Research Article

Interpolation and Best Approximation for Spherical Radial Basis Function Networks

Shaobo Lin, Jinshan Zeng, and Zongben Xu

Institute for Information and System Sciences, School of Mathematics and statistics, Xian Jiaotong University, Xian, Shaanxi 710049, China

Correspondence should be addressed to Zongben Xu; zbxu@mail.xjtu.edu.cn

Received 1 May 2013; Revised 17 September 2013; Accepted 18 September 2013

Academic Editor: Mieczysław Mastyło

Copyright © 2013 Shaobo Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Within the conventional framework of a native space structure, a smooth kernel generates a small native space, and radial basis functions stemming from the smooth kernel are intended to approximate only functions from this small native space. In this paper, we embed the smooth radial basis functions in a larger native space generated by a less smooth kernel and use them to interpolate the samples. Our result shows that there exists a linear combination of spherical radial basis functions that can both exactly interpolate samples generated by functions in the larger native space and near best approximate the target function.

1. Introduction

Many scientific questions boil down to synthesizing an unknown but definite function from finitely many samples \((x_i, y_i)_{i=1}^m\). The purpose is to find a functional model that can effectively represent or approximate the underlying relation between the input \(x_i\) and the output \(y_j\). If the unknown functions are defined on spherical domains, then data collected by satellites or ground stations are usually not restricted on any regular region and are scattered. Thus, any numerical method relying on the structure of a “grid” is doomed to fail.

The success of the radial basis function networks methodology in Euclidean space derives from its ability to generate estimators from data with essentially unstructured geometry. Therefore, it is natural to borrow this idea to deal with spherical scattered data. This method, called the spherical radial basis function networks (SRBFNs), has been extensively used in gravitational phenomenon [1, 2], image processing [3, 4], and learning theory [5].

Mathematically, the SRBFN can be represented as

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

where \(c_i \in \mathbb{R}\), \(\xi_i \in S^d\), and \(\phi : [-1, 1] \to \mathbb{R}\) are called the connection weight, center, and activation function in the terminology of neural networks, respectively. Here and hereafter, \(S^d\) denotes the unit sphere embedded in the \(d + 1\)-dimensional Euclidean space, \(\mathbb{R}^{d+1}\). Both the connection weights \(c_i\) and the centers \(\xi_i\) are adjustable in the process of training. We denote by \(\Phi_n\) the collection of functions formed as (1).

A basic and classical approach for training SRBFN is to construct exact interpolant based on the given samples \((x_i, y_i)_{i=1}^m\), that is, to find a function \(S_n\) in \(\Phi_n\) such that

\[
S_n(x_i) = y_i, \quad i = 1, \ldots, m. \tag{2}
\]

If the activation function \(\phi\) is chosen to be positive definite [2], then the matrix \(A := (\phi(x_i \cdot x_j))_{i,j=1}^m\) is nonsingular. Thus, the system of (2) can be easily solved by taking the scattered data as the centers of SRBFN. This method has been already named the spherical basis function (SBF) method. Under this circumstance, the connection weights are determined via

\[
c = A^{-1} y, \tag{3}
\]

where \(c = (c_1, c_2, \ldots, c_m)^T\), \(y = (y_1, \ldots, y_m)^T\), \(U^T\) denotes the transpose of the matrix (or vector) \(U\), and \(A^{-1}\) denotes the inverse matrix of \(A\).

