1}{4n}\right) \eta\left(f, \frac{b-a}{n}\right), \quad (18)$$

where c_i is defined in Theorem 2 and l_i^s is equal to ω_i in Theorem 2.

Proof. We make fractional transformation $H : x \rightarrow u$, where $u = (x-z)^s/|x-z|^{p-s}$, $x \in [a, b]$. Since $z \notin [a, b]$, thus H is a continuous transformation. Hence $f([a, b]) = U$ is also a bounded close set; moreover, $0 \notin U$.

There exists an inverse transformation for fractional transformation $H^{-1} : u \rightarrow x$, where $x = z + u^{1/s}|u|^{(s-p)/sp}$, $u \in U$. So $g(u) = f(x) = f(z + u^{1/s}|u|^{(s-p)/sp})$ is a continuous real-valued function defined on bounded close set U . For $g(u)$, from Theorem 2, there exists a neural network with two weights and one hidden layer $N_n : [a, b] \rightarrow R$ defined by

$$N_n(u) = \sum_{i=1}^n c_i \phi(l_i^s u - \theta_i) \quad (19)$$

such that

$$\begin{aligned} \hat{g}(u) &= \sum_{i=1}^n c_i \phi(l_i^s u - \theta_i) \\ &= \sum_{i=1}^n c_i \phi\left(l_i^s \frac{(x-z)^s}{|(x-z)|^s} |x-z|^p - \theta_i\right) \\ &= \hat{f}(x) \end{aligned} \quad (20)$$

which satisfies

$$\sup |g(u) - \hat{g}(u)| \leq \left(1 + \frac{n+1}{4n}\right) \eta\left(f, \frac{b-a}{n}\right). \quad (21)$$

This ends the proof of Theorem 4. $\square$

Remark 5. When $p < s$, $p < 0$, the network with two weights is not BP and RBF neural network; thus, Theorem 4 validates that the neural network with two weights has the same approximation ability as BP and RBF neural network on any bounded close set. When $n \rightarrow \infty$, from Theorem 4, this neural network can approximate any continuous function defined on a bounded close set.

Theorem 6. Suppose ϕ is a bounded, strictly monotonously increasing, and odd function defined on R ; D is an unbounded subset of R (all accumulation points are in D beside ∞) and $f \in C(D)$, $z \notin D$, $\lim_{x \rightarrow \infty} f(x) = b(\text{constant})$ and then there exists a neural network with two weights and one hidden layer:

$$N_n(x) = \sum_{i=1}^n c_i \phi\left(l_i^s \frac{(x-z)^s}{|(x-z)|^s} |x-z|^p - \theta_i\right) \quad (22)$$

with $p < s, p < 0$ such that

$$\hat{f}(x) = \sum_{i=1}^n c_i \phi \left(l_i^s \frac{(x-z)^s}{|(x-z)|^s} |x-z|^p - \theta_i \right) \quad (23)$$

which satisfies

$$\sup_{x \in [a,b]} |\hat{f}(x) - f(x)| < \left(1 + \frac{n+1}{4n}\right) \eta \left(f, \frac{b-a}{n}\right). \quad (24)$$

Proof. We make fractional transformation $H: x \rightarrow u$, where $u = (x-z)^s/|x-z|^{p-s}$, $x \in D$. Since $z \notin D$, thus H is a continuous transformation on D and as $x \rightarrow \infty$, $u \rightarrow 0$. Let $H(\infty) = 0$; $H(x)$ is a continuous function on $D \cup \{\infty\}$. We prove that $H(D \cup \{\infty\}) = H(D) \cup \{0\}$ is a close set. If v is an arbitrary accumulation point of $H(D) \cup \{0\}$, then $v = 0$ or v is an accumulation point of $H(D)$. When $v = 0$, $z \in H(D) \cup \{0\}$. When v is an accumulation point of $H(D)$, then there exists a sequence $u^k \in H(D)$ and another sequence $x^k \in D$, such that $x^k \rightarrow u^k$, $u^k \rightarrow v \neq 0$ ($k \rightarrow \infty$). From fractional transformation, there exists $\bar{u}, z = \bar{u}$, $x^k \rightarrow \bar{x}$ such that $u^k \rightarrow \bar{u}$. From $u^k \rightarrow v \neq 0$ ($k \rightarrow \infty$) and inverse transformation $H^{-1}: x = z + u^{1/s}|u|^{(s-p)/sp}$, $u \in U$, it follows that $x \rightarrow \infty$ does not hold. Hence $\bar{x}$ is a finite point of D . Therefore $\bar{x} \in D$, $v = \bar{u} \in H(D)$. So, $H(D \cup \{\infty\}) = H(D) \cup \{0\}$ is a close set. Let

