Korean J. Math. **23** (2015), No. 3, pp. 327–336
http://dx.doi.org/10.11568/kjm.2015.23.3.327

CONSTRUCTIVE APPROXIMATION BY NEURAL NETWORKS WITH POSITIVE INTEGER WEIGHTS

Bum Il Hong and Nahmwoo Hahm†

Abstract. In this paper, we study a constructive approximation by neural networks with positive integer weights. Like neural networks with real weights, we show that neural networks with positive integer weights can even approximate arbitrarily well for any continuous functions on compact subsets of $\mathbb{R}$. We give a numerical result to justify our theoretical result.

1. Introduction and Preliminaries

Because of its applications in engineering such as robotics and signal processing, neural network approximation has been investigated by many mathematicians ([3], [4], [6], [7], [8]). A general form of feedforward neural network with one hidden layer is

$$\sum_{i=1}^{n} c_i \sigma(a_i x + b_i), \tag{1.1}$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is a univariate activation function and $a_i, b_i, c_i \in \mathbb{R}$ for $i = 1, 2, \cdots, n$. In (1.1), a_i's are called the weights and b_i's are called the thresholds. Although we may choose the Gaussian function

Received June 29, 2015. Revised July 30, 2015. Accepted July 31, 2015.
2010 Mathematics Subject Classification: 41A25, 41A46.
Key words and phrases: Neural network, Positive integer weight, Sigmoidal function.
† This research was supported by Incheon National University Fund, 2014.
† Corresponding author.

$\sigma(x) = e^{-x^2}$ or the generalized multiquadrics $\sigma(x) = (1+x^2)^\beta, \beta \notin \mathbb{Z}$ as an activation function, Kalman and Kwasny [5] pointed out the importance of a sigmoidal function as an activation function in hardware implementations of back propagation.

Definition 1.1. A sigmoidal function is a function $\sigma : \mathbb{R} \to \mathbb{R}$ such that

$$\lim_{x\to-\infty} \sigma(x) = 0 \quad \text{and} \quad \lim_{x\to\infty} \sigma(x) = 1. \tag{1.2}$$

The followings are examples of sigmoidal functions.

$$\text{The squashing function } : \sigma(x) = 1/(1+e^{-x}).$$

$$\text{The Heaviside function} : \sigma(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

Chen [1] proved that any continuous function on a compact subset of $\mathbb{R}$ was approximated well by a neural network with a bounded sigmoidal function and a fixed weight. Medvedeva [9] showed a neural network approximation using Taylor's theorem and infinitely differentiable activation function. In [2], Chui and Li suggested that a neural network with a continuous sigmoidal function with integer weights could approximate any continuous function on a compact subset of $\mathbb{R}$, but their proofs are not constructive. Unlike the previous papers, we suggest a constructive approximation by a neural network with a continuous monotone sigmoidal function and positive integer weights using convolution in this paper.

Note that a sigmoidal function σ is not improper integrable on $\mathbb{R}$, since $\lim_{x\to\infty} \sigma(x) = 1$. In order to obtain an improper integrable function on $\mathbb{R}$ from a sigmoidal function, we define

$$\phi_\alpha(x) = \sigma(x+\alpha) - \sigma(x) \tag{1.3}$$

for $\alpha > 0$ and a continuous monotone sigmoidal function σ. Then

$$\phi_\alpha(x) \geq 0 \tag{1.4}$$

for all $x \in \mathbb{R}$ and $\alpha > 0$.

Lemma 1.2. *If σ is a continuous monotone sigmoidal function, ϕ_α in* (1.3) is improper integrable on $\mathbb{R}$ for $\alpha > 0$.

