 ISBN 978-9937-0-9019-3

Approximation by Sigmoid Functions with respect to Linear Transformation

Shreejana Koirala
Researcher
Hazesoft Pvt. Ltd.
Kathmandu, Nepal
shreejana@hazesoft.com

Sudeep K. Rayamajhi
Researcher
Hazesoft Pvt. Ltd.
Kathmandu, Nepal
sudeep.rayamajhi@hazesoft.com

***Abstract*—In this paper, we discuss about the approximation of any continuous function by the sigmoid functions on a unit cuboid and simplify the approximation by choosing some appropriate linear transformation B_h.**

***Index Terms*—Function Approximation, Linear Transformation**

Mathematics Subject Class [2020]$\{46N99, 68T07\}$

I. Introduction

Let Ω be a set in $\mathbb{R}^n$. The space $\overline{C}(\Omega)$ is the space of all uniformly continuous functions on Ω, with norm

$$||f||_{C(\Omega)} := \sup_{x\in\Omega} |f(x)| < \infty.$$

For an arbitrary function from $\overline{C}(\Omega)$ we have

$$\lim_{h\to 0} ||f(x+h) - f(x)||_{C(\Omega_h)} = 0;$$

where $h \in \mathbb{R}^n$ and $\Omega_h := \Omega \cap (\Omega - h)$, $(\Omega - h)$ is the translation of Ω by the vector $-h$. The property follows from the definition of uniform continuity. The property is also known as continuity with respect to translation.

In 1989, Bajracharya and Burenkov [1] obtained the necessary and sufficient conditions for the continuity with respect to a non-linear translation, $B_h : \Omega \to \mathbb{R}^n$.

Theorem 1.1 (Bajracharya & Burenkov):

Let $\Omega \subset \mathbb{R}^n$ be an open set , $\delta > 0$, and for all $h \in \mathbb{R}^n$, $|h| < \delta$, let a transformation $B_h : \Omega \to \mathbb{R}^n$. For all $f \in \overline{C}(\Omega)$

$$\lim_{h\to 0} ||f(B_h) - f||_{C(\Omega_h)} = 0$$

it is necessary and sufficient that

$$\lim_{h\to 0} |||B_h(x) - x|||_{C(\Omega_h)} = 0$$

where $\Omega_{\{h\}} = \{x \in \Omega : B_h(x) \in \Omega\}$

Definition 1.1: A sigmoid function is a continuous monotone function with $\lim_{t\to -\infty} \sigma(t) = 0$ and $\lim_{t\to\infty} \sigma(t) = 1$

In 1989, Cybenko [2] showed that any continuous function on a compact subset of $[0,1]^n$ can be approximated by a feed forward neural network with only one single hidden layer and a finite number of neurons, i.e. by superpositions of sigmoid functions.

The functions of form $\sum^N \sigma(Wx + b)$ approximate a continuous function on a cuboid.

Theorem 1.2 (Cybenko): Let σ be a continuous monotone function with $\lim_{t\to -\infty} \sigma(t) = 0$ and $\lim_{t\to\infty} \sigma(t) = 1$. Then, the set of functions of the form $\sum \alpha_j \sigma(W_j^T x + b_j)$ is dense in $C([0,1]^n)$.

Definition 1.2: Let V and W be vector spaces over the same field K. A function $f : V \to W$ is said to be a linear transformation if for any two vectors $u, v \in V$ and any scalar $c \in K$ the following two conditions are satisfied:

Additivity: $f(u+v) = f(u) + f(v)$

Homogenity: $f(cu) = cf(u)$.

Then our main result is the following.

Theorem 1.3: Let f be a continuous function on $[0,1]^n, h \in \mathbb{R}^n, B_h : \mathbb{R}^n \to \mathbb{R}^n$ a linear transformation and σ a uniformly continuous sigmoid function. If

$$\lim_{||h||\to 0} ||B_h(x) - x|| = 0$$

then, for every $h \in R^n$, for some $N \in \mathbb{N}$,

$$\lim_{||h||\to 0} |f(x) - \sum^N \alpha_j \sigma(B_h(x)| = 0.$$

II. Approximation with respect to Linear Transfomation

Let B_h be a linear transformation $B_h : \mathbb{R}^n \to \mathbb{R}^n$.
By Theorem 1.1

$$\lim_{||h||\to 0} ||B_h(x) - x|| = 0$$

iff

$$\lim_{||h||\to 0} ||\sigma(B_h(x)) - \sigma(x)|| = 0.$$

By theorem 1.2, for any function $f \in C[0,1]^n$ we have for $\epsilon > 0$, some $N \in \mathbb{N}$ such that

$$||f(x) - \sum^N \alpha_j \sigma(W(x) + b_j)|| < \epsilon.$$

 ISBN 978-9937-0-9019-3

Now,

$$\lim_{||h||\to 0} ||\sigma(B_h(x)) - \sigma(x)|| = 0$$

$\implies$

$$\lim_{||h||\to 0} ||\sum^{N} \alpha_j \sigma(B_h(x)) - \sum^{N} \alpha_j \sigma(x)|| = 0$$

Thus, for every function f and $\epsilon > 0$, there is some N such that

$$||f(x) - \sum^{N} \alpha_j \sigma(B_h(x)|| < \epsilon.$$

Moreover,

$$\lim_{||h\to 0||} ||f(x) - \sum^{N} \alpha_j \sigma(B_h(x)|| = ||f(x) - \sum^{N} \alpha_j \sigma(x)|| < \epsilon.$$

Example 2.1: Consider $B_h(x) = (I - |h|W)x + |h|b)$ where W is a matrix of order n; $h, b \in R^n$ and I is an identity matrix, then

$$\lim_{||h||\to 0} ||B_h(x) - I(x)|| = 0$$

Problem 2.1: Let B_h be a square matrix of order n and $x, h \in R^n$, we look for general conditions for B_h such that

$$\lim_{||h||\to 0} ||B_h(x) - I(x)|| = 0$$

Problem 2.2: What additional assumptions are required to approximate discontinuous functions?

III. Conclusion

In this paper, we simplified the approximation of continuous function by uniform continuous sigmoid functions with respect to some appropriate linear transformations.

Acknowledgment

We would like to express our very great appreciation to Prof. Dr. Prakash Muni Badzhachariya for his valuable and constructive suggestions. We would also like to thank HazeSoft Pvt. Ltd. for their valuable support.

References

[1] P. M. Badzhrachariya, V. I. Burenkov, “Necessary and sufficient conditions for continuity with respect to a nonlinear translation for various function spaces”, Proc. Steklov Inst. Math., 194 (1993), 1–12

[2] Cybenko, G. Approximation by superpositions of a sigmoidal function. Math. Control Signal Systems 2, 303–314 (1989). https://doi.org/10.1007/BF02551274