Journal of Complexity 35 (2016) 86–101

Contents lists available at ScienceDirect

Journal of Complexity

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

Linear and nonlinear approximation of spherical radial basis function networks☆

Shaobo Lin

Institute of Intelligent System and Decision, Wenzhou University, Wenzhou 325035, China

ARTICLE INFO

Article history:
Received 18 September 2015
Accepted 12 February 2016
Available online 27 February 2016

Keywords:
Spherical radial basis function networks
Linear approximation
Nonlinear approximation
Pseudo-dimension
Kolmogorov *n*-width

ABSTRACT

In this paper, the center-selection strategy of spherical radial basis function networks (SRBFNs) is considered. To approximate functions in the Bessel-potential Sobolev classes, we provide two lower bounds of nonlinear SRBFN approximation. In the first one, we prove that, up to a logarithmic factor, the lower bound of SRBFN approximation coincides with the Kolmogorov *n*-width. In the other one, we prove that if a pseudo-dimension assumption is imposed on the activation function, then the logarithmic factor can even be omitted. These results together with the well known Jackson-type inequality of SRBFN approximation imply that the center-selection strategy does not affect the approximation capability of SRBFNs very much, provided the target function belongs to the Bessel-potential Sobolev classes. Thus, we can choose centers only for the algorithmic factor. Hence, a linear SRBFN approximant whose centers are specified before the training is recommended.

1. Introduction

Fitting spherical data arising from sampling an unknown function defined on the sphere comes up frequently in applied problems. Examples include the study of seismic signals, gravitational phenomenon, solar corona and medical imaging of the brain. A common procedure to fitting spherical data can boil down to two steps: choosing a specific class of functions to build up the candidates and selecting the final estimate from the candidates by using the spherical data (this process is also called

☆ The research was supported by the National Natural Science Foundation of China (Grant Nos. 61502342, 11401462).
E-mail address: sblin1983@gmail.com.

http://dx.doi.org/10.1016/j.jco.2016.02.003
0885-064X/

as "training"). Therefore, the performance of the final estimate depends heavily on the quality of the candidates, which can be measured by the approximation capability of the selected class of functions.

The success of the radial basis function networks methodology in Euclidean space derives from its ability to generate approximants from data having arbitrary geometry. Thus, it is natural to introduce spherical radial basis function networks (SRBFNs) to tackle spherical data. This method has been extensively used in gravitational phenomenon [7], image processing [35] and learning theory [27]. An SRBFN can be mathematically represented as

$$S_n(x) := \sum_{i=1}^{n} c_i \phi(\xi_i \cdot x), \quad x \in \mathbf{S}^d \tag{1.1}$$

where $c_i \in \mathbf{R}$ is the connection weight, ϕ is the activation function, $\{\xi_i\}_{i=1}^n \subset \mathbf{S}^d$ is the set of centers, and $\mathbf{S}^d$ is the unit sphere in $\mathbf{R}^{d+1}$.

Obviously, the approximation capability of SRBFNs depends on the activation function and centers. A seminal paper concerning the activation function selection is [34], in which Sun and Cheney deduced the sufficient and necessary conditions of the activation function to guarantee the universal approximation property of corresponding SRBFNs. Consequently, Mhaskar et al. [25] deduced a Jackson-type inequality of SRBFN approximation under the condition that the Fourier–Legendre coefficients of the activation function are not trivial. Later, Le Gia et al. [11] provided an upper bound error estimate for least-square SRBFN approximation with positive definition activation function by using the topological relation between $\mathbf{S}^d$ and the $d+1$-dimensional unit ball $\mathbf{B}^{d+1}$. For more details on this topic, the readers are referred to [5,12,13,15,28,30,29].

Compared with the activation function selection, the center-selection strategy of SRBFNs is more important and difficult, since it determines the computational burden of the training process. To be detailed, if centers are specified before the training, then solving a simple linear optimization problem can deduce the final estimate. If centers need to be tuned in the process of training, then we should tackle a nonlinear optimization problem that usually requires more computation. Up till now, there are roughly three categories of center-selection strategies. The first one is the spherical basis function (SBF) method (or zonal function networks method) that uses the linear combination of kernels located at points in a given scattered data. It follows from the definition that this type of networks is a linear approximant. Thus, we can use a linear algorithm to get the globally optimal solution. In particular, it was pointed out in [11,27] that solutions to the regularized least squares and support vector machine algorithms are SBFs. The second one is the minimal energy method that focuses on selecting centers by minimizing some quantities concerning the energy of the points. Examples include the Riesz minimal energy [9], ϕ-Riesz minimum energy [33] and other low discrepancy energies [32]. It should be highlighted that whether the SRBFN whose centers are minimal energy points is a linear approximant depends on the definition of energy. If the energy is independent of the data, such as the Riesz minimal energy, then the corresponding SRBFN is a linear approximant. The last one aims to select centers via training, which naturally results two types of parameters, the connection weights and centers, and makes the corresponding SRBFN be a nonlinear approximant. The main advantages of this approach is that the nonlinear SRBFN sometimes leads to better approximation capability [14,26] and may circumvent the well known curse of dimensionality [1]. However, due to their nonlinearity, the implementation and training of the nonlinear SRBFN are much more difficult than its linear counterpart.

Our focus in this paper is not on selecting the most appropriate centers for a specified learning task, but on quantifying different approximation capabilities between linear and nonlinear SRBFNs. To this end, we should at first provide an answer to the following question: Is the approximation capability of the nonlinear SRBFN essentially better than that of linear SRBFN? Such a question is not new in the classical neural network approximation. For instance, in [17,18], Maiorov proved that the approximation capabilities of the nonlinear neural network and radial basis function network manifolds are better than that of arbitrary linear space with the same number of parameters, as the approximation errors of these nonlinear manifolds are essentially smaller than the Kolmogorov n-width [31]. A similar conclusion can also be found in [10]. The main novelty of this paper is to prove that similar conclusion is not valid for SRBFNs. In fact, we derive two lower bounds of nonlinear