For the SBF method, if the target function \(f\) belongs to the native space \(N_\phi\) associated with the activation function \(\phi\),
then the best SBF approximant of \( f \) is characterized by the exact interpolation:
\[
\left\| f - \sum_{i=1}^{m} \phi_i \right\|_{N_p} = \inf_{g \in \Phi_m} \left\| f - g \right\|_{N_p},
\]
(4)

where \( \{\phi_i\}_{i=1}^{m} \) are determined by (3). This property makes the SBF interpolation strategy popular in spherical scattered data fitting [6–14]. However, there are also two disadvantages for the SBF method. On one hand, since the centers of SRBFN interpolants are chosen as the scattered data, the interpolation capability depends heavily on their geometric distributions. This implies that we cannot obtain a satisfactory interpolation error estimate if the data are “badly” located on the sphere. On the other hand, the well known “native space barrier” [11,12,15] shows that (4) only holds for a small class of smooth functions if \( \phi \) is smooth. Therefore, for functions outside \( N_p \), the SBF interpolants are not guaranteed to be the best approximants.

Along the flavor of the previous papers [7,11,12], we use SRBFNs to interpolate functions in a large native space \( N_p \) associated with the kernel \( \psi \) which is less smooth than \( \phi \) and study its interpolation capability. Different from the previous work, the centers are chosen in advance to be quasi-uniform located on spheres, which makes the interpolation error depend on the number rather than the geometric distribution of centers. Our purpose is not to give the detailed error estimate of the SBF interpolaton. Instead, we focus on investigating the relation between interpolation and approximation for SRBFN. Indeed, we find that there exists an SRBFN interpolant which is also the near best and approximation for SRBFN. Indeed, we find that there exists an SRBFN interpolant which is also the near best approximant of functions in \( N_p \), where the number of centers and geometric distribution of the scattered data satisfy a certain assumption. That is, for a suitable choice of \( n \), there exists a function \( g_n \in \Phi_n \) such that

(1) \( g \) exactly interpolates the samples \( (x_j, y_j)_{j=1}^{m} \);

(2) \( \| f - g_n \|_{N_p} \leq C \inf_{g \in \Phi_m} \| f - g \|_{N_p} \),

where \( C \) is a constant depending only on \( d, \alpha \), and \( \beta \).

\[2.\text{ Positive Radial Basis Function on the Sphere}\]

It is easy to deduce that the volume of \( S^d \), \( \Omega_d \), satisfies
\[
\Omega_d := \int_{S^d} d\omega = \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)},
\]
(5)

where \( d\omega \) denotes the surface area element on \( S^d \). For integer \( k \geq 0 \), the restriction to \( S^d \) of a homogeneous harmonic polynomial of degree \( k \) on the unit sphere is called a spherical harmonic of degree \( k \). The span of all spherical harmonics of degree \( k \) is denoted by \( H^d_k \), and the class of all spherical harmonics (or spherical polynomials) of degree \( k \leq s \) is denoted by \( \Pi^d_s \). It is obvious that \( \Pi^d_s = \oplus_{k=0}^s H^d_k \). The dimension of \( H^d_k \) is given by
\[
D_k^d := \dim H^d_k = \begin{cases} 2k + d - 1 & k \geq 1; \\ k & k = 0, \end{cases}
\]
(6)

and that of \( \Pi^d_s \) is
\[
\sum_{k=0}^s D_k^d = D_{s+1}^d \sim s^d.
\]

Denote by \( \{Y_{k,j}\}_{j=1}^{D_k^d} \) an orthonormal basis of \( H^d_k \); then the following addition formula [16,17] holds
\[
\sum_{i=1}^{D_k^d} Y_{k,i}(x) Y_{k,i}(y) = \frac{D_k^d}{\Omega_d} P_k^{d+1}(x \cdot y),
\]
(7)

where \( P_k^{d+1} \) is the Legendre polynomial with degree \( k \) and dimension \( d + 1 \). The Legendre polynomial \( P_k^{d+1} \) can be normalized such that \( P_k^{d+1}(1) = 1 \) and satisfies the orthogonality relation
\[
\int_{-1}^{1} p_k^{d+1}(t) p_j^{d+1}(t) (1 - t^2)^{(d-2)/2} dt = \frac{\Omega_d}{\Omega_{d-1} D_k^d} \delta_{k,j},
\]
(8)

where \( \delta_{k,j} \) is the usual Kronecker symbol.