Proof. Since $\lim_{x\to-\infty}\sigma(x)=0$ and $\lim_{x\to\infty}\sigma(x)=1$, σ is bounded and uniformly continuous on $\mathbb{R}$. Thus there exists $M>0$ such that $|\sigma(x)|\leq M$ for all $x\in\mathbb{R}$. For $p\in\mathbb{R}$, we have

$$\begin{aligned}\int_p^\infty \phi_\alpha(x)dx &= \lim_{t\to\infty}\int_p^t \phi_\alpha(x)dx\\ &= \lim_{t\to\infty}\int_p^t (\sigma(x+\alpha)-\sigma(x))dx\\ &= \lim_{t\to\infty}(\int_{p+\alpha}^{t+\alpha}\sigma(x)dx-\int_p^t\sigma(x)dx)\\ &\leq \lim_{t\to\infty}(\int_t^{t+\alpha}Mdx+\int_p^{p+\alpha}Mdx)\\ &< \infty.\end{aligned} \tag{1.5}$$

Similarly, we get, for $p\in\mathbb{R}$,

$$\int_{-\infty}^p \phi_\alpha(x)dx<\infty. \tag{1.6}$$

Therefore we complete the proof. □

2. Main results

For $\alpha>0$ and a continuous monotone sigmoidal function σ, ϕ_α in (1.3) is uniformly continuous on $\mathbb{R}$. In addition,

$$\lim_{x\to-\infty}\phi_\alpha(x)=0 \quad\text{and}\quad \lim_{x\to\infty}\phi_\alpha(x)=0. \tag{2.1}$$

Since ϕ_α is improper integrable on $\mathbb{R}$ by Lemma 1.2, we set c_α be the constant such that

$$\int_{\mathbb{R}} c_\alpha\phi_\alpha(x)dx=1. \tag{2.2}$$

For $n\in\mathbb{N}$, we define the dilation function

$$\psi_{n,\alpha}(x)=nc_\alpha\phi_\alpha(nx) \tag{2.3}$$

for $x\in\mathbb{R}$. Note that

$$\int_{\mathbb{R}}\psi_{n,\alpha}(x)dx=\int_{\mathbb{R}}nc_\alpha\phi_\alpha(nx)dx=\int_{\mathbb{R}}c_\alpha\phi_\alpha(x)dx=1 \tag{2.4}$$

by substitution.

THEOREM 2.1. *Let $\alpha > 0$ and let σ be a continuous monotone sigmoidal function on $\mathbb{R}$. For a given $\delta > 0$ and $\psi_{n,\alpha}$ in* (2.3), we have

$$\lim_{n\to\infty}\int_{\{x\in\mathbb{R}:|x|\geq\delta\}}\psi_{n,\alpha}(x)dx = 0. \tag{2.5}$$

Proof. Since $\int_{\mathbb{R}} c_\alpha\phi_\alpha(x)dx = 1 < \infty$, we get

$$\lim_{k\to\infty}\int_{\{x\in\mathbb{R}:|x|\geq k\}}c_\alpha\phi_\alpha(x)dx = 0. \tag{2.6}$$

By substitution and (2.6), we have

$$\begin{aligned}\lim_{n\to\infty}\int_{\{x\in\mathbb{R}:|x|\geq\delta\}}\psi_{n,\alpha}(x)dx &= \lim_{n\to\infty}\int_{\{x\in\mathbb{R}:|x|\geq\delta\}}nc_\alpha\phi_{n,\alpha}(nx)dx\\ &= \lim_{n\to\infty}\int_{\{x\in\mathbb{R}:|x|\geq n\delta\}}c_\alpha\phi_\alpha(x)dx\\ &= 0.\end{aligned} \tag{2.7}$$

Therefore we complete the proof. □

Using Theorem 2.1, we obtain the following.

THEOREM 2.2. *Let g be a continuous function on $\mathbb{R}$ with compact support. For a given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $n \geq N$, we have*

$$||\psi_{n,\alpha} * g - g||_{\infty,\mathbb{R}} < \epsilon, \tag{2.8}$$

where $\alpha > 0$ and $\psi_{n,\alpha}$ is a function satisfying (2.3) for a continuous monotone sigmoidal function σ.

