



ELSEVIER

Contents lists available at ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

Multivariate Jackson-type inequality for a new type neural network approximation [☆]

Shaobo Lin ^{*}, Yuanhua Rong, Zongben Xu

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China

ARTICLE INFO

Article history:

Received 18 July 2012

Received in revised form 21 March 2014

Accepted 20 May 2014

Available online 29 May 2014

Keywords:

Neural networks

Jackson-type inequality

Error estimate

Sigmoidal function

ABSTRACT

In this paper, we introduce a new type neural networks by superpositions of a sigmoidal function and study its approximation capability. We investigate the multivariate quantitative constructive approximation of real continuous multivariate functions on a cube by such type neural networks. This approximation is derived by establishing multivariate Jackson-type inequalities involving the multivariate modulus of smoothness of the target function. Our networks require no training in the traditional sense.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Feed-forward neural networks (FNNs) can be formally described as devices producing input–output functions depending on flexible parameters. Often input–output functions have the form of linear combinations of functions computable by units specific for the given type of networks. Both coefficients of the linear combinations and parameters of the computational units are adjustable in the process of learning. Mathematically, the FNN can be represented by

$$\sum_{i=1}^M c_i \sigma(a_i \cdot x + b_i), \quad x \in \mathbf{R}^d, \quad (1.1)$$

where $a_i \in \mathbf{R}^d$, $c_i \in \mathbf{R}$, $b_i \in \mathbf{R}$ are the inner weight, outer weight, and threshold of the FNN, respectively.

Theoretically, any continuous functions defined on a compact set in \mathbf{R}^d can be approximated by neural networks to any desired accuracy by increasing the number of hidden neurons. This result is usually called the density problem of FNNs. This problem has been tackled in [1–7] and references therein. Compared to the density problem, a related and more important problem is the complexity: to determine how many neurons are necessary to yield a prescribed degree of approximation. There have been many studies for this problem. We refer the readers to Anastassiou [8], Barron [9], Ferrari and Stengel [10], Maiorov and Meir [11], Makovoz [12], Mhaskar and Micchelli [13] for more information about the complexity problem.

On the other hand, the Jackson-type inequality which describes the relation between the smoothness of the target function and the rate of approximation has been extensively used in approximation theory and neural networks. If the activation

[☆] The research was supported by the National 973 Programming (2013CB329404), the Key Program of National Natural Science Foundation of China (Grant No. 11131006).

^{*} Corresponding author. Tel.: +86 18392137967.

E-mail address: sblin1983@gmail.com (S. Lin).

function σ is analytic and non-polynomial, Mhaskar [14] proved that the Jackson-type inequality held for neural networks formed as (1.1). If σ is a sigmoidal function, i.e.,

$$\lim_{t \rightarrow \infty} \sigma(t) = 1, \quad \lim_{t \rightarrow -\infty} \sigma(t) = 0,$$

then Chen [15] established a Jackson-type inequality for neural networks (1.1) with $d = 1$. He also proposed an open question that whether the Jackson-type inequality held for neural network (1.1) with $d \geq 2$ if σ is sigmoidal. In the recent paper [8], Anastassiou devoted to giving an answer to this question. However, he only proved that if the activation function is a combination of sigmoidal function, then the Jackson-type inequality held. In this paper, we will give an answer to the above question in another direction. We will prove that if we give a little change for the structure of neural networks by introducing a distance based on a partition of the cube, then the multivariate Jackson-type inequality holds for the new type of neural network with sigmoidal activation function.

This paper is organized as follows. In the next section, we will introduce a partition-based distance on and give the construction of the neural networks. Our main result will be given in the third section, where a Jackson-type error estimates for approximation by neural network will be given. In Section 4, we will verify our statement by two simulation experiments. In the last section, we will draw a conclusion of this paper.

2. Construction of neural networks

In this section, we introduce a new type of neural networks. Let $\mathbf{I}^d := [0, 1]^d$. In [15], Chen proved that the neural networks formed as

$$N_n^*(x) = c_0 + \sum_{i=1}^{n-1} c_i \sigma(a_i x + b_i), \quad (2.1)$$

possessed prominent approximation capability when $d = 1$ and σ is a bounded sigmoidal function. An obvious extension for this type of neural networks to multivariate case is

$$N_n^{(1)}(\mathbf{x}) := c_0 + \sum_{i=1}^{n-1} c_i \sigma(\mathbf{a}_i \cdot \mathbf{x} + b_i), \quad (2.2)$$

where $\mathbf{x} \cdot \mathbf{y}$ denotes the inner product between the vectors \mathbf{x} and \mathbf{y} . However, to the best of our knowledge, it is not very easy to establish a Jackson-type inequality for such type of network if $d \geq 2$ and σ is sigmoidal. The main reason is that there is not a strict order for any points in \mathbf{I}^d when $d \geq 2$, but there are strict orders for approximation by neural networks [16]. So, we turn to another type of extension of (2.1). For $u = 0$, we have

$$a_i x + b_i = a_i(x - v_i) = a_i((x - u) - (v_i - u)) = a_i(d(x, u) - d(v_i, u)),$$

where $d(x, y)$ denotes the Euclidean distance between x and y . Thus, we can rewrite (2.1) as

$$N_n^*(x) = c_0 + \sum_{i=1}^{n-1} c_i \sigma(a_i(d(x, u) - d(v_i, u))).$$

Noting this, we define

$$N_n(\mathbf{x}) := c_0 + \sum_{i=1}^{n-1} c_i \sigma(a_i(\bar{d}(\mathbf{x}, \mathbf{u}) - \bar{d}(\mathbf{v}_i, \mathbf{u}))), \quad (2.3)$$

where $\bar{d}(\mathbf{x}, \mathbf{y})$ is a partition-based distance between \mathbf{x} and \mathbf{y} which is introduced to guarantee the order for points in \mathbf{I}^d . Thus, as far as the strict order is concerned, the special neural networks (2.3) is a more suitable extension of the univariate counterpart (2.1). Based on this property, we can deduce a multivariate Jackson inequality for neural networks approximation, which can be regarded as an extension of Chen's [15] result. The proposed network $N_n(\cdot)$ can be interpreted as a model of feed-forward neural networks with four layers:

- The first one is the input layer with the input \mathbf{x} ($\mathbf{x} \in \mathbf{I}^d$).
- The second one is the pre-handling layer, which transform an input \mathbf{x} into the partition-based distance between \mathbf{u} and \mathbf{x} , $\bar{d}(\mathbf{x}, \mathbf{u})$.
- The third one is the handling layer with n neurons in it.
- The last one is the output layer.

Before giving a concrete definition for $\bar{d}(\cdot, \cdot)$, we need give a division of \mathbf{I}^d . For the sake of brevity, we only study it for $d = 2$. Divide \mathbf{I}^2 into n^2 small square with length $\frac{1}{n}$, J_k , $k = 1, \dots, n^2$. Let $\mathbf{x}_1, \dots, \mathbf{x}_{n^2}$ be the centers of J_1, \dots, J_{n^2} , and $d(\mathbf{x}_k, \mathbf{x}_{k+1}) = \frac{1}{n}$. For example, if n is odd, then we have (see Fig. 1).

Download English Version:

<https://daneshyari.com/en/article/1703547>

Download Persian Version:

<https://daneshyari.com/article/1703547>

[Daneshyari.com](https://daneshyari.com)