Journal of Analysis and Applications
Vol. 20 (2022), No.2, pp.123-134
ISSN: 0972-5954
© SAS International Publications
URL : www.sasip.net

Degree of multivariate approximation by superposition of a Sigmoidal function*

N.W. Hahm

Abstract. Multivariate approximation by superposition of a sigmoidal function has been investigated by many authors ([2], [3], [4]). Cybenko [3] suggested a non-constructive proof of multivariate approximation and Hahm [4] showed the density result of multivariate approximation using a constructive proof.

In this paper, we examine the complexity result of multivariate approximation by superposition of a sigmoidal function and suggest an approximation order using the modulus of continuity. Our proofs are constructive.

AMS Subject Classification (2020): 65D15, 41A25

Keywords: Constructive approximation, degree of approximation, superposition, Sigmoidal function

1. Introduction

The approximation capability by neural networks have been investigated by many researchers ([1], [5], [6], [7]). A neural network with one hidden layer is given by

$$\sum_{i=1}^{n} a_i \sigma(\mathbf{b}_i \cdot \mathbf{x} + c_i),$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is an activation function and $\mathbf{b}_i, \mathbf{x} \in \mathbb{R}^m$ and $a_i, c_i \in \mathbb{R}$. In many applications such as deep learning algorithm and binary logistic

*This research was supported by Incheon National University Research Grant in 2020.

regression, a sigmoidal function is generally chosen as an activation function in neural networks.

Definition 1. A sigmoidal function is a function $\sigma : \mathbb{R} \to \mathbb{R}$ which satisfies

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

Cybenko [3] suggested a non-constructive proof of multivariate approximation by neural networks with one hidden layer. Chen [1] showed a constructive proof to a function in $C[0, 1]$ by superposition of a sigmoidal function but it only considered the univariate case.

Note that the degree of approximation is almost the same as the complexity problem in neural network approximation.

In this paper, we investigate a degree of multivariate approximation to a continuous function on $[0, 1]^2$ by superposition of a sigmoidal function.

2. Preliminaries

Throughout the paper, i and j denote nonnegative integers, and k, l, m and n denote positive integers. We put

$$x_i = \frac{i}{n} \quad \text{and} \quad y_j = \frac{j}{n}$$

for $0 \le i, j \le n$.

In addition, we set

$$\bar{x}_i = \frac{x_{i-1} + x_i}{2} \quad \text{and} \quad \bar{y}_j = \frac{y_{j-1} + y_j}{2}$$

for $1 \le i, j \le n$.

For a continuous function F on $[0, 1]^2$ and a bounded sigmoidal function σ, we define

$$||F||_\infty := ||F||_{\infty,[0,1]^2} = \sup\big\{|F(x, y)| : x, y \in [0, 1]\big\}$$

and

$$||\sigma||_\infty := ||\sigma||_{\infty,\mathbb{R}} = \sup\{|\sigma(x)| : x \in \mathbb{R}\}.$$

Note that $||\sigma||_\infty \geq 1$ and $[0,1]^2$ is divided by n^2 disjoint subregions as follows.

$$[0,1]^2 = \bigcup_{i,j=1}^{n} \big(A_i \times B_j\big),$$

where

$$A_i = [x_{i-1}, x_i) \text{ and}$$
$$B_j = [y_{j-1}, y_j)$$

for $1 \leq i, j \leq n-1$, and

$$A_n = [x_{n-1}, x_n] \text{ and}$$
$$B_n = [y_{n-1}, y_n].$$

For $(p,q), (r,s) \in \mathbb{R}^2$, we define

$$(p,q)\cdot(r,s) = pr + qs \text{ and}$$
$$||(p,q)-(r,s)||_\infty = \max\{|p-r|, |q-s|\}.$$