Proof. Since g is continuous on $\mathbb{R}$ with compact support, g is uniformly continuous on $\mathbb{R}$. Hence, for a given $\epsilon > 0$, there exists $\delta > 0$ such that for any $x, y \in \mathbb{R}$ with $|x - y| < \delta$, we have

$$|g(x) - g(y)| < \frac{\epsilon}{2}. \tag{2.9}$$

By (1.4) and (2.4), for $x \in \mathbb{R}$, we get

$$\begin{aligned} |(\psi_{n,\alpha} * g)(x) - g(x)| &\leq \int_{\mathbb{R}} \psi_{n,\alpha}(y)|g(x-y) - g(x)|dy \\ &= \int_{\mathbb{R}-\{y\in\mathbb{R}:|y|<\delta\}} \psi_{n,\alpha}(y)|g(x-y) - g(x)|dy \\ &\quad + \int_{\{y\in\mathbb{R}:|y|<\delta\}} \psi_{n,\alpha}(y)|g(x-y) - g(x)|dy. \end{aligned} \tag{2.10}$$

Hence by Theorem 2.1, there exists $N \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ with $n \geq N$, we have

$$\int_{\{y\in\mathbb{R}:|y|\geq\delta\}} \psi_{n,\alpha}(y)dy < \frac{\epsilon}{4\|g\|_{\mathbb{R},\infty}}. \tag{2.11}$$

Thus, for any $n \in \mathbb{N}$ with $n \geq N$, we get

$$\begin{aligned} &\int_{\mathbb{R}-\{y\in\mathbb{R}:|y|<\delta\}} \psi_{n,\alpha}(y)|g(x-y) - g(x)|dy \\ &= \int_{\{y\in\mathbb{R}:|y|\geq\delta\}} \psi_{n,\alpha}(y)|g(x-y) - g(x)|dy \\ &\leq 2\|g\|_{\mathbb{R},\infty} \int_{\{y\in\mathbb{R}:|y|\geq\delta\}} \psi_{n,\alpha}(y)dy \\ &< \frac{\epsilon}{2} \end{aligned} \tag{2.12}$$

by (2.11). On the other hand, we also get

$$\begin{aligned} &\int_{\{y\in\mathbb{R}:|y|<\delta\}} \psi_{n,\alpha}(y)|g(x-y) - g(x)|dy \\ &< \frac{\epsilon}{2} \int_{\{y\in\mathbb{R}:|y|<\delta\}} \psi_{n,\alpha}(y)dy \\ &\leq \frac{\epsilon}{2} \int_{\mathbb{R}} \psi_{n,\alpha}(y)dy \\ &= \frac{\epsilon}{2} \end{aligned} \tag{2.13}$$

by (2.4) and (2.9). From (2.12) and (2.13), we finally have

$$|(\psi_{n,\alpha} * g)(x) - g(x)| < \epsilon \tag{2.14}$$

for all $x \in \mathbb{R}$. Therefore we complete the proof. □

For a continuous monotone sigmoidal function σ, we set

$$\Phi_{\sigma,n} = \{\sum_{i=1}^{n} c_i\sigma(a_i x + b_i) : a_i \in \mathbb{N}, b_i, c_i \in \mathbb{R}\}. \tag{2.15}$$

Next theorem shows that any continuous function on $[a, b]$ can be approximated arbitrarily well by a neural network with a continuous monotone sigmoidal function and positive integer weights.

Theorem 2.3. *Let f be continuous on $[a, b]$ and let σ be a continuous monotone sigmoidal function on $\mathbb{R}$. For a given $\epsilon > 0$, there exists $h_{m,n,\alpha} \in \Phi_{\sigma,2n}$ such that*

$$\|f - h_{m,n,\alpha}\|_{\infty,[a,b]} < \epsilon, \tag{2.16}$$

where $m, n \in \mathbb{N}$ and $\alpha > 0$.

