Some density results by deep Kantorovich type neural network operator

Manju Sharma, Uaday Singh*
Manju_s@ma.iitr.ac.in,Uaday.singh@ma.iitr.ac.in

SIGMA (Signal, Image, Geometry, Modeling, Approximation) 28 October-1 November 2024
CIRM Luminy, Marseille, France

YEARS OF IIT ROORKEE Estd. 1847

Introduction

Neural networks are widely applied in numerous areas such as machine learning algorithms, artificial intelligence, solving partial differential equations etc. These applications of neural networks revolve around approximating a typically unknown function f from a given function class. One of the important problems of approximation of functions by neural network architectures is the density problem. The density problems are based on whether it is possible to approximate a function arbitrarily well by a function defined using neural network architectures.

Preliminaries and known results

A sigmoidal function η : R → R is a measurable function which satisfies

$$\lim_{x\to-\infty} \eta(x) = 0 \text{ and } \lim_{x\to\infty} \eta(x) = 1.$$

Throughout the paper, we consider η as a non-decreasing sigmoidal function with η(3) > η(1) also satisfying the following properties:

P1: η(x)−1/2 is an odd function.

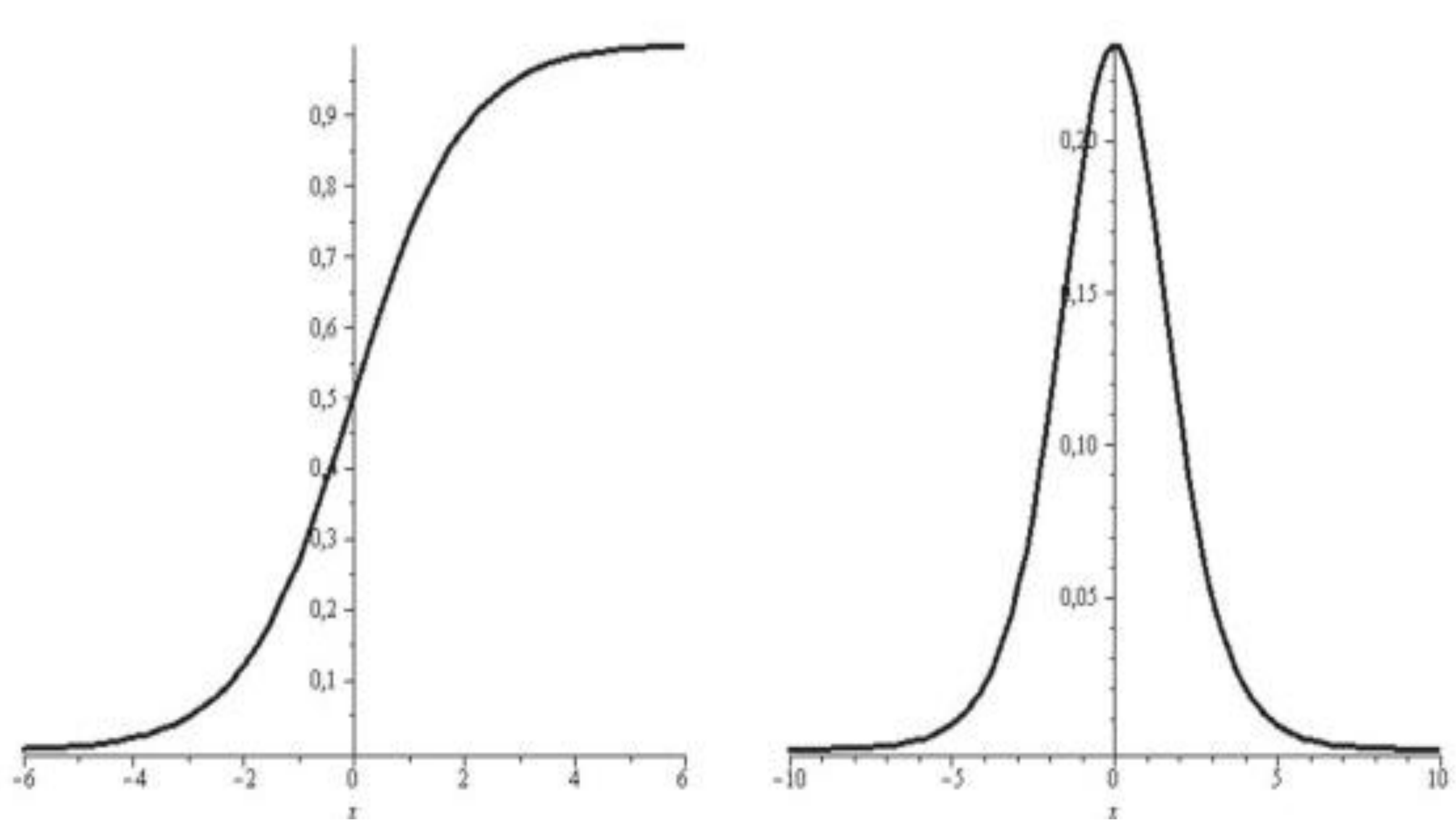
P2: $\eta \in C^2(R)$ (space of twice continuously differentiable functions) and is concave for x ≥ 0.

