Journal of Analysis and Applications
Vol. 23 (2025), No.2, pp.61-71
ISSN: 0972-5954

URL : www.sasip.net

Constructive localized approximation of multivariate functions by neural networks*

N.W. Hahm

Abstract. In this paper, we discuss a constructive localized approximation of multivariate functions by generalized translation networks. Using the idea of B-spline functions, we suggest an approximation order of multivariate functions by the tensor product of B-spline functions using the modulus of continuity. Then we show the approximation possibility of the tensor product of B-spline functions by generalized translation networks. Finally, we demonstrate our theory using numerical results.

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

Keywords: Localized approximation, neural networks, multivariate functions, B-spline function

1. Introduction

Recently, research on neural network approximation has obtained significant interest in a variety of disciplines including mathematics ([2], [3], [4], [5]). Since the number of neurons required for neural network approximation grows exponentially with the dimensionality of the domain of target functions, Chui, Li and Mhaskar [1] suggested a localized approximation to solve this problem. The explanation for the motivation behind a localized approximation is as follows. When the target functions change over a small portion of a region, a localized approximation offers the advantage

*This research was supported by Incheon National University Research Grant in 2023.

of requiring the retraining of neural networks only for that specific region, rather than for the entire region [8].

In [2], Hahm and Hong suggested a localized neural network approximation with one hidden layer but the target functions were univariate. In this paper, we locally approximate continuous multivariate target functions with compact support by generalized translation networks.

2. Preliminaries

Let s be a natural number. Note that a generalized translation network is of the form

$$\sum_{i=1}^{n} c_i \phi(\mathbf{a}_i \cdot \mathbf{x} + b_i),$$

where ϕ is a univariate activation function, $\mathbf{a}_i, \mathbf{x} \in \mathbb{R}^s$ and $b_i, c_i \in \mathbb{R}$. When F is continuous on $[a,b]^s$, the modulus of continuity of F defined by

$$\Omega(F,t) := \sup\{|F(\mathbf{x}) - F(\mathbf{y})| : ||\mathbf{x} - \mathbf{y}||_\infty \leq t,\ \mathbf{x}, \mathbf{y} \in [a,b]^s\}$$

satisfies the following for $t > 0$.

Lemma 1 [7]. *For $F \in C([a,b]^s)$ and $t > 0$, the modulus of continuity Ω of F satisfies the followings.*

1. (1)] $\Omega(F,t) \to 0$ *as* $t \to 0$.
2. (2)] $\Omega(F,\alpha t) \leq (\alpha+1)\Omega(F,t)$ *for* $\alpha > 0$.

For a nonnegative integer m, B-spline function of order m is defined recursively by

$$B_m(x) = \int_0^1 B_{m-1}(x-t)dt,$$

where

$$B_0(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Note that $B_2(x)$ is a continuously differentiable function with compact support $[0,3]$. We denote the tensor product of $B_m(x)$ and $B_m(y)$ by $B_m(x,y)$. In addition, we set $C_c([a,b]^s) := \{F \mid F \text{ is continuous on } [a,b]^s \text{ and } \mathrm{supp}(F) \subset [a,b]^s\}$.

3. Main results

First, we approximate $F \in C_c([0,1]^2)$ by a superposition of the tensor product of B-spline functions and suggest an approximation order using the modulus of continuity.

Theorem 2. *Let $F \in C_c([0,1]^2)$ and let Ω be the modulus of continuity of F. For a given integer n with $n \geq 2$ and $x, y \in [0,1]$, we set*

$$P_n(x,y) = n^2 \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left(\int_{\frac{j}{n}}^{\frac{j+1}{n}} \int_{\frac{i}{n}}^{\frac{i+1}{n}} F(s,t)dsdt \right) B_2(nx-i, ny-j). \quad (1)$$

Then we have

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

Proof. Let $\tilde{F}$ be a continuous extension of F on $[-2,2]^2$ with $\tilde{F}(x,y) = 0$ on $[-2,2]^2 - [0,1]^2$. Since $\tilde{F}$ is continuous on $[-2,2]^2$, we have

$$|\tilde{F}(\mathbf{x}) - \tilde{F}(\mathbf{y})| \leq \Omega(\tilde{F}, \frac{1}{n}) = \Omega(F, \frac{1}{n}) \quad (2)$$

for any $\mathbf{x}, \mathbf{y} \in [-2,2]^2$ with $\|\mathbf{x} - \mathbf{y}\|_\infty \leq \frac{1}{n}$.