Proof. Let

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } \ x \in [a, b] \\ \text{linear} & \text{if } \ x \in [a-1, a] \cup [b, b+1] \\ 0 & \text{otherwise.} \end{cases} \tag{2.17}$$

Then $\tilde{f}$ is a continuous extension of f on $\mathbb{R}$. For $n \in \mathbb{N}$ and $\alpha > 0$, we define $\psi_{n,\alpha}(x) = nc_\alpha(\sigma(nx + \alpha) - \sigma(nx))$ on $\mathbb{R}$, where c_α is a constant such that $\int_{\mathbb{R}} \psi_{n,\alpha}(x)dx = 1$. Since $\tilde{f}$ is continuous on $\mathbb{R}$ with compact support $[a-1, b+1]$, there exists $N \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ with $n \geq N$,

$$\|\tilde{f} - \psi_{n,\alpha} * \tilde{f}\|_{\infty,[a-1,b+1]} = \|\tilde{f} - \psi_{n,\alpha} * \tilde{f}\|_{\infty,\mathbb{R}} < \frac{\epsilon}{2} \tag{2.18}$$

by Theorem 2.1. Hence we get

$$\|f - \psi_{n,\alpha} * \tilde{f}\|_{\infty,[a,b]} \leq \|\tilde{f} - \psi_{n,\alpha} * \tilde{f}\|_{\infty,[a-1,b+1]} < \frac{\epsilon}{2}. \tag{2.19}$$

For $m \in \mathbb{N}$, we set $y_i = (a-1) + \frac{b-a+2}{m} i$ for $i = 1, 2, \cdots, m$. Then

$$\begin{aligned} & \frac{b-a+2}{m} \sum_{i=1}^{m} \psi_{n,\alpha}(x - y_i)\tilde{f}(y_i) \\ = & \frac{b-a+2}{m} \sum_{i=1}^{m} n\tilde{f}(y_i)c_\alpha(\sigma(nx - ny_i + \alpha) - \sigma(nx - ny_i)) \end{aligned} \tag{2.20}$$

is a Riemann sum for $\int_{a-1}^{b+1} \psi_{n,\alpha}(x-y)\tilde{f}(y)dy$. Thus there exists $m_0 \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ with $n \geq N$,

(2.21)

$$\|\frac{b-a+2}{m_0}\sum_{i=1}^{m_0} n\tilde{f}(y_i)c_\alpha(\sigma(n\cdot - ny_i+\alpha)-\sigma(n\cdot - ny_i))-\psi_{n,\alpha}*\tilde{f}\|_{\infty,[a,b]} < \frac{\epsilon}{2}.$$

Let

(2.22)

$$h_{m_0,n,\alpha}(x) := \frac{b-a+2}{m_0}\sum_{i=1}^{m_0} n\tilde{f}(y_i)c_\alpha(\sigma(nx-ny_i+\alpha)-\sigma(nx-ny_i))$$

for $x \in \mathbb{R}$ and $n \in \mathbb{N}$ with $n \geq N$. Then $h_{m_0,n,\alpha} \in \Phi_{\sigma,2n}$ and

$$\begin{aligned} & \|f-h_{m_0,n,\alpha}\|_{\infty,[a,b]} \\ \leq \; & \|f-\psi_{n,\alpha}*\tilde{f}\|_{\infty,[a,b]} + \|\psi_{n,\alpha}*\tilde{f}-h_{m_0,n,\alpha}\|_{\infty,[a,b]} \qquad (2.23) \\ < \; & \epsilon \end{aligned}$$

by (2.19) and (2.21). Therefore we complete the proof. □

3. Numerical results

In order to justify our theory, we suggest numerical results implemented by MATHEMATICA. We choose, for $x \in [-1,1]$,

$$f(x) = x^2 + 2x + 2 \qquad (3.1)$$

as a target function. We extend f on $\mathbb{R}$ as follows :

$$\tilde{f}(x) = \begin{cases} 0 & \text{if } x \leq -2 \\ x+2 & \text{if } -2 \leq x \leq -1 \\ x^2+2x+2 & \text{if } -1 \leq x \leq 1 \\ -5x+10 & \text{if } 1 \leq x \leq 2 \\ 0 & \text{if } 2 \leq x \end{cases} \qquad (3.2)$$

