Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2022-0151

Vol. **32**, No. 5, pp. 1361-1400
November 2022

A New Function Space from Barron Class and Application to Neural Network Approximation

Yan Meng[1] and Pingbing Ming[2,3,*]

[1] *School of Mathematics, Renmin University of China, Beijing 100872, China.*
[2] *LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, AMSS, Chinese Academy of Sciences, No. 55, East Road Zhong-Guan-Cun, Beijing 100190, China.*
[3] *School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China.*

Received 27 May 2022; Accepted (in revised version) 12 October 2022

Abstract. We introduce a new function space, dubbed as the Barron spectrum space, which arises from the target function space for the neural network approximation. We give a Bernstein type sufficient condition for functions in this space, and clarify the embedding among the Barron spectrum space, the Bessel potential space, the Besov space and the Sobolev space. Moreover, the unexpected smoothness and the decaying behavior of the radial functions in the Barron spectrum space have been investigated. As an application, we prove a dimension explicit L^q error bound for the two-layer neural network with the Barron spectrum space as the target function space, the rate is dimension independent.

AMS subject classifications: 65N30, 65M12, 41A46, 35J25

Key words: Fourier transform, Besov space, Sobolev space, radial function, neural network.

1 Introduction

Several target function spaces have been scattered in the literature for the neural network approximation, such as the potential space, the reproducing kernel space [23, 29, 42], the Sobolev space [44], the Besov space [17] and its variant of the dominating mixed smoothness [54], among many others. In Barron's seminal work [2], he proved that for a function that has a finite first moment of the magnitude of the Fourier transform, the convergence rate for a feedforward artificial neural network with sigmoidal nonlinearity is

*Corresponding author. *Email addresses:* `mengyan@ruc.edu.cn` (Y. Meng), `mpb@lsec.cc.ac.cn` (P. B. Ming)

$\mathcal{O}(n^{-1/2})$ with n the number of the neurons. Roughly speaking, such function class may be rephrased as

$$f \quad \text{has Fourier transform} \quad \widehat{f} \quad \text{with} \quad \int_{\mathbb{R}^d} |x||\widehat{f}(x)|\mathrm{d}x < \infty. \tag{1.1}$$

For any $f \in L^1(\mathbb{R}^d)$, its Fourier transform is defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x \cdot \xi} \mathrm{d}x.$$

The novelty of this result lies in the fact that the convergence rate is independent of the dimension of the ambient space, and cracks the curse of the dimensionality [4, 16]. The spectrum norm was extended to second order to study the approximation rate for hinging functions by Breiman in [9]. Hornik et al [26] introduced the spectrum norm of arbitrary positive integer order m:

$$\int_{\mathbb{R}^d} \max\{1, |x|^m\} |\widehat{f}(x)| \mathrm{d}x.$$

Several works have been devoted to study the spectrum norm in the literature; see [10, §7.2] and [16, 45]. In a series of work [19, 20], E et al have defined a function class, which is dubbed as Barron space. Barron space contains *infinitely wide* neural networks with certain controls over the parameters, which depends on the activation function used in the neural network. The relation among Barron spaces and the so-called *Fourier-analytic Barron space* have been established in [11, §7]. They have studied the pointwise behavior of the functions in Barron space, and have proved the direct and inverse approximation theorems for the two-layer neural network approximation. It is worth mentioning that the approximation class for the deep neural network have recently been investigated in [19, 24].

Motivated by (1.1), for $s \in \mathbb{R}$ and $1 \le p \le 2$, we introduce a new function space $\mathcal{B}_{s,p}(\mathbb{R}^d)$, called Barron spectrum space, which consists of $f \in L^p(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} |x|^s |\widehat{f}(x)| \mathrm{d}x < \infty$ (see Definition 2.1). Compared with the original Barron class or the Fourier-analytic Barron space, we include the L^p-norm in addition to the spectrum norm and our definition is independent of the choice of the activation functions. It is a Banach space as shown in the next part. We shall study the properties of $\mathcal{B}_{s,p}(\mathbb{R}^d)$ and clarify the relationship among $\mathcal{B}_{s,p}(\mathbb{R}^d)$ and certain commonly used target function spaces for the neural network approximation. For any $s > -d/p$ and $1 \le p \le 2$, we prove the embedding

$$B^{s+d/p}_{p,1}(\mathbb{R}^d) \hookrightarrow \mathcal{B}_{s,p}(\mathbb{R}^d) \hookrightarrow B^s_{\infty,1}(\mathbb{R}^d)$$

holds, where $B^\alpha_{p,q}(\mathbb{R}^d)$ is the Besov space with $\alpha \in \mathbb{R}$ and $1 \le p, q \le \infty$. With the aid of this inclusion, we establish the embedding between $\mathcal{B}_{s,p}(\mathbb{R}^d)$ and the Sobolev space. Moreover, we prove the relation between $\mathcal{B}_{s,p}(\mathbb{R}^d)$ and the Bessel potential space. Another