IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

A New Class of 2–parametric Deterministic Activation Functions

Nikolay Kyurkchiev[1,2], Anton Iliev[1,2]

Faculty of Mathematics and Informatics, University of Plovdiv Paisii Hilendarski, 24, Tzar Asen Str., 4000 Plovdiv, Bulgaria[1]

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 8, 1113 Sofia, Bulgaria[2]

Abstract: We will explore the interesting methodological task for constructing new 2–parametric deterministic activation function – (2PDAF). We prove upper and lower estimates for the Hausdorff approximation of the sign function by means of this new activation functions. Numerical examples, illustrating our results are given.

Keywords: 2–parametric deterministic activation function – (2PDAF), Sign function, Hausdorff distance, Upper and lower bounds.

I. INTRODUCTION

The interesting task of approximating the function $sign(t)$ with activation functions is important in the treatment of questions related to the study of the ”super saturation” - the object of the research in various fields - neural networks, bioinformatics, nucleation theory, population dynamics, engineering sciences and others. A new activation function using ”correcting amendments” for example from a combination of amendment of ”Gompertz - type” and ”Hyperbolic Tangent - type” is considered in [3]:

$$\varphi(t;k) = \frac{e^{-e^{e^{\cdot^{\cdot^{\cdot^{e^{-at}}}}}}} - e^{-e^{e^{\cdot^{\cdot^{\cdot^{e^{at}}}}}}}}{e^{-e^{e^{\cdot^{\cdot^{\cdot^{e^{-at}}}}}}} + e^{-e^{e^{\cdot^{\cdot^{\cdot^{e^{at}}}}}}}}, \tag{1}$$

where k means the number of recursive insertion of exp (given with sign $''+''$ in (1).
Following this idea, in this note we construct 2–parametric deterministic activation function – (2PDAF).

II. PRELIMINARIES

Definition 1. The sign function of a real number t is defined as follows:

$$sgn(t) = \begin{cases} -1, & \text{if} \quad t < 0, \\ 0, & \text{if} \quad t = 0, \\ 1, & \text{if} \quad t > 0. \end{cases} \tag{2}$$

Definition 2.[1] The Hausdorff distance (the H–distance) $\rho(f,g)$ between two interval functions f,g on $\Omega \subseteq \mathbb{R}$, is the distance between their completed graphs $F(f)$ and $F(g)$ considered as closed subsets of $\Omega \times \mathbb{R}$. More precisely,

$$\rho(f,g) = \max\{ \sup_{A\in F(f)} \inf_{B\in F(g)} \|A-B\|, \sup_{B\in F(g)} \inf_{A\in F(f)} \|A-B\|\}, \tag{3}$$

wherein $\|.\|$ is any norm in $\mathbb{R}^2$, e. g. the maximum norm $\|(t,x)\| = \max\{|t|,|x|\}$; hence the distance between the points $A=(t_A,x_A)$, $B=(t_B,x_B)$ in $\mathbb{R}^2$ is $\|A-B\| = max(|t_A - t_B|, |x_A - x_B|)$.

Definition 3. We define the following ”2-parametric deterministic activation function” (2PDAF):

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

$$\varphi_8(t;a,b) = \frac{a^{-b^{a^{-t}}} - a^{-b^{a^{t}}}}{a^{-b^{a^{-t}}} + a^{-b^{a^{t}}}} \quad (4)$$

III. MAIN RESULTS

The H-distance $d(sgn(t), \varphi_8(t))$ between the sgn function and the function φ_8 satisfies the relation:

$$\varphi_8(d) = \frac{a^{-b^{a^{-d}}} - a^{-b^{a^{d}}}}{a^{-b^{a^{-d}}} + a^{-b^{a^{d}}}} = 1-d. \quad (5)$$

The following Theorem gives upper and lower bounds for d

Theorem. For the Hausdorff distance d between the sgn function and the function φ_8 the following inequalities hold for

$$b(\ln a)^2 \ln b > \frac{e^2}{2} - 1$$

$$d_l = \frac{1}{2\left(1+b(\ln a)^2 \ln b\right)} < d < \frac{\ln\left(2\left(1+b(\ln a)^2 \ln b\right)\right)}{2\left(1+b(\ln a)^2 \ln b\right)} = d_r. \quad (6)$$

Proof. We define the functions

$$F(d) = \frac{a^{-b^{a^{-d}}} - a^{-b^{a^{d}}}}{a^{-b^{a^{-d}}} + a^{-b^{a^{d}}}} - 1 + d, \quad (7)$$

$$G(d) = -1 + (1+b(\ln a)^2 \ln b)d. \quad (8)$$

From Taylor expansion we find (see, Fig. 1)

$$F(d) - G(d) = O(d^3).$$

In addition $G'(d) > 0$.