We choose the squashing function $\sigma(x) = 1/(1+e^{-x})$ as an activation function of a neural network, since it is continuous monotone sigmoidal function on $\mathbb{R}$. We also choose $\alpha = 1 > 0$. Then $c_\alpha = 1$ in (2.2) since

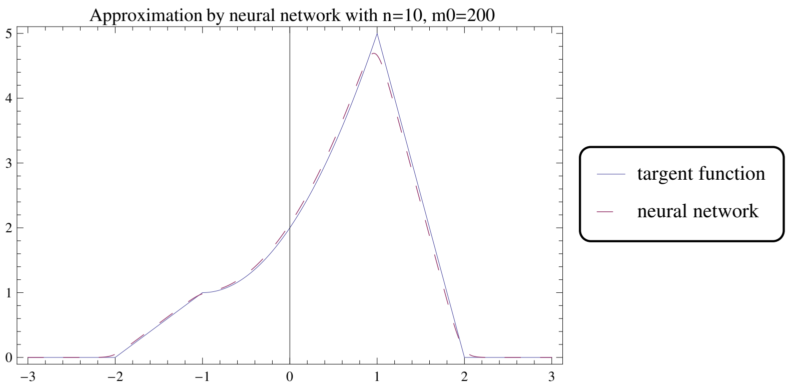

FIGURE 1. neural network with $n = 10, m_0 = 200$

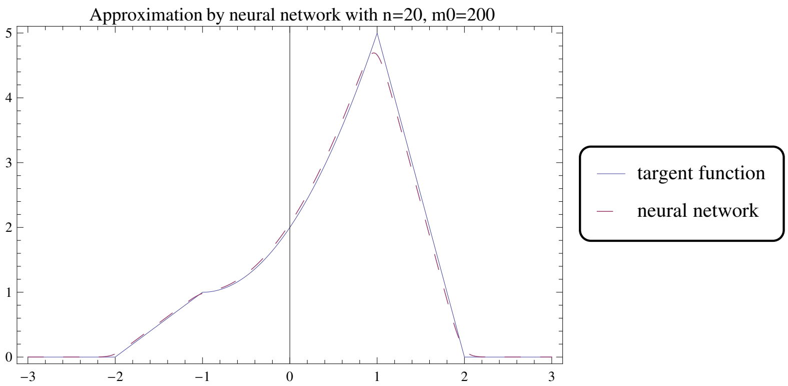

FIGURE 2. neural network with $n = 20, m_0 = 200$

$\int_{\mathbb{R}}(\sigma(x+1) - \sigma(x))dx = 1$. Thus the dilation function $\psi_{n,\alpha}$ in (2.3) is given by

$$\psi_{n,1}(x) = n\phi_1(nx) = n(\sigma(nx+1) - \sigma(nx)). \tag{3.3}$$

A neural network with the squashing activation function and positive integer weights in (2.16) of Theorem 2.3 is

$$h_{m_0,n,1} = \sum_{i=1}^{m_0} \frac{4}{m_0}\tilde{f}(-2+\frac{4i}{m_0})(n\phi_1(n(x-(-2+\frac{4i}{m_0})))). \tag{3.4}$$

We select $m_0 = 200$. Figure 1, Figure 2, Figure 3 and Figure 4 show the target function and neural networks with positive integer weights $n = 10$, $n = 20$, $n = 40$ and $n = 80$, respectively. As seen in figures,