Note that

$$\int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{j}{n}}^{\frac{j+1}{n}} \tilde{F}(s,t)dsdt = 0 \quad (3)$$

for $i, j = -3, -2, -1$ and

$$\sum_{i=-3}^{n-1} \sum_{j=-3}^{n-1} B_2(nx-i, ny-j) = 1 \quad (4)$$

by the property of B-spline function. Thus we have, by (2), (3) and (4),

$$
\begin{aligned}
& \left|F(x,y) - P_n(x,y)\right| \\
= & \left|F(x,y) - n^2 \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \left(\int_{\frac{j}{n}}^{\frac{j+1}{n}} \int_{\frac{i}{n}}^{\frac{i+1}{n}} F(s,t)dsdt\right) B_2(nx-i, ny-j)\right| \\
= & \left|F(x,y) - n^2 \sum_{i=-3}^{n-1}\sum_{j=-3}^{n-1} \left(\int_{\frac{j}{n}}^{\frac{j+1}{n}} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \tilde{F}(s,t)dsdt\right) B_2(nx-i, ny-j)\right| \\
\leq & \; n^2 \sum_{i=-3}^{n-1}\sum_{j=-3}^{n-1} \left(\int_{\frac{j}{n}}^{\frac{j+1}{n}} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left|\tilde{F}(x,y) - \tilde{F}(s,t)\right| dsdt\right) B_2(nx-i, ny-j) \\
\leq & \; \Omega(F, \frac{1}{n}) \sum_{i=-3}^{n-1}\sum_{j=-3}^{n-1} B_2(nx-i, ny-j) \\
= & \; \Omega(F, \frac{1}{n})
\end{aligned}
$$

for $x, y \in [0,1]$. Therefore we complete the proof. □

Note that $B_2(x,y)$ is a piecewise polynomial of degree 4 on $[0,3]^2$. Thus we denote $B_2(x,y)$ on $I_i \times I_j \subset [0,3]^2$ by

$$\sum_{p=0}^{2}\sum_{q=0}^{2} a_{p,q}^{i,j} x^p y^q, \tag{5}$$

where $I_i = [i, i+1)$ and $I_j = [j, j+1)$ for $0 \leq i, j \leq 2$. Now, we show that P_n in (1) is approximated locally on $[0,1]^2$ by the generalized translation network.

Theorem 3. *Let σ be an infinitely differentiable function on $\mathbb{R}$ and let $b \in \mathbb{R}$ with $\sigma^{(n)}(b) \neq 0$ for all $n \in \mathbb{N}$. We set*

$$
\begin{aligned}
& Q_{\sigma,p,q,h}(x,y) \\
& = \frac{1}{h^{p+q}\sigma^{(p+q)}(b)} \sum_{c=0}^{p}\sum_{d=0}^{q} (-1)^{p+q-(c+d)} \binom{p}{c}\binom{q}{d} \sigma(h(cx+dy)+b)
\end{aligned}
$$

for $h > 0$. For a given $\epsilon > 0$, there exists $h > 0$ such that the generalized *translation network given by*

$$N_{\sigma,n,2,h}(x,y) = n^2 \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{j}{n}}^{\frac{j+1}{n}} F(s,t)dsdt\Big)$$
$$\sum_{p=0}^{2}\sum_{q=0}^{2} a_{p,q}^{i,j} Q_{\sigma,p,q,h}(nx-i, ny-j)$$

satisfies

$$\|P_n - N_{\sigma,n,2,h}\|_{\infty,[0,1]^2} < \epsilon,$$

where $a_{p,q}^{i,j}$*'s are coefficients in (5).*

Proof. Note that for $h > 0$,

$$Q_{\sigma,p,q,h}(x,y)$$
$$= \frac{1}{h^{p+q}\sigma^{(p+q)}(b)} \sum_{c=0}^{p}\sum_{d=0}^{q}(-1)^{p+q-(i+j)}\binom{p}{c}\binom{q}{d}\sigma(h(cx+dy)+b)$$

represents a divided difference for $x^p y^q$ on $[0,1]^2$. Hence we have

$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\Big| B_2(nx-i, ny-j)$$
$$- \sum_{p=0}^{2}\sum_{q=0}^{2} a_{p,q}^{i,j} Q_{\sigma,p,q,h}(nx-i, ny-j)\Big| \leq C_{\sigma,n} h,$$

where $C_{\sigma,n}$ is a positive constant depending on σ and n.