For $\mathbf{x} \in [0,1]^2$, we set

$$\begin{aligned}\Phi_{n,\sigma} &:= \Phi_{n,\sigma,[0,1]^2} \\ &= \Big\{ \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\sigma(\mathbf{q}_{ij}\cdot\mathbf{x} + r_{ij}) : \mathbf{q}_{ij} \in \mathbb{R}^2,\ \ p_{ij}, r_{ij} \in \mathbb{R}\Big\}.\end{aligned}$$

The degree of approximation to a continuous function F on $[0,1]^2$ by $\Phi_{n,\sigma}$ is given by

$$E_{n,\sigma}(F) := E_{n,\sigma,[0,1]^2}(F) = \inf\{||F-G||_{\infty,[0,1]^2} : G \in \Phi_{n,\sigma}\}.$$

In order to estimate a degree of approximation by superposition of a sigmoidal function, we use the modulus of continuity.

Definition 2. For a continuous function F on a compact region A and $\delta > 0$,

$$\Omega(F, \delta, A) = \sup\Big\{|F(p,q) - F(r,s)| : (p,q), (r,s) \in A \text{ with } ||(p,q) - (r,s)||_\infty < \delta\Big\}$$

is called the modulus of continuity of F.

If $A = [0,1]^2$, we replace $\Omega(F, \delta, [0,1]^2)$ with $\Omega(F, \delta)$ for simplicity.

Note that Ω is a positive, continuous and increasing function.

3. Main results

First of all, we estimate an upper approximation bound of multivariate approximation to a continuous function F with compact support on $[0,1]^2$ by superposition of a sigmoidal function. Since $\mathrm{supp}(F) \subset [0,1]^2$, we have

$$F(x, y_0) = 0 = F(x, y_n)$$

and

$$F(x_0, y) = 0 = F(x_n, y)$$

for any $x, y \in [0,1]$.

Theorem 3. *Let F be a continuous function with compact support on $[0,1]^2$ and let σ be a bounded sigmoidal function. For a positive integer n, we have*

$$E_{n,\sigma}(F) \leq 2||\sigma||_\infty \Omega(F, \frac{1}{n}).$$

Proof. Let $\epsilon > 0$ be given and let n be fixed. We set

$$\alpha = \frac{\epsilon}{n^2(\Omega(F, 1/n) + 1)}.$$

Since σ is a sigmoidal function, there exists $\beta > 0$ such that

$$|\sigma(x) - 1| < \alpha$$

for any $x \geq \beta$ and

$$|\sigma(x)| < \alpha$$

for any $x \leq -\beta$.

Note that there exists $\gamma \in \mathbb{R}$ such that

$$\frac{\gamma}{2n} > \beta.$$

We define $\mathbf{b}_{ij} \in \mathbb{R}^2$ by

$$\mathbf{b}_{ij} := \mathbf{b}_{ij}(x,y) = \begin{cases} (0,\gamma), & \text{if } x \in A_i \text{ and } y \in [0,1], \\ (\gamma,0), & \text{otherwise.} \end{cases}$$

Now, we define

$$G_n(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n} \big(F(x_i,y_j) - F(x_i,y_{j-1})\big)\sigma \\ \big(\mathbf{b}_{ij}\cdot\big((x,y)-(\bar{x}_i,\bar{y}_j)\big)\big)$$

so that $G_n \in \Phi_{n,\sigma}$.

Note that G_n is a superposition of a sigmoidal function σ. For positive integers k,l with $1 \leq k,l \leq n$, we set

$$M_{kl}(x,y) = F(x_k,y_{l-1}) + \big(F(x_k,y_l) - F(x_k,y_{l-1})\big)\sigma \\ \big(\mathbf{b}_{kl}\cdot\big((x,y)-(\bar{x}_k,\bar{y}_l)\big)\big).$$

