Chapter 3
Multivariate Sigmoidal Neural Network Quantitative Approximation

Here we present the multivariate quantitative constructive approximation of real and complex valued continuous multivariate functions on a box or $\mathbb{R}^N$, $N \in \mathbb{N}$, by the multivariate quasi-interpolation sigmoidal neural network operators. The "right" operators for the goal are fully and precisely described. This approximation is obtained by establishing multidimensional Jackson type inequalities involving the multivariate modulus of continuity of the engaged function or its high order partial derivatives. The multivariate operators are defined by using a multidimensional density function induced by the logarithmic sigmoidal function. Our approximations are pointwise and uniform. The related feed-forward neural network is with one hidden layer. This chapter is based on [5].

3.1 Introduction

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

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

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 $\sigma$ is the activation function of the network. In many fundamental network models, the activation function is the sigmoidal function of logistic type.

To achieve our goals the operators here are more elaborate and complex, please see (3.2) and (3.3) for exact definitions.
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 proved by Cybenko [12] and Funahashi [14], that any continuous function 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 [16], have shown that any measurable function can be approached with such a network. Furthermore, these authors proved in [17], 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: [19] by Leshno et al., [23] by Mhaskar and Micchelli, [11] by Chui and Li, [9] by Chen and Chen, [15] 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 \( \epsilon \).

Barron [6] proves that if the function is supposed 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 \( O(\epsilon^{-2}) \) neurons are needed to achieve the order of approximation \( \epsilon \). Some other authors have published similar results on the complexity of FNN approximations: Mhaskar and Micchelli [24], Suzuki [25], Maiorov and Meir [21], Makovoz [22], Ferrari and Stengel [13], Xu and Cao [27], Cao et al. [8], etc.

The author in [1], [2] and [3], see chapters 2-5, was the first to obtain neural network approximations to continuous functions with rates by very specifically defined neural network operators of Cardaliaget-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 [3] he gives the \( N \)th 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 [10] by Z. Chen and F. Cao, also by [4].

The author here performs multivariate sigmoidal neural network approximations to continuous functions over boxes or over the whole \( \mathbb{R}^N, N \in \mathbb{N} \), then he extends his results to complex valued multivariate functions. All convergences here are with rates expressed via the multivariate modulus of continuity of the involved function or its high order partial derivatives, and given by very tight multidimensional Jackson type inequalities.