$$g(u) = f(x) = f\left(z + u^{1/s}|u|^{(s-p)/sp}\right). \quad (25)$$

Then $g(u)$ is a continuous function on $H(D)$. Let $g(0) = b = \lim_{x \rightarrow \infty} f(x)$; it is easy to prove that $g(u)$ is a continuous function on $H(D) \cup \{0\}$. For a continuous function $g(u)$ defined on $H(D) \cup \{0\}$, by means of Theorem 4, there exists a neural network with two weights $\sum_{i=1}^n c_i \phi(l_i^s((x-z)^s/|(x-z)|^s)|x-z|^p - \theta_i)$ such that

$$\begin{aligned} \hat{g}(u) &= \sum_{i=1}^n c_i \phi \left(\omega_i \frac{(x-z)^s}{|(x-z)|^s} |x-z|^p - \theta_i \right) \\ &= \hat{f}(x) \end{aligned} \quad (26)$$

which satisfies

$$\sup |g(u) - \hat{g}(u)| \leq \left(1 + \frac{n+1}{4n}\right) \eta \left(f, \frac{b-a}{n}\right). \quad (27)$$

This ends the proof of Theorem 6. $\square$

Remark 7. When $p < s, p < 0$, the network with two weights is not BP and RBF neural network; thus, Theorem 6 validates that the neural network with two weights has the approximation ability that BP and RBF neural network do not have on any unbounded set. Namely, the network with two weights has a better approximation ability than BP and RBF neural network on any unbounded set. When $n \rightarrow \infty$, from Theorem 6, this neural network can approximate any continuous function defined on unbounded set.

3. Conclusions

In this paper, we construct a new BP neural network and prove that the network has the approximation ability to any nonlinear continuous function. In our result, the threshold values and direction weight values in [2] are different from those in our paper. Second, we prove that neural network with two weights has the same approximation ability as BP neural network and RBF neural network to any continuous function defined on any bounded close set; furthermore, we prove that neural network with two weights has a better approximation ability than BP neural network and RBF neural network to any continuous function defined on any unbounded set.

References

- [1] J. Wang and Z. Xu, "New study on neural networks: the essential order of approximation," *Neural Networks*, vol. 23, no. 5, pp. 618–624, 2010.
- [2] F. Cao, T. Xie, and Z. Xu, "The estimate for approximation error of neural networks: a constructive approach," *Neurocomputing*, vol. 71, no. 4–6, pp. 626–630, 2008.
- [3] Z.-X. Chen, F.-L. Cao, and J.-W. Zhao, "The construction and approximation of some neural networks operators," *Applied Mathematics*, vol. 27, no. 1, pp. 69–77, 2012.
- [4] H. L. Tolman, V. M. Krasnopolsky, and D. V. Chalikov, "Neural network approximations for nonlinear interactions in wind wave spectra: direct mapping for wind seas in deep water," *Ocean Modelling*, vol. 8, no. 3, pp. 253–278, 2005.
- [5] B. Lianas and F. J. Sainz, "Constructive approximate interpolation by neural networks," *Journal of Computational and Applied Mathematics*, vol. 188, no. 2, pp. 283–308, 2006.
- [6] Z. Chen and F. Cao, "The approximation operators with sigmoidal functions," *Computers and Mathematics with Applications*, vol. 58, no. 4, pp. 758–765, 2009.
- [7] Z. Xu and F. Cao, "The essential order of approximation for neural networks," *Science in China F*, vol. 47, no. 1, pp. 97–112, 2004.
- [8] G.-W. Yang, S.-J. Wang, and Q.-X. Yan, "Research of fractional linear neural network and its ability for nonlinear approach," *Chinese Journal of Computers*, vol. 30, no. 2, pp. 189–199, 2007.
- [9] S. Wang and X. Zhao, "Biomimetic pattern recognition theory and its applications," *Chinese Journal of Electronics*, vol. 13, no. 3, pp. 373–377, 2004.
- [10] Y. Cao and X.-T. Zhao, "Data fitting based on a new double weights neural network," *Chinese Journal of Electronics*, vol. 32, no. 10, pp. 1671–1673, 2004.