



Contents lists available at ScienceDirect

Journal of Complexity

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



Complexity of Gaussian-radial-basis networks approximating smooth functions

Paul C. Kainen^a, Věra Kůrková^b, Marcello Sanguineti^{c,*}

^a Department of Mathematics, Georgetown University, Washington, DC, 20057-1233, USA

^b Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 2, 182 07, Prague 8, Czech Republic

^c Department of Communications, Computer, and System Sciences (DIST), University of Genoa, Via Opera Pia 13, 16145 Genova, Italy

ARTICLE INFO

Article history:

Received 4 November 2007

Accepted 20 August 2008

Available online 7 September 2008

Keywords:

Gaussian-radial-basis-function networks

Rates of approximation

Model complexity

Variation norms

Bessel and Sobolev norms

Tractability of approximation

ABSTRACT

Complexity of Gaussian-radial-basis-function networks, with varying widths, is investigated. Upper bounds on rates of decrease of approximation errors with increasing number of hidden units are derived. Bounds are in terms of norms measuring smoothness (Bessel and Sobolev norms) multiplied by explicitly given functions $a(r, d)$ of the number of variables d and degree of smoothness r . Estimates are proven using suitable integral representations in the form of networks with continua of hidden units computing scaled Gaussians and translated Bessel potentials. Consequences on tractability of approximation by Gaussian-radial-basis function networks are discussed.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Radial-basis function (RBF) networks with Gaussian computational units are known to be able to approximate with an arbitrary accuracy all continuous and all $\mathcal{L}^2$ -functions on compact subsets of $\mathbb{R}^d$ [10,23,27,30,31]. In such approximations, the number n of RBF units plays the role of a measure of model complexity and its size determines the feasibility of network implementation.

Several authors have investigated rates of approximation by RBF networks with n Gaussians units of *fixed width*. Girosi and Anzellotti [9] derived an asymptotic upper bound of order $n^{-1/2}$ on approximation error measured by the supremum norm for band-limited functions with continuous derivatives up to order r with $r > d/2$, where d is the number of variables [9, p. 106]. Using results

* Corresponding author.

E-mail addresses: kainen@georgetown.edu (P.C. Kainen), vera@cs.cas.cz (V. Kůrková), marcello@dist.unige.it (M. Sanguineti).

from statistical learning theory, Girosi [8] extended these bounds to more general classes of kernels. For Gaussians of varying widths, Kon, Raphael, and Williams [14, Corollary 3] obtained bounds on a weighted $\mathcal{L}^\infty$ -distance from the target function to a linear combination of Gaussians.

Some bounds improve on the exponent of $-1/2$. Mhaskar [24,25] and Narcovich et al. [29] obtained bounds of order $n^{-r/2d}$, and in one special case, Maiorov [21] found $n^{-r/(d-1)}$. Although order with respect to n improves, the remaining multiplicative factors in such bounds involve constants that are unknown and these upper bounds increase as d increases. Also, they apply to different classes of target and approximating functions. Moreover, dependence on parameters may differ, and approximation error is computed with respect to different metrics. Thus, it is not easy to compare these bounds.

In this paper, we approximate smooth functions by Gaussian RBF networks with units of varying widths, using $\mathcal{L}^2$ -distance with respect to the Lebesgue measure. We derive upper bounds on rates of approximation in terms of the Bessel and Sobolev norms of the functions to be approximated. Bessel norms are defined in terms of convolutions with Bessel-potential kernels, while Sobolev norms use integrals of partial derivatives. The Bessel norm $\|\cdot\|_{L^{2,r}}$ and the Sobolev norm $\|\cdot\|_{W^{2,r}}$ are equivalent but the ratio between them depends on the number d of variables.

Our estimates hold for all numbers n of hidden units and all degrees $r > d/2$ of Bessel potentials. The estimates are of the form $n^{-1/2}$ times the Bessel norm $\|f\|_{L^{1,r}}$ of the function f to be approximated times a factor $k(r, d)$. For a fixed $c > 0$ and the degree $r_d = d/2 + c$, the factor $k(r_d, d)$ decreases to zero exponentially fast. We also derive estimates in terms of $\mathcal{L}^2$ Bessel and Sobolev norms. Our results show that reasonably smooth functions can be approximated quite efficiently by Gaussian-radial-basis networks. A preliminary version of the results appeared in [13].

The paper is organized as follows. Section 2 presents some concepts, notations, and auxiliary results for studying approximation by Gaussian RBF networks. Section 3 derives upper bounds on rates of approximation of Bessel potentials by linear combinations of scaled Gaussians in terms of variation norms obtained from integral representations of Bessel potentials and their Fourier transforms. In Section 4, for functions representable as convolutions with Bessel potentials, upper bounds are derived in terms of Bessel-potential norms. These bounds are then combined with estimates of variational norms from the previous section to obtain bounds for approximation by Gaussian RBFs in terms of Bessel norms. Section 5 uses the relationship between Sobolev and Bessel norms to obtain bounds in terms of Sobolev norms. In Section 6, we discuss consequences for tractability of multivariate approximation by Gaussian-radial-basis networks.

2. Approximation by Gaussian RBF networks

For $\Omega \subseteq \mathbb{R}^d$, $\mathcal{L}^2(\Omega)$ denotes the space of real-valued functions on Ω with norm $\|f\|_{\mathcal{L}^2(\Omega)} = (\int_\Omega |f(x)|^2 dx)^{1/2}$. Two functions are identified if they differ only on a set of Lebesgue-measure zero. When $\Omega = \mathbb{R}^d$, we omit it in the notation.

For nonzero $f \in \mathcal{L}^2$, $f^0 = f/\|f\|_{\mathcal{L}^2}$ denotes the *normalization* of f ; for convenience, we put $0^0 = 0$. For $F \subset \mathcal{L}^2$, $F|_\Omega$ denotes the set of functions from F restricted to Ω , $\hat{F}$ the set of Fourier transforms of functions in F , and F^0 the set of their normalizations. For $n \geq 1$, define

$$\text{span}_n F := \left\{ \sum_{i=1}^n w_i f_i \mid f_i \in F, w_i \in \mathbb{R} \right\}.$$

In this paper, we investigate accuracy measured by $\mathcal{L}^2$ -norm with respect to the Lebesgue measure λ in approximation by Gaussian-radial-basis-function networks.

A *Gaussian-radial-basis-function unit* with d inputs computes all *scaled and translated Gaussian functions* on $\mathbb{R}^d$. For $b > 0$, let $\gamma_b : \mathbb{R}^d \rightarrow \mathbb{R}$ denote the Gaussian function of *width* b defined by

$$\gamma_b(x) = e^{-b\|x\|^2}.$$

A simple calculation shows that

$$\|\gamma_b\|_{\mathcal{L}^2} = (\pi/2b)^{d/4}. \quad (1)$$

Download English Version:

<https://daneshyari.com/en/article/4609031>

Download Persian Version:

<https://daneshyari.com/article/4609031>

[Daneshyari.com](https://daneshyari.com)