If $(x,y) \in [0,1]^2$, then $(x,y) \in A_{i_0} \times B_{j_0}$ for some i_0, j_0 with $1 \leq i_0, j_0 \leq n$. So,

$$\begin{aligned} &|G_n(x,y) - M_{i_0j_0}(x,y)| \\ &\leq \Big| \sum_{i=1,i\neq i_0}^{n} \sum_{j=1}^{n} \big(F(x_i,y_j) - F(x_i,y_{j-1})\big)\sigma\big(\mathbf{b}_{ij}\cdot\big((x,y)-(\bar{x}_i,\bar{y}_j)\big)\big)\Big| \\ &+ \Big| \sum_{j=1}^{j_0-1} \big(F(x_{i_0},y_j) - F(x_{i_0},y_{j-1})\big)\Big(\sigma\big(\mathbf{b}_{i_0j}\cdot\big((x,y)-(\bar{x}_{i_0},\bar{y}_j)\big)\big) - 1\Big)\Big| \\ &+ \Big| \sum_{j=j_0+1}^{n} \big(F(x_{i_0},y_j) - F(x_{i_0},y_{j-1})\big)\sigma\big(\mathbf{b}_{i_0j}\cdot\big((x,y)-(\bar{x}_{i_0},\bar{y}_j)\big)\big)\Big|, \end{aligned} \tag{1}$$

since

$$F(x_{i_0}, y_0) = F(x_{i_0}, 0) = 0$$

and

$$F(x_{i_0}, y_{j_0-1}) = \sum_{j=1}^{j_0-1} \big(F(x_{i_0}, y_j) - F(x_{i_0}, y_{j-1})\big).$$

First, we compute an upper approximation bound of (1).

Since $x \in A_{i_0}$, we have

$$\mathbf{b}_{ij} \cdot \big((x, y) - (\bar{x}_i, \bar{y}_j)\big) = \gamma(x - \bar{x}_i) \geq \frac{\gamma}{2n} > \beta$$

for $1 \leq i \leq i_0 - 1$ and $1 \leq j \leq n$, and

$$\mathbf{b}_{ij} \cdot \big((x, y) - (\bar{x}_i, \bar{y}_j)\big) = \gamma(x - \bar{x}_i) \leq -\frac{\gamma}{2n} < -\beta$$

for $i_0 + 1 \leq i \leq n$ and $1 \leq j \leq n$.

Using the fact that

$$\begin{aligned} \sum_{j=1}^{n} \big(F(x_i, y_j) - F(x_i, y_{j-1})\big) &= F(x_i, y_n) - F(x_i, y_0) \\ &= F(x_i, 1) - F(x_i, 0) = 0 \end{aligned}$$

for any positive integer i with $1 \leq i \leq n$, we have

$$\begin{aligned} &\Big| \sum_{i=1, i \neq i_0}^{n} \sum_{j=1}^{n} \big(F(x_i, y_j) - F(x_i, y_{j-1})\big)\sigma\big(\mathbf{b}_{ij} \cdot \big((x, y) - (\bar{x}_i, \bar{y}_j)\big)\big) \Big| \\ &\leq \sum_{i=1}^{i_0-1} \sum_{j=1}^{n} \Big| \big(F(x_i, y_j) - F(x_i, y_{j-1})\big)\Big(\sigma\big(\mathbf{b}_{ij} \cdot \big((x, y) - (\bar{x}_i, \bar{y}_j)\big)\big) - 1\Big) \Big| \\ &+ \sum_{i=i_0+1}^{n} \sum_{j=1}^{n} \Big| \big(F(x_i, y_j) - F(x_i, y_{j-1})\big)\sigma\big(\mathbf{b}_{ij} \cdot \big((x, y) - (\bar{x}_i, \bar{y}_j)\big)\big) \Big| \\ &< (i_0 - 1)n\Omega\Big(F, \frac{1}{n}\Big)\alpha + (n - i_0)n\Omega\Big(F, \frac{1}{n}\Big)\alpha \\ &< 2\epsilon. \end{aligned} \tag{2}$$

Since $(x,y) \in A_{i_0} \times B_{j_0}$, we have

$$\mathbf{b}_{i_0 j} \cdot \big((x,y) - (\bar{x}_{i_0}, \bar{y}_j)\big) = \gamma(y - \bar{y}_j) \geq \frac{\gamma}{2n} > \beta$$

for $1 \leq j \leq j_0 - 1$ and

$$\mathbf{b}_{i_0 j} \cdot \big((x,y) - (\bar{x}_{i_0}, \bar{y}_j)\big) = \gamma(y - \bar{y}_j) \leq -\frac{\gamma}{2n} < -\beta$$

for $j_0 + 1 \leq j \leq n$.