We look for two reals d_l and d_r such that $G(d_l) < 0$ nad $G(d_r) > 0$ (leading to $G(d_l) < G(d) < G(d_r)$ and thus $d_l < d < d_r$).

Trying $d_l = \frac{1}{2\left(1+b(\ln a)^2 \ln b\right)}$ and $d_r = \frac{\ln\left(2\left(1+b(\ln a)^2 \ln b\right)\right)}{2\left(1+b(\ln a)^2 \ln b\right)}$ we obtain for $b(\ln a)^2 \ln b > \frac{e^2}{2} - 1$

$$G(d_l) < 0; \; G(d_r) > 0.$$

This completes the proof of the inequalities (6).

Approximation of the $sgn(t)$ by (2PDAF)–function for $a = 3.9$, $b = 3.8$ is visualized on Fig. 2.

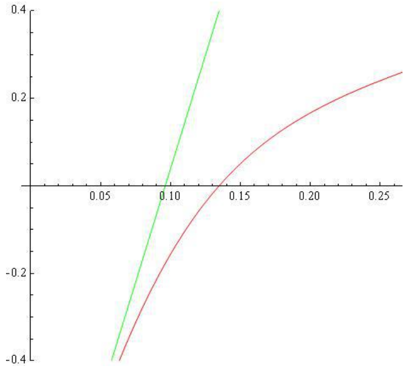

Fig. 1: The functions $F(d)$ and $G(d)$ for $a = 3.9$, $b = 3.8$.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

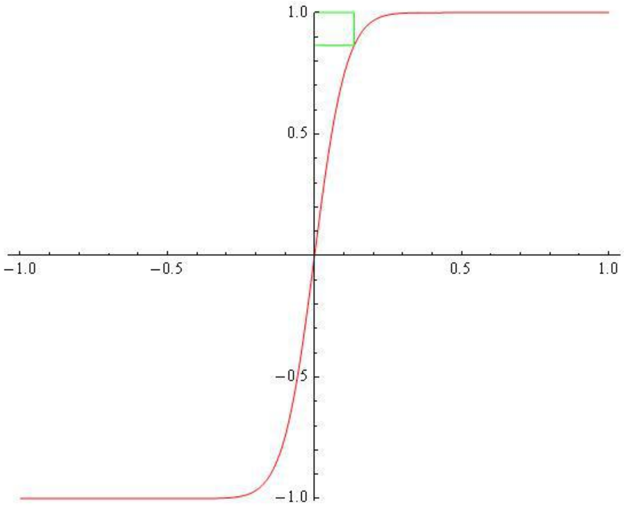

Fig 2: Approximation of the $sgn(t)$ by (2PDAF) for $a = 3.9$, $b = 3.8$; Hausdorff distance: $d = 0.134639$; $d_l = 0.0480931$; $d_r = 0.145944$.

From the graphics it can be seen that the "saturation" is faster. For other results, see [2]–[5].

REMARK

The reader can be formulate the "General case" using k recursive insertion of a and b in (5) as well to explore respectively approximation task.

ACKNOWLEDGMENT

This work has been supported by the project FP17-FMI-008 of Department for Scientific Research, Paisii Hilendarski University of Plovdiv.

REFERENCES

[1] Hausdorff F. (1962) Set Theory (2 ed.), New York, Chelsea Publ.
[2] Kyurkchiev N., A. Iliev, S. Markov (2017) Some techniques for recurrence generating of activation functions, LAP LAMBERT Academic Publishing; ISBN 978-3-330-33143-3.
[3] Kyurkchiev, N., A new class of activation functions based on the correcting amendments, Int. J. for Scientific Research and Development, 6 (2), 2018, (accepted).
[4] Kyurkchiev, N., The new transmuted C.D.F. based on Gompertz function, Biomath Communications, 5 (1), 2018.
[5] Pavlov, N., G. Spasov, A. Rahnev, N. Kyurkchiev, Some deterministic reliability growth curves for software error detection: approximation and modeling aspects, IJPAM, 118, 2018, (accepted).