P3: $\eta(x) = O(|x|^{-\alpha})$ as x → −∞ for some α > 0.

For this sigmoidal function, we define the density function

$$\psi_\eta(x) := \frac{1}{2}(\eta(x+1) - \eta(x-1)),\ x \in R.$$

Example : Logistic function $\eta(x) = \frac{1}{1+e^{-x}}$ satisfies all the above conditions. The plot of logistic function with its corresponding density function are shown in the figure.



Now, we recall single-layered(means single hidden layered) Kantorovich type NNOs (K_n). Let η be a sigmoidal function as defined above. For a locally integrable function $f : I \to R$ the operator K_n is defined as

$$K_n(f;x) = \frac{\sum_{k=-n}^{n-1}\left(n\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(u)du\right)\psi_\eta(nx-k)}{\sum_{k=-n}^{n-1}\psi_\eta(nx-k)}, \quad x \in I,$$

n ∈ N. The approximation properties of the operator K_n stated as under.

Theorem 1. Let $f : I \to R$ be bounded. Then $\lim_{n\to\infty} K_n$(f, x) = f (x) at each point x ∈ I where f is continuous. Moreover, if $f \in C(I)$, then

$$\lim_{n\to\infty} \| K_n(f,.) - f(.)\|_\infty = 0.$$

Theorem 2. For every f ∈ $L^p(I)$, 1 ≤ p < ∞, we have

$$\lim_{n\to\infty} \| K_n(f,.) - f \|_p = 0.$$

In this paper, we introduce the deeper version of the Kantorovich type neural network operators (K_n) and prove the density results in the spaces C(I) and $L^p(I)$, p ≥ 1.

Approximation by two layer deep neural network operator

For a locally integrable function $f : I \to R$ and sigmoidal function satisfying the above assumptions, we define the two-layer NNOs as follows

$$(K_{(n_1,n_2)}f)(x) := \frac{\sum_{k=-n_1}^{n_1-1}\psi_\eta\left(n_1\frac{\sum_{j=-n_2}^{n_2-1}\frac{2j+1}{2n_2}\psi_\eta(n_2x-j)}{\sum_{j=-n_2}^{n_2-1}\psi_\eta(n_2x-j)} - k\right)\left(n_1\int_{\frac{k}{n_1}}^{\frac{k+1}{n_1}} f(u)du\right)}{\sum_{k=-n_1}^{n_1-1}\psi_\eta(n_1x-k)},$$

x ∈ I, n_1, n_2 ∈ N with $n_1 < n_2$. We establish density results for the space of continuous functions and p-integrable functions.

Theorem 3: Let $f \in C(I)$ and η be a sigmoidal function satisfying the above assumptions. For any ε > 0, ∃ n_1, n_2 ∈ N such that

$$\|K_{(n_1,n_2)}f - f\|_\infty < \epsilon.$$

Theorem 4: Let $f \in L^p(I)$ and η be a sigmoidal function satisfying the above assumptions. For any ε > 0, ∃ n_1, n_2 ∈ N such that

$$\|K_{(n_1,n_2)}f - f\|_p < \epsilon.$$

Multi-layers neural network operators

For a locally integrable function $f : I \to R$ and for sigmoidal function η satisfying the above assumptions. For 2 ≤ m ∈ N, we define the m-layer NNOs as follows

$$\left(K_{\underline{n}_{[m]}}f\right)(x) := \frac{\sum_{k=-n_1}^{n_1-1}\left(n_1\int_{\frac{k}{n_1}}^{\frac{k+1}{n_1}} f(u)du\right)\psi_\eta\left(n_1\left(K_{\underline{n}_{[m-1]}}\theta\right)(x) - k\right)}{\sum_{k=-n_1}^{n_1-1}\psi_\eta(n_1x-k)},$$

where θ is the identity function defined on I, $\underline{n}_{[m]} = (n_1, n_2, n_3, \ldots, n_m)$ $\in N^m$ and $\underline{n}_{[m-1]} = (n_2, n_3, \ldots, n_m)$ $\in N^{m-1}$. We establish density results for the space of continuous functions and p-integrable functions.

Theorem 5: Let $f \in C(I)$ and η be a sigmoidal function satisfying the above assumptions. Then, for every ε > 0, ∃ $\underline{n}_{[m]} \in N^m$, 2 ≤ m ∈ N, such that

$$\|K_{\underline{n}_{[m]}}f - f\|_\infty < \epsilon.$$

Theorem 6: Let $f \in L^p(I)$ and η be a sigmoidal function satisfying the above assumptions. Then, for every ε > 0, ∃ $\underline{n}_{[m]} \in N^m$, 2 ≤ m ∈ N, such that

$$\|K_{\underline{n}_{[m]}}f - f\|_p < \epsilon.$$

Conclusions

In this study, we have modified the definition of Kantorovich type NNOs and proved the density results by deep neural network for the space of continuous functions and p-integrable functions. The study can be applied to measure the rate of approximation when f ∈ $L^p(I)$ is approximated by the operators $K_{\underline{n}_{[m]}}$. One can also investigate the role of number of layers in an NNO in improving the rate of approximation.

Journal of Mathematical Analysis and Applications 533 (2024), 128009, https://doi.org/10.1016/j.jmaa.2023.128009