Thus we have

$$\begin{aligned}
&\Big|\sum_{j=1}^{j_0-1} \big(F(x_{i_0}, y_j) - F(x_{i_0}, y_{j-1})\big)\Big(\sigma\big(\mathbf{b}_{i_0 j} \cdot \big((x,y) - (\bar{x}_{i_0}, \bar{y}_j)\big)\big) - 1\Big)\Big| \\
&\leq \sum_{j=1}^{j_0-1} \Big|\big(F(x_{i_0}, y_j) - F(x_{i_0}, y_{j-1})\big)\Big(\sigma\big(\mathbf{b}_{i_0 j} \cdot \big((x,y) - (\bar{x}_{i_0}, \bar{y}_j)\big)\big) - 1\Big)\Big| \\
&< (j_0 - 1)\Omega\bigg(F, \frac{1}{n}\bigg)\alpha \\
&< \epsilon \qquad (3)
\end{aligned}$$

and

$$\begin{aligned}
&\Big|\sum_{j=j_0+1}^{n} \big(F(x_{i_0}, y_j) - F(x_{i_0}, y_{j-1})\big)\sigma\big(\mathbf{b}_{i_0 j} \cdot \big((x,y) - (\bar{x}_{i_0}, \bar{y}_j)\big)\big)\Big| \\
&\leq \sum_{j=j_0+1}^{n} \Big|\big(F(x_{i_0}, y_j) - F(x_{i_0}, y_{j-1})\big)\sigma\big(\mathbf{b}_{i_0 j} \cdot \big((x,y) - (\bar{x}_{i_0}, \bar{y}_j)\big)\big)\Big| \\
&< (n - j_0)\Omega\bigg(F, \frac{1}{n}\bigg)\alpha \\
&< \epsilon. \qquad (4)
\end{aligned}$$

From (2), (3) and (4), we get

$$|G_n(x,y) - M_{i_0 j_0}(x,y)| < 4\epsilon. \qquad (5)$$

Second, we compute an upper approximation bound of $|M_{i_0 j_0}(x,y) -$

$F(x,y)|$ for $(x,y) \in A_{i_0} \times B_{j_0}$. Since $||\sigma||_\infty \geq 1$, we have

$$\begin{aligned}
&\big|M_{i_0 j_0}(x,y) - F(x,y)\big| \\
&\leq \big|F(x_{i_0}, y_{j_0-1}) - F(x,y)\big| \\
&+ \big|\big(F(x_{i_0}, y_{j_0}) - F(x_{i_0}, y_{j_0-1})\big)\sigma\big(\mathbf{b}_{i_0 j_0}\big((x,y) - (\bar{x}_{i_0}, \bar{y}_{j_0})\big)\big)\big| \\
&\leq \Omega\bigg(F, \frac{1}{n}\bigg) + ||\sigma||_\infty \Omega\bigg(F, \frac{1}{n}\bigg) \\
&\leq 2||\sigma||_\infty \Omega\bigg(F, \frac{1}{n}\bigg). \qquad (6)
\end{aligned}$$