Positive definite radial basis functions on spheres were introduced and characterized by Schoenberg [18]. Namely, a radial basis function \( \varphi \) is positive definite if and only if its expansion
\[
\varphi(x \cdot y) = \sum_{k=0}^{\infty} e_k \frac{D_k^d}{\Omega_d} \phi_k^{d+1}(x \cdot y)
\]
(9)

has all Fourier-Legendre coefficients \( e_k \geq 0 \) and \( \sum_{k=0}^{\infty} e_k (D_k^d / \Omega_d) < \infty \). We define the native space \( N_p \) as
\[
N_p := \left\{ f(x) = \sum_{k=0}^{\infty} \sum_{j=1}^{D_k^d} \hat{f}_{k,j} Y_{k,j}(x) : \sum_{k=0}^{\infty} e_k \sum_{j=1}^{D_k^d} \hat{f}_{k,j}^2 < \infty \right\},
\]
(10)

with its inner product
\[
\langle f, g \rangle_{N_p} := \sum_{k=0}^{\infty} e_k \sum_{j=1}^{D_k^d} \hat{f}_{k,j} \hat{g}_{k,j},
\]
(11)

\[3.\text{ Interpolation and Near Best Approximation}\]

Let \( \Lambda := \{\eta_i\}_{i=1}^{n} \subset S^d \) be a set of points and \( d(x, y) = \arccos x \cdot y \) be the spherical distance between \( x \) and \( y \). We denote by \( h_{\Lambda} := \max_{x \in \Lambda} \min_{\eta_i \in \Lambda} d(x, \eta_i) \), \( d_{\Lambda} := (1/2) \min_{x \in \Lambda} \max_{k \leq s} (\delta_{\eta_i, \eta_j})_k \), and \( r_{\Lambda} := h_{\Lambda} / h_{\Lambda} \) the mesh norm, separation radius, and mesh ratio of \( \Lambda \), respectively. It is easy to check that these three quantities describe the geometric distribution of points in \( \Lambda \).
The $\tau$-uniform set $F_\tau := F_\tau(S^d)$ is defined by the family of all centers sets $\Xi$ with $\tau \Lambda \cdots$ (21).

**Lemma 3.** and [19, Example 2.10], respectively. which can be found in [12, Proposition 5.2] and activation function $\phi$.

**Remark 2.** Similar results have been considered for spherical polynomials both in $C(S^d)$ and $N_\psi$. Narcowich et al. [11, 12] proved that there exists a spherical polynomial interpolant of degree at most $L \geq C_{S^d}$ which can also best approximate the target both in $C(S^d)$ and $N_\psi$.

To prove Theorem 1, we need the following three lemmas, which can be found in [12, Proposition 5.2], [12, Theorem 5.5], and [19, Example 2.10], respectively.

**Lemma 3.** Let $\mathcal{Y}$ be a (possibly complex) Banach space, $\mathcal{Y}$ a subspace of $\mathcal{Y}$, and $Z^*$ a finite-dimensional subspace of $\mathcal{Y}^*$, the dual of $\mathcal{Y}$. If for every $z^* \in Z^*$ and some $\gamma > 1$, $y$ independent of $z^*$,

$$\|z^*\|_{\mathcal{Y}^*} \leq \gamma\|z^*\|_{\mathcal{Y}^*}. \tag{15}$$

then for $y \in \mathcal{Y}$ there exists $v \in \mathcal{Y}$ such that $v$ interpolates $y$ on $Z^*$, that is, $z^*(y) = z^*(v)$ for all $z^* \in Z^*$. In addition, $v$ approximates $y$ in the sense that $\|y - v\|_{\mathcal{Y}} \leq (1 + 2\gamma) \text{dist}_{\mathcal{Y}}(y, \mathcal{Y}).$
Thus, Lemma 3 with $\gamma = 2$ yields that in order to prove (18), it suffices to prove
\[
\left\| \sum_{i=1}^{m} \xi_i \psi_{x_i} - s \right\|_{N_p} \leq \frac{\sqrt{3}}{2} . \tag{22}
\]