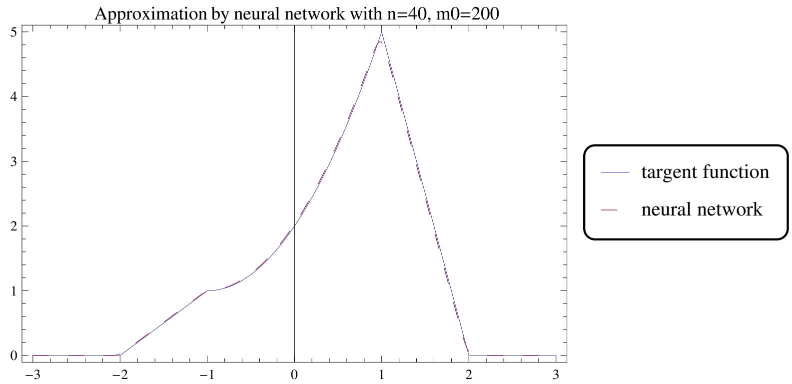

FIGURE 3. neural network with $n = 40, m_0 = 200$

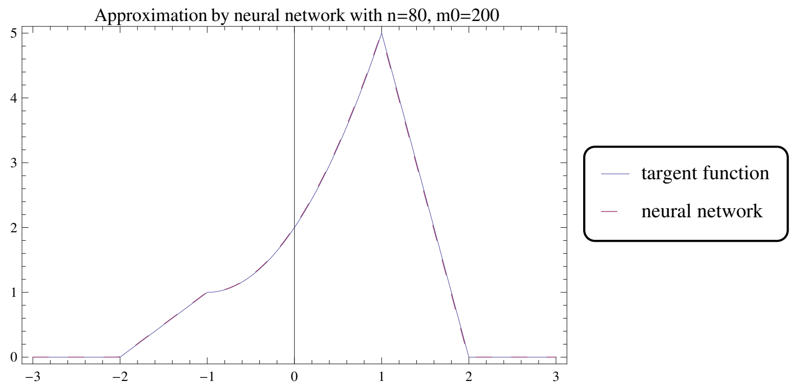

FIGURE 4. neural network with $n = 80, m_0 = 200$

neural networks approximate $\tilde{f}$ very well on $[-3, 3]$ as n increases. Since the target function f is the restriction of $\tilde{f}$ on $[-1, 1]$, it is trivial that neural networks can approximate the target function f on $[-1, 1]$.

The following numerical computation table shows that the maximum errors between the target function f and and a neural network decrease as n increases for fixed $m_0 = 200$. Therefore these results imply that the maximum errors between the target function and neural networks with integer weight decrease as the values of positive integer weight increase just like neural networks with real weights.

n	Maximum Error
10	0.67373
20	0.38583
40	0.16786
80	0.07982

References

[1] D. Chen, *Degree of Approximation by Superpositions of a Sigmoidal Function*, Approx. Theory and Appl., **9** (3) (1993), 17–28.

[2] C. K. Chui and X. Li, *Approximation by Ridge Functions and Neural Networks with One Hidden Layer*, J. Approx. Theory, **70** (1992), 17–28.

[3] N. Hahm and B. I. Hong, *Degree of Approximation by Periodic Neural Networks*, Korean J. Math., **22** (2) (2014), 307–315.

[4] N. Hahm and B. I. Hong, *The Capability of Localized Neural Network Approximation*, Honam Math. J., **34** (3) (2012), 341–349.

[5] B. L. Kalman and S. C. Kwasny, *Why Tanh : Choosing a Sigmoidal Function*, Lect. Notes Comput. Sc., **600** (1995), 303–312.

[6] J. Killian and H. T. Siegelmann, *The Dynamic Universality of Sigmoidal Neural Networks*, Inform. Comput., **128** (1996), 48–56.

[7] M. Leshno, V. Y. Lin, A. Pinkus and S. Schoken, *Multilayer Feedforward Networks with a Nonpolynomial Activation Function Can Approximate Any Function*, Neural Networks, **7** (1995), 861–867.

[8] G. Lewicki and G. Marino, *Approximation of Functions of Finite Variation by Superpositions of a Sigmoidal Function*, Appl. Math. Lett. **17**, (2004), 1147–1152.

[9] M. V. Medvedeva, *On Sigmoidal Functions*, Mosc. Univ. Math. Bull., **53** (1) (1998), 16–19.

Bum Il Hong
Department of Applied Mathematics
and The Institute of Natural Sciences
Kyung Hee University
Yongin 446-701, Korea
E-mail: bihong@khu.ac.kr

Nahmwoo Hahm
Department of Mathematics
Incheon National University
Incheon 406-772, Korea
E-mail: nhahm@inu.ac.kr