INTELLIGENT SYSTEMS REFERENCE LIBRARY
Volume 19

George A. Anastassiou

Intelligent Systems: Approximation by Artificial Neural Networks

Springer

George A. Anastassiou

Intelligent Systems: Approximation by Artificial Neural Networks

Intelligent Systems Reference Library, Volume 19

Editors-in-Chief

Further volumes of this series can be found on our homepage: springer.com

Vol. 1. Christine L. Mumford and Lakhmi C. Jain (Eds.)
Computational Intelligence: Collaboration, Fusion and Emergence, 2009
ISBN 978-3-642-01798-8

Vol. 2. Yuehui Chen and Ajith Abraham
Tree-Structure Based Hybrid Computational Intelligence, 2009
ISBN 978-3-642-04738-1

Vol. 3. Anthony Finn and Steve Scheding
Developments and Challenges for Autonomous Unmanned Vehicles, 2010
ISBN 978-3-642-10703-0

Vol. 4. Lakhmi C. Jain and Chee Peng Lim (Eds.)
Handbook on Decision Making: Techniques and Applications, 2010
ISBN 978-3-642-13638-2

Vol. 5. George A. Anastassiou
Intelligent Mathematics: Computational Analysis, 2010
ISBN 978-3-642-17097-3

Vol. 6. Ludmila Dymowa
Soft Computing in Economics and Finance, 2011
ISBN 978-3-642-17718-7

Vol. 7. Gerasimos G. Rigatos
Modelling and Control for Intelligent Industrial Systems, 2011
ISBN 978-3-642-17874-0

Vol. 8. Edward H.Y. Lim, James N.K. Liu, and Raymond S.T. Lee
Knowledge Seeker – Ontology Modelling for Information Search and Management, 2011
ISBN 978-3-642-17915-0

Vol. 9. Menahem Friedman and Abraham Kandel
Calculus Light, 2011
ISBN 978-3-642-17847-4

Vol. 10. Andreas Tolk and Lakhmi C. Jain
Intelligence-Based Systems Engineering, 2011
ISBN 978-3-642-17930-3

Vol. 11. Samuli Niiranen and Andre Ribeiro (Eds.)
Information Processing and Biological Systems, 2011
ISBN 978-3-642-19620-1

Vol. 12. Florin Gorunescu
Data Mining, 2011
ISBN 978-3-642-19720-8

Vol. 13. Witold Pedrycz and Shyi-Ming Chen (Eds.)
Granular Computing and Intelligent Systems, 2011
ISBN 978-3-642-19819-9

Vol. 14. George A. Anastassiou and Oktay Duman
Towards Intelligent Modeling: Statistical Approximation Theory, 2011
ISBN 978-3-642-19825-0

Vol. 15. Antonino Freno and Edmondo Trentin
Hybrid Random Fields, 2011
ISBN 978-3-642-20307-7

Vol. 16. Alexiei Dingli
Knowledge Annotation: Making Implicit Knowledge Explicit, 2011
ISBN 978-3-642-20322-0

Vol. 17. Crina Grosan and Ajith Abraham
Intelligent Systems, 2011
ISBN 978-3-642-21003-7

Vol. 18. Achim Zielesny
From Curve Fitting to Machine Learning, 2011
ISBN 978-3-642-21279-6

Vol. 19. George A. Anastassiou
Intelligent Systems: Approximation by Artificial Neural Networks, 2011
ISBN 978-3-642-21430-1

George A. Anastassiou

Intelligent Systems: Approximation by Artificial Neural Networks

Springer

Prof. George A. Anastassiou
University of Memphis
Department of Mathematical Sciences
Memphis, TN 38152
USA
E-mail: ganastss@memphis.edu

ISBN 978-3-642-21430-1 e-ISBN 978-3-642-21431-8

DOI 10.1007/978-3-642-21431-8

Intelligent Systems Reference Library ISSN 1868-4394

Library of Congress Control Number: 2011930524

Typeset & *Cover Design:* Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

This brief monograph is the first one to deal exclusively with the quantitative approximation by artificial neural networks to the identity-unit operator. Here we study with rates the approximation properties of the "right" sigmoidal and hyperbolic tangent artificial neural network positive linear operators. In particular we study the degree of approximation of these operators to the unit operator in the univariate and multivariate cases over bounded or unbounded domains. This is given via inequalities and with the use of modulus of continuity of the involved function or its higher order derivative. We examine the real and complex cases.

For the convenience of the reader, the chapters of this book are written in a self-contained style.

This treatise relies on author's last two years of related research work.