By (5) and (6), we finally have, for any $(x,y) \in [0,1]^2$,

$$\begin{aligned}
&\big|G_n(x,y) - F(x,y)\big| \\
&\leq \big|G_n(x,y) - M_{i_0 j_0}(x,y)\big| + \big|M_{i_0 j_0}(x,y) - F(x,y)\big| \\
&\leq 4\epsilon + 2||\sigma||_\infty \Omega\bigg(F, \frac{1}{n}\bigg).
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we get

$$E_{n,\sigma}(F) \leq ||G_n - F||_{\infty,[0,1]^2} \leq 2||\sigma||_\infty \Omega\bigg(F, \frac{1}{n}\bigg).$$

Therefore we complete the proof. □

Note that $\Omega(F,\delta) \to 0$ as $\delta \to 0$, since F is uniformly continuous on $[0,1]^2$. So, for a given $\epsilon > 0$, there exists a sufficiently large positive integer n such that

$$\Omega\bigg(F, \frac{1}{n}\bigg) < \frac{\epsilon}{2||\sigma||_\infty}. \qquad (7)$$

Thus, the next theorem that is the main theorem of [4] can be obtained directly from Theorem 3 and (7).

Theorem 4. *Let F be a continuous function with compact support on $[0,1]^2$ and let σ be a bounded sigmoidal function. For a given $\epsilon > 0$, there exist a positive integer n and $G_n \in \Phi_{n,\sigma}$ such that*

$$||G_n - F||_{\infty,[0,1]^2} < \epsilon. \qquad □$$

If F is continuous on a compact subregion R of $\mathbb{R}^2$, there exists a continuous extension $\tilde{F}$ with compact support on $[a,b]^2$ such that $\tilde{F}(x,y) = F(x,y)$ on R with $R \subset [a,b]^2$.

Theorem 5. *Let F be continuous on a compact subregion R of $\mathbb{R}^2$ and let σ be a bounded sigmoidal function. For each positive integer n, we have*

$$E_{n,\sigma,R}(F) \leq 2||\sigma||_\infty \Omega\left(\tilde{F}, \frac{b-a}{n}, [a,b]^2\right),$$

where $\tilde{F}$ is a continuous extension of F with compact support on $[a,b]^2$.

Proof. Let n be fixed. Since $\tilde{F}$ is continuous extension of F with compact support on $[a,b]^2$, we have $\tilde{F}(x,y) = F(x,y)$ on R and $\text{supp}(\tilde{F}) \subset [a,b]^2$. By Theorem 3, there exists $\tilde{G_n} \in \Phi_{n,\sigma,[a,b]^2}$ such that

$$||\tilde{G_n} - \tilde{F}||_{\infty,[a,b]^2} \leq 2||\sigma||_\infty \Omega\left(\tilde{F}, \frac{b-a}{n}, [a,b]^2\right).$$

Hence we get

$$\begin{aligned} ||\tilde{G_n} - F||_{\infty,R} = ||\tilde{G_n} - \tilde{F}||_{\infty,R} &\leq ||\tilde{G_n} - \tilde{F}||_{\infty,[a,b]^2} \\ &\leq 2||\sigma||_\infty \Omega\left(\tilde{F}, \frac{b-a}{n}, [a,b]^2\right) \end{aligned}$$

and so

$$E_{n,\sigma,R}(F) \leq 2||\sigma||_\infty \Omega\left(\tilde{F}, \frac{b-a}{n}, [a,b]^2\right).$$

Therefore we complete the proof. □

If we use the Heaviside function as a sigmoidal function, we can obtain a lower approximation bound for a certain continuous function with compact support on $[0,1]^2$.

For a positive integer n and $1 \leq i,j \leq n$, we set $F_{ij}(x,y)$ be the quadrangular pyramid with vertices $(x_{i-1}, y_{j-1}, 0)$, $(x_{i-1}, y_j, 0)$, $(x_i, y_{j-1}, 0)$, $(x_i, y_j, 0)$ and $(\bar{x}_i, \bar{y}_j, (-1)^{i+j})$.