Since F is continuous on $[0,1]^2$, there exists $M > 0$ such that $|F(x,y)| \leq M$ for all $x, y \in [0,1]$. We set

$$Q_{\sigma,n,2,i,j,h}(x,y) = \sum_{p=0}^{2}\sum_{q=0}^{2} a_{p,q}^{i,j} Q_{\sigma,p,q,h}(nx-i, ny-j). \tag{6}$$

By (6), we have

$$\begin{aligned}
&\Big|P_n(x,y) - N_{\sigma,n,2,h}(x,y)\Big| \\
\le\ & n^2 \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\left(\int_{\frac{j}{n}}^{\frac{j+1}{n}}\int_{\frac{i}{n}}^{\frac{i+1}{n}} M dsdt\right) \\
&\Big|\big(B_2(nx-i,ny-j) - Q_{\sigma,n,2,i,j,h}(x,y)\big)\Big| \\
=\ & M \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\Big|B_2(nx-i,ny-j) - Q_{\sigma,n,2,i,j,h}(x,y)\Big| \\
\le\ & MC_{\sigma,n}h \\
<\ & \epsilon
\end{aligned}$$

for sufficiently small $h > 0$. Thus we complete the proof. □

Using Theorem 2 and Theorem 3, we obtain the following result that is the main theorem of this paper.

Theorem 4. *Let σ be an infinitely differentiable function on $\mathbb{R}$ and let $b \in \mathbb{R}$ be a real number such that $\sigma^{(n)}(b) \neq 0$ for all $n \in \mathbb{N}$. For a given $\epsilon > 0$, the generalized translation network given by*

$$\begin{aligned}
&N_{\sigma,n,2,h}(x,y) \\
&= n^2 \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{j}{n}}^{\frac{j+1}{n}} F(s,t)dsdt\Big) Q_{\sigma,n,2,i,j,h}(x,y)
\end{aligned}$$

satisfies

$$\|F - N_{\sigma,n,2,h}\|_{\infty,[0,1]^2} \le \Omega(F, \frac{1}{n}),$$

where $a_{p,q}^{i,j}$'s are coefficients in (5) and $Q_{\sigma,n,2,i,j,h}$ is a function in (7).

Proof. Let $\epsilon > 0$ be given and let P_n be a function in (1). By Theorem 2 and Theorem 3, we have, for sufficiently small $h > 0$,

$$
\begin{aligned}
& \|F - N_{\sigma,n,2,h}\|_{\infty,[0,1]^2} \\
\leq \; & \|F - P_n\|_{\infty,[0,1]^2} + \|P_n - N_{\sigma,n,2,h}\|_{\infty,[0,1]^2} \\
\leq \; & \Omega(F, \frac{1}{n}) + \epsilon.
\end{aligned} \tag{7}
$$

Since $\epsilon > 0$ is arbitrary, we finish the proof. □

4. Numerical results

In order to verify our results, we provide numerical data generated by MATHEMATICA. We choose

$$
F(x,y) = \begin{cases} \sin\big(\pi(2x - \frac{1}{2})\big)\cos\big(\pi(2x-1)\big), & \text{if } \frac{1}{4} \leq x, y \leq \frac{3}{4} \\ 0, & \text{otherwise} \end{cases}
$$

as the target function.

Note that $F \in C_c([0,1]^2)$ and the graph of F on $[0,1]^2$ is as follows.

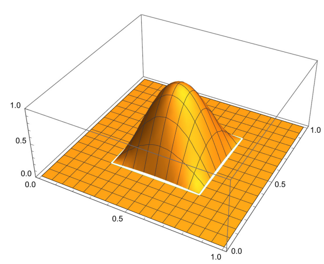

Figure 1: $F(x, y)$

Now, we approximate F by a superposition of the tensor product of B-spline functions $B_2(x)$ and $B_2(y)$, where

$$B_2(x) = \begin{cases} 0 & \text{if } x < 0 \\ \dfrac{x^2}{2} & \text{if } 0 \le x < 1 \\ \dfrac{1}{2}(-2x^2 + 6x - 3) & \text{if } 1 \le x < 2 \\ \dfrac{1}{2}(x^2 - 6x + 9) & \text{if } 2 \le x < 3 \\ 0 & \text{if } 3 \le x. \end{cases}$$

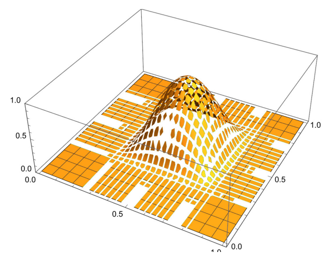

Figure 2: $P_{32}(F)$

Figure 2 demonstrates that P_{32} approximates F closely on $[0,1]^2$. In addition, numerical computation shows that the maximum errors between F and P_n decrease as n increases. Table 1 shows this.