Advanced courses and seminars can be taught out of this brief book. All necessary background and motivations are given per chapter. A related list of references is given also per chapter. My book's results appeared for the first time in my published articles which are mentioned throughout the references. They are expected to find applications in many areas of computer science and applied mathematics, such as neural networks, intelligent systems, complexity theory, learning theory, vision and approximation theory, etc. As such this monograph is suitable for researchers, graduate students, and seminars of the above subjects, also for all science libraries.

The preparation of this booklet took place during 2010-2011 in Memphis, Tennessee, USA.

I would like to thank my family for their dedication and love to me, which was the strongest support during the writing of this book.

March 1, 2011

George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis, TN, U.S.A.

Contents

Chapter 1
Univariate Sigmoidal Neural Network Quantitative Approximation

Here we give the univariate quantitative approximation of real and complex valued continuous functions on a compact interval or all the real line by quasi-interpolation sigmoidal neural network operators. This approximation is obtained by establishing Jackson type inequalities involving the modulus of continuity of the engaged function or its high order derivative. The operators are defined by using a density function induced by the logarithmic sigmoidal function. Our approximations are pointwise and with respect to the uniform norm. The related feed-forward neural network is with one hidden layer. This chapter relies on [4].

1.1 Introduction

Feed-forward neural networks (FNNs) with one hidden layer, the only type of networks we deal with in this chapter, are mathematically expressed as

$$N_n(x) = \sum_{j=0}^{n} c_j \sigma\left(\langle a_j \cdot x \rangle + b_j\right), \quad x \in \mathbb{R}^s, \quad s \in \mathbb{N},$$

where for $0 \leq j \leq n$, $b_j \in \mathbb{R}$ are the thresholds, $a_j \in \mathbb{R}^s$ are the connection weights, $c_j \in \mathbb{R}$ are the coefficients, $\langle a_j \cdot x \rangle$ is the inner product of a_j and x, and σ is the activation function of the network. In many fundamental network models, the activation function is the sigmoidal function of logistic type.

It is well known that FNNs are universal approximators. Theoretically, any continuous function defined on a compact set can be approximated to any desired degree of accuracy by increasing the number of hidden neurons. It was shown by Cybenko [11] and Funahashi [13], that any continuous function

G.A. Anastassiou: Intelligent Systems: Approximation by ANN, ISRL 19, pp. 1–32.
springerlink.com

can be approximated on a compact set with uniform topology by a network of the form $N_n(x)$, using any continuous, sigmoidal activation function. Hornik et al. in [15], have proved that any measurable function can be approached with such a network. Furthermore, these authors established in [16], that any function of the Sobolev spaces can be approached with all derivatives. A variety of density results on FNN approximations to multivariate functions were later established by many authors using different methods, for more or less general situations: [18] by Leshno et al., [22] by Mhaskar and Micchelli, [10] by Chui and Li, [8] by Chen and Chen, [14] by Hahm and Hong, etc.

Usually these results only give theorems about the existence of an approximation. A related and important problem is that of complexity: determining the number of neurons required to guarantee that all functions belonging to a space can be approximated to the prescribed degree of accuracy ϵ.

Barron [5] shows that if the function is assumed to satisfy certain conditions expressed in terms of its Fourier transform, and if each of the neurons evaluates a sigmoidal activation function, then at most $O\left(\epsilon^{-2}\right)$ neurons are needed to achieve the order of approximation ϵ. Some other authors have published similar results on the complexity of FNN approximations: Mhaskar and Micchelli [23], Suzuki [24], Maiorov and Meir [20], Makovoz [21], Ferrari and Stengel [12], Xu and Cao [26], Cao et al. [7], etc.

The author in [1] and [2], see chapters 2-5, was the first to establish neural network approximations to continuous functions with rates by very specifically defined neural network operators of Cardaliagnet-Euvrard and "Squashing" types, by employing the modulus of continuity of the engaged function or its high order derivative, and producing very tight Jackson type inequalities. He treats there both the univariate and multivariate cases. The defining these operators "bell-shaped" and "squashing" function are assumed to be of compact support. Also in [2] he gives the Nth order asymptotic expansion for the error of weak approximation of these two operators to a special natural class of smooth functions, see chapters 4-5 there.

For this chapter the author is greatly motivated by the important article [9] by Z. Chen and F. Cao.

He presents related to it work and much more beyond however [9] remains the initial point. So the author here performs univariate sigmoidal neural network approximations with rates to continuous functions over compact intervals of the real line or over the whole $\mathbb{R}$, then he extends his results to complex valued functions. All convergences here are with rates expressed via the modulus of continuity of the involved function or its high order derivative, and given by very tight Jackson type inequalities.