Let $L \in \mathbb{N}$, $P \in \Pi_{L}$ be the best polynomial approximation of \( \sum_{i=1}^{m} \xi_i \psi_{x_i} \) in the metric of $N_p$; then for arbitrary $s' \in \Phi_{n}$, we have
\[
\left\| \sum_{i=1}^{m} \xi_i \psi_{x_i} - s' \right\|_{N_p} \leq \left\| \sum_{i=1}^{m} \xi_i \psi_{x_i} - P \right\|_{N_p} + \left\| P - s' \right\|_{N_p}. \tag{23}
\]

It follows from [12, Page 382] that there exists a constant, $\kappa$, depending only on $\beta$ such that for arbitrary $L \geq \kappa q_X^{-1}$, there holds
\[
\left\| \sum_{i=1}^{m} \xi_i \psi_{x_i} - P \right\|_{N_p} \leq \frac{\sqrt{3}}{4} \sum_{i=1}^{m} \left\| \xi_i \psi_{x_i} \right\|_{N_p}. \tag{24}
\]

Let $s'$ be the best $\Phi_n$ approximation of $P$ in the metric of $N_p$. Then it follows from Lemma 4 and the well-known Bernstein inequality [17] that
\[
\left\| P - s' \right\|_{N_p} \leq C h_n^\alpha \left\| P \right\|_{L^2} \leq C h_n^\alpha \beta L^\beta \left\| P \right\|_{L^2}. \tag{25}
\]

Since $P$ is the best polynomial approximation of $\sum_{i=1}^{m} \xi_i \psi_{x_i}$ in the metric of $N_p$, a simple computation yields
\[
\left\| P \right\|_{L^2} = \sum_{i=1}^{m} \xi_i^2. \tag{26}
\]

Furthermore, it follows from Lemma 5 that
\[
\left\| \sum_{i=1}^{m} \xi_i \psi_{x_i} \right\|_{N_p}^2 \geq C d_X^2 \sum_{i=1}^{m} \xi_i^2. \tag{27}
\]

Then, $\Xi_n \in F$, together with $L = \max\{|\kappa q_X^{-1}| + 1, [3^{\beta/2}/(4\xi^2)]\}$ yields that
\[
\left\| P - s' \right\|_{N_p} \leq C h_n^\alpha \beta L^\beta \left\| \sum_{i=1}^{m} \xi_i \psi_{x_i} \right\|_{N_p} \leq C \left( n^{-1/d} q_X \right)^{\alpha-\beta} \left\| \sum_{i=1}^{m} \xi_i \psi_{x_i} \right\|_{N_p}, \tag{28}
\]

where $C$ is a constant depending only on $\beta, d$, and $r$. Thus, there exists a constant $C_r$ depending only on $\alpha, \beta, d$, and $r$ such that for arbitrary $n \geq C_r q_X^{-d}$, there holds
\[
\left\| P - s' \right\|_{N_p} \leq \frac{\sqrt{3}}{4} \sum_{i=1}^{m} \left\| \xi_i \psi_{x_i} \right\|_{N_p}. \tag{29}
\]

Inserting (24) and (29) into (23), we finish the proof of (22) and then complete the proof of Theorem 1. \qed

Acknowledgments

An anonymous referee has carefully read the paper and has provided to us numerous constructive suggestions. As a result, the overall quality of the paper has been noticeably enhanced, to which we feel much indebted and are grateful. The research was supported by the National 973 Programing (2013CB329404), the Key Program of National Natural Science Foundation of China (Grant no. 11131006), and the National Natural Science Foundations of China (Grant no. 61075054).

References


Submit your manuscripts at http://www.hindawi.com