Numer. Math. Theor. Meth. Appl. Vol. **16**, No. 1, pp. 58-78
doi: 10.4208/nmtma.OA-2022-0104 February 2023

Sparse Deep Neural Network for Nonlinear Partial Differential Equations

Yuesheng Xu[1] and Taishan Zeng[2,3,*]

[1] *Department of Mathematics and Statistics, Old Dominion University, Norfolk, Virginia 23529, USA*
[2] *School of Mathematical Science, South China Normal University, Guangzhou 510631, China*
[3] *Guangdong Key Laboratory of Big Data Analysis and Processing, Guangzhou 510006, China*

Received 29 June 2022; Accepted (in revised version) 22 July 2022

Abstract. More competent learning models are demanded for data processing due to increasingly greater amounts of data available in applications. Data that we encounter often have certain embedded sparsity structures. That is, if they are represented in an appropriate basis, their energies can concentrate on a small number of basis functions. This paper is devoted to a numerical study of adaptive approximation of solutions of nonlinear partial differential equations whose solutions may have singularities, by deep neural networks (DNNs) with a sparse regularization with multiple parameters. Noting that DNNs have an intrinsic multi-scale structure which is favorable for adaptive representation of functions, by employing a penalty with multiple parameters, we develop DNNs with a multi-scale sparse regularization (SDNN) for effectively representing functions having certain singularities. We then apply the proposed SDNN to numerical solutions of the Burgers equation and the Schrödinger equation. Numerical examples confirm that solutions generated by the proposed SDNN are sparse and accurate.

AMS subject classifications: 68T07, 65D15, 65F50, 47A52, 65N50
Key words: Sparse approximation, deep learning, nonlinear partial differential equations, sparse regularization, adaptive approximation.

1. Introduction

The goal of this paper is to develop a sparse regularization deep neural network model for numerical solutions of nonlinear partial differential equations whose solutions may have singularities. We will mainly focus on designing a sparse regularization model by employing multiple parameters to balance sparsity of different layers and the

*Corresponding author. *Email addresses:* zengtsh@m.scnu.edu.cn (T. Zeng), y1xu@odu.edu (Y. Xu)

overall accuracy. The proposed ideas are tested in this paper numerically to confirm our intuition and more in-depth theoretical studies will be followed in a future paper.

Artificial intelligence especially deep neural networks (DNN) has received great attention in many research fields. From the approximation theory point of view, a neural network is built by functional composition to approximate a continuous function with arbitrary accuracy. Deep neural networks are proven to have better approximation by practice and theory due to their relatively large number of hidden layers. Deep neural network has achieved state-of-the-art performance in a wide range of applications, including speech recognition [11], computer vision [28], natural language processing [14], and finance [8]. For an overview of deep learning the readers are referred to monograph [20]. Recently, there was great interest in applying deep neural networks to the field of scientific computing, such as discovering the differential equations from observed data [34], solving the partial differential equation (PDE) [21, 29, 30, 35], and problem aroused in physics [16]. Mathematical understanding of deep neural networks received much attention in the applied mathematics community. A universal approximation theory of neural network for Borel measurable function on compact domain is established in [9]. Some recent research studies the expressivity of deep neural networks for different function spaces [15], for example, Sobolev spaces, Barron functions, and Hölder spaces. There are close connections between deep neural network and traditional approximation methods, such as splines [13, 37], compressed sensing [1], and finite elements [22, 26]. Convergence of deep neural networks and deep convolutional neural networks are studied in [40] and [41] respectively. Some work aims at understanding the training process of DNN. For instance, in paper [10], the training process of DNN is interpreted as learning adaptive basis from data.

Traditionally, deep neural networks are dense and over-parameterized. A dense network model requires more memory and other computational resources during training and inference of the model. Increasingly greater amounts of data and related model sizes demand the availability of more competent learning models. Compared to dense models, sparse deep neural networks require less memory, less computing time and have better interpretability. Hence, sparse deep neural network models are desirable. On the other hand, animal brains are found to have hierarchical and sparse structures [19]. The connectivity of an animal brain becomes sparser as the size of the brain grows larger. Therefore, it is not only necessary but also natural to design sparse networks. In fact, it was pointed out in [25] that the future of deep learning relies on sparsity. Furthermore, over-parameterized and dense models tend to lead to overfitting and weakening the ability to generalize over unseen examples. Sparse models can improve accuracy of approximation. Sparse regularization is a popular way to learn the sparse solutions [5, 38, 39, 42]. The readers are referred to [24] for an overview of sparse deep learning.

Although much progress has been made in theoretical research of deep learning, it remains a challenging issue to construct an effective neural network approximation for general function spaces using as few neuron connections or neurons as possible. Most of existing network structures are specific for a particular class of functions. In this