The author presents here the correct and precisely defined quasi-interpolation neural network operator related to compact intervals, and among others, improves results from [9]. The compact intervals are not necessarily symmetric to the origin. Some of the upper bounds to error quantity are very flexible and general. In preparation to establish our results we prove further properties of the basic density function defining our operators.

1.2 Background and Auxiliary Results

We consider here the sigmoidal function of logarithmic type

$$s(x) = \frac{1}{1+e^{-x}}, \quad x \in \mathbb{R}.$$

It has the properties $\lim_{x \to +\infty} s(x) = 1$ and $\lim_{x \to -\infty} s(x) = 0$.

This function plays the role of an activation function in the hidden layer of neural networks, also has application in biology, demography, etc. ([6, 17]).

As in [9], we consider

$$\Phi(x) := \frac{1}{2}\left(s(x+1) - s(x-1)\right), \quad x \in \mathbb{R}.$$

It has the following properties:

i) $\Phi(x) > 0, \ \forall\, x \in \mathbb{R}$,

ii) $\sum_{k=-\infty}^{\infty} \Phi(x-k) = 1, \ \forall\, x \in \mathbb{R}$,

iii) $\sum_{k=-\infty}^{\infty} \Phi(nx-k) = 1, \ \forall\, x \in \mathbb{R}; \, n \in \mathbb{N}$,

iv) $\int_{-\infty}^{\infty} \Phi(x)\, dx = 1$,

v) Φ is a density function,

vi) Φ is even: $\Phi(-x) = \Phi(x), \ x \geq 0$.

We observe that ([9])

$$\Phi(x) = \left(\frac{e^2-1}{2e}\right) \frac{e^{-x}}{\left(1+e^{-x-1}\right)\left(1+e^{-x+1}\right)} = \left(\frac{e^2-1}{2e^2}\right) \frac{1}{\left(1+e^{x-1}\right)\left(1+e^{-x-1}\right)},$$

and

$$\Phi'(x) = \left(\frac{e^2-1}{2e^2}\right)\left[-\frac{\left(e^x - e^{-x}\right)}{e\left(1+e^{x-1}\right)^2\left(1+e^{-x-1}\right)^2}\right] \leq 0, \ x \geq 0.$$

Hence

vii) Φ is decreasing on $\mathbb{R}_+$, and increasing on $\mathbb{R}_-$.

Let $0 < \alpha < 1$, $n \in \mathbb{N}$. We see the following

$$\sum_{\begin{cases} k=-\infty \\ :|nx-k|>n^{1-\alpha} \end{cases}}^{\infty} \Phi(nx-k) = \sum_{\begin{cases} k=-\infty \\ :|nx-k|>n^{1-\alpha} \end{cases}}^{\infty} \Phi(|nx-k|) \leq$$

$$\left(\frac{e^2-1}{2e^2}\right) \int_{(n^{1-\alpha}-1)}^{\infty} \frac{1}{(1+e^{x-1})(1+e^{-x-1})} dx \leq$$

$$\left(\frac{e^2-1}{2e}\right) \int_{(n^{1-\alpha}-1)}^{\infty} e^{-x} dx = \left(\frac{e^2-1}{2e}\right) \left(e^{-\left(n^{1-\alpha}-1\right)}\right)$$

$$= \left(\frac{e^2-1}{2}\right) e^{-n^{(1-\alpha)}} = 3.1992 e^{-n^{(1-\alpha)}}.$$

We have found that:

viii) for $n \in \mathbb{N}$, $0 < \alpha < 1$, we get

$$\sum_{\begin{cases} k=-\infty \\ :|nx-k|>n^{1-\alpha} \end{cases}}^{\infty} \Phi(nx-k) < \left(\frac{e^2-1}{2}\right) e^{-n^{(1-\alpha)}} = 3.1992 e^{-n^{(1-\alpha)}}.$$

Denote by $\lceil \cdot \rceil$ the ceiling of a number, and by $\lfloor \cdot \rfloor$ the integral part of a number. Consider $x \in [a,b] \subset \mathbb{R}$ and $n \in \mathbb{N}$ such that $\lceil na \rceil \leq \lfloor nb \rfloor$.

We observe that

$$1 = \sum_{k=-\infty}^{\infty} \Phi(nx-k) > \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \Phi(nx-k) =$$

$$\sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \Phi(|nx-k|) > \Phi(|nx-k_0|),$$

for any $k_0 \in [\lceil na \rceil, \lfloor nb \rfloor] \cap \mathbb{Z}$.

Here we can choose $k_0 \in [\lceil na \rceil, \lfloor nb \rfloor] \cap \mathbb{Z}$ such that $|nx-k_0| < 1$.