Table 1: Maximum errors between F and P_n

Maximum error	n=4	n=8	n=16	n=32
$\|\|F - P_n\|\|_{\infty,[0,1]^2}$	0.898679	0.640863	0.350335	0.187611

Now, we choose $\sigma(x) = e^x$ as the activation function of generalized translation network.

Note that $\sigma^{(m)}(0) = 1 \neq 0$ for any $m \in \mathbb{N}$ and P_n is a superposition of the tensor product of B-spline functions. In order to avoid the round-off error, we show the approximation of $B_2(x,y)$ on $[0,3]^2$ by a generalized translation network to justify the result of Theorem 3. Let

$$B_2(x,y) = \sum_{i=0}^{2}\sum_{j=0}^{2}\Big(\sum_{p=0}^{2}\sum_{q=0}^{2} a_{p,q}^{i,j} x^p y^q\Big),$$

where $a_{p,q}^{i,j}$'s are coefficients in (5). Then Figure 4 shows that the generalized translation network

$$Net_{\sigma,2,h}(x,y) = \sum_{i=0}^{2}\sum_{j=0}^{2}\Big(\sum_{p=0}^{2}\sum_{q=0}^{2} a_{p,q}^{i,j} Q_{\sigma,p,q,h}(x,y)\Big)$$

approximate $B_2(x,y)$ on $[0,3]^2$ closely for sufficiently small $h>0$.

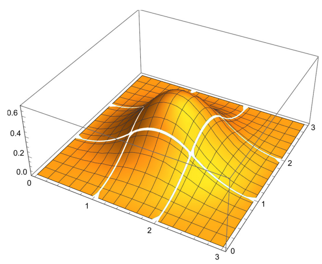

Figure 3: $B_2(x,y)$

Table 2 demonstrates that the maximum errors between B_2 and $Net_{\sigma,2,h}$ decrease as $h \to 0$.

Table 2: Maximum errors between B_2 and $Net_{\sigma,2,h}$

Maximum error	h=0.1	h=0.01	h=0.001
$\|\|B_2(\cdot,\cdot) - Net_{\sigma,2,h}\|\|_{\infty,[0,3]^2}$	0.169415	0.0175478	0.00156952

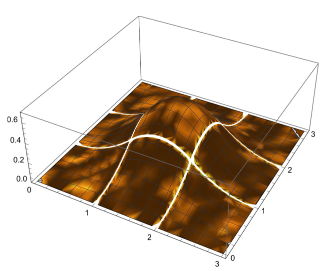

Figure 4: $Net_{\sigma,2,0.001}$

4. Conclusion

We have proved a localized neural network approximation of multivariate functions using the tensor product of B-spline functions. In this study, we expanded a localized neural network approximation to continuous multivariate functions and obtained an approximation order using the modulus of continuity. We will explore the approximation order of differentiable multivariate functions locally by neural networks in the future.

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

References

[1] C.K. Chui, X. Li and H. N. Mhaskar, *Neural networks for localized approximation*, Math. Comput., 63 (1994), 607-623.

[2] N. Hahm and B. I. Hong, *The Capability of localized neural network approximation*,Honam Math. J., (2013), 729-738.

[3] N. Hahm and B. Hong, *Simultaneous approximation algorithm using a feedforward neural network with a single hidden layer*, J. Korean Phys. Soc., 54 (2009), 2219-2224.

[4] N. Hahm, B. Hong and I. Ryoo, *Approximation algorithms using generalized translation networks with one hidden layer*, J. Korean Phys. Soc., 58 (2011), 174-181.

[5] N. Hahm, *Simultaneous approximation of multivariate functions by superposition of a sigmoidal function*, J. Anal. Appl., 21 (2023), 65-76.

[6] N. Hahm, *Degree of simultaneous approximation of multivariate functions by neural networks with one hidden layer*, J. Anal. Appl., 22 (2024), 69-79.

[7] L.L. Schumaker, Spline functions : Basic theory, John Wiley and Sons, 1981.

[8] H.N. Mhaskar, *Neural networks for localized approximation of real functions*, IEEE-SP Workshop Proc., (1993), pp.190-196.

[9] H.N. Mhaskar, *Versatile Gaussian networks*,IEEE Workshop on Nonlinear Image and Signal Proc., (1995), 70-73.

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

(*Received: April, 2025; Revised: June, 2025*)