Define L_n on $[0,1]^2$ by

$$L_n(x,y) = F_{ij}(x,y) \quad \text{if} \ \ (x,y) \in A_i \times B_j. \tag{8}$$

Then L_n is continuous and piecewise linear on $[0,1]^2$ and $\text{supp}(L_n) = [0,1]^2$ since $L_n(0,y) = 0, L_n(1,y) = 0, L_n(x,0) = 0$ and $L_n(x,1) = 0$.

Theorem 6. *For a positive integer n and the Heaviside function H, the continuous function L_n with compact support on $[0,1]^2$ in (8) satisfies*

$$||H||_\infty \Omega\left(L_n, \frac{1}{2n}\right) \leq E_{n,H}(L_n).$$

Proof. It is clear that $\Omega(L_n, 1/2n) = 1$ and $||H||_\infty = 1$. Note that G_n in $\Phi_{n,H}$ is of the form

$$G_n(x,y) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} H(\mathbf{b}_{ij} \cdot \mathbf{x} + c_{ij}),$$

where $\mathbf{b}_{ij} \in \mathbb{R}^2$ and $\mathbf{x} \in [0,1]^2$. Then G_n is a step function with at most n^2 different values.

Hence G_n is constant on $[\bar{x}_{i_0-1}, \bar{x}_{i_0}] \times [\bar{y}_{j_0-1}, \bar{y}_{j_0}]$ for some i_0, j_0 with $1 \leq i_0, j_0 \leq n$. Then

$$1 \leq ||L_n - G_n||_{\infty,[\bar{x}_{i_0-1},\bar{x}_{i_0}]\times[\bar{y}_{j_0-1},\bar{y}_{j_0}]}.$$

Therefore, we have

$$\begin{aligned} ||H||_\infty \Omega\left(L_n, \frac{1}{2n}\right) &= 1 \leq ||L_n - G_n||_{\infty,[\bar{x}_{i_0-1},\bar{x}_{i_0}]\times[\bar{y}_{j_0-1},\bar{y}_{j_0}]} \\ &\leq ||L_n - G_n||_{\infty,[0,1]^2}. \end{aligned} \tag{9}$$

By choosing infimum on (9), we complete the proof. □

4. Conclusion

In this paper, we explored a degree of approximation to a continuous multivariate function on $[0,1]^2$ by superposition of a sigmoidal function.

Although we used superposition of a sigmoidal function, it is not a neural network with one hidden layer since the weights in superposition of a sigmoidal function vary along with $(x, y) \in [0, 1]^2$ as we mentioned in [4]. But our constructive proofs offer the motivation of a degree of constructive multivariate approximation by neural networks with one hidden layer and so we will study it in future.

Acknowledgement. The author would like to express our sincere gratitude to the referees for his valuable suggestions and comments which improved the paper.

References

[1] D. Chen, *Degree of approximation by superpositions of a sigmoidal function*, Approx. Theory Appl., 9 (1993), 17-28.

[2] D. Costarelli and R. Spigler, *Multivariate neural network operators with sigmoidal activation functions*, Neural Networks, 48 (2013), 72-77.

[3] G. Cybenko, *Approximation by superposition of a sigmoidal function*, Math. Control Signals systems, 2 (1989), 303-314.

[4] N. Hahm, Constructive multivariate approximation by superposition of a sigmoidal function, J. Anal. Appl., 18 (2020), 119-131.

[5] B. I. Hong and N. W. Hahm, *Approximation by neural networks with a generalized sigmoidal function*, J. Anal. Appl., 17 (2019), 21-34.

[6] B. I. Hong and N. Hahm, *Approximation order to a function in* $\bar{C}(\mathbb{R})$ *by superposition of a sigmoidal function*, Comput. Math. Appl., 47 (2004), 1897-1903.

[7] M. V. Medvedeva, *On sigmoidal functions*, Moscow Uni. Math. Bull., 53 (1998), 16-19.

Department of Mathematics
Incheon National University
Incheon 22012
Republic of Korea
E-mail: nhahm@inu.ac.kr

(*Received: January, 2022; Revised: February, 2022*)