Applied Mathematics and Computation 330 (2018) 170–184

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Sparse radial basis function approximation with spatially variable shape parameters

Valentin Stolbunov, Prasanth B. Nair*

Institute for Aerospace Studies, University of Toronto, Toronto, ON M3H 5T6, Canada

ARTICLE INFO

Keywords:
Function approximation
Parameterized dictionary learning
Radial basis functions
Greedy algorithm
Shape parameter tuning
Surrogate modeling

ABSTRACT

We present an efficient greedy algorithm for constructing sparse radial basis function (RBF) approximations with spatially variable shape parameters. The central idea is to incrementally construct a sparse approximation by greedily selecting a subset of basis functions from a parameterized dictionary consisting of RBFs centered at all of the training points. An incremental thin QR update scheme based on the Gram–Schmidt process with reorthogonalization is employed to efficiently update the weights of the sparse RBF approximation at each iteration. In addition, the shape parameter of the basis function chosen at each iteration is tuned by minimizing the ℓ_2-norm of the training residual, while an approximate leave-one-out error metric is used as the dominant stopping criterion. Numerical studies are presented for a range of test functions to demonstrate that the proposed algorithm enables the efficient construction of RBF approximations with spatially variable shape parameters. It is shown that, compared to a classical RBF model with a single tunable shape parameter and Gaussian process models with an anisotropic Gaussian covariance function, the proposed algorithm can provide significant improvements in accuracy, cost, and sparsity, particularly for high-dimensional datasets.

1. Introduction

The problem of constructing an approximation model using a set of scattered observational data arises in many application areas, including machine learning [1,2], forecasting and control of dynamical systems [3], numerical methods for partial differential equations [4,5], design optimization [6,7], inverse problems [8], stochastic modeling [9] and computer graphics [10]. Approximation models constructed using an observational or training dataset are also commonly referred to in the literature as surrogate models, emulators, metamodels, or interpolants. The field of engineering design optimization in particular has benefited greatly from approximation models and motivates many of the new research directions [11].

The task of constructing a surrogate/metamodel/emulator can be viewed as approximation through "learning" – using an available dataset to "teach" a model how to predict output values for various input values by tuning model-specific parameters. A wide range of numerical methods can be found in the literature for modeling input–output datasets, for example, radial basis functions [12], Gaussian process models [13], support vector machines [1] and neural networks [2]. Research in the field is mostly concerned with developing modeling techniques that can scale well to large, high-dimensional datasets

* Corresponding author.
E-mail addresses: valentin.stolbunov@mail.utoronto.ca (V. Stolbunov), pbn@utias.utoronto.ca (P.B. Nair).

https://doi.org/10.1016/j.amc.2018.02.001
0096-3003/

and/or improve the generalization ability of the approximation [14]; see, for example, [15] for a comparative review of various techniques.

Many of the popular nonparametric approximation techniques involve the use of expansions in radial basis functions (RBFs); see, for example, [1,2,13,16]. The use of RBF expansions can be theoretically justified by the representer theorem in approximation theory [17]. RBF models are widely used in practical applications since they can be used in any arbitrary space dimension, work for scattered datasets, and allow the construction of interpolants with a specified degree of smoothness [18]. Hardy [19] provided a detailed discussion of RBFs used for models in fields such as geophysics and signal processing, while Poggio and Girosi [20] presented the relationship between RBF models and regularization networks in the field of statistical learning theory. Even though RBF models have a number of attractive features, this approach often results in large-scale ill-conditioned linear systems that can be difficult to solve efficiently. This has motivated the development of numerical schemes for tackling the ill-conditioning issue and reducing the computational complexity associated with large-scale datasets [10,18].

The present work is concerned with the development of efficient numerical schemes for constructing RBF models of moderate to large-scale datasets such as those arising from design of computer experiments. More specifically, we propose a sequential greedy algorithm for efficiently constructing RBF approximations with spatially variable shape parameters. The idea of using greedy algorithms [21] to improve the efficiency of RBF approximation is not new and has been successfully applied to speedup RBF based methods for function approximation and pattern recognition; see, for example, [22–24]. The orthogonal matching pursuit (OMP) algorithm [25] approximated functions recursively using a dictionary of wavelets and later Natarajan [26] improved on OMP by introducing order recursion (ORMP). Floater and Iske [27] applied a greedy algorithm to the problem of generating evenly distributed subsets of scattered data. Greedy algorithms have also been applied to RBF collocation methods for partial differential equations [28,29].

The proposed greedy algorithm is an incremental forward greedy strategy wherein a new basis function is appended to the RBF approximation at each iteration. A key difference between the proposed greedy algorithm and those found in the literature lies in the fact that we work with a parameterized dictionary of basis functions since our goal is to construct an RBF expansion with spatially variable shape parameters. This approach provides a number of significant advantages over existing greedy algorithms for RBF approximation. Firstly, this feature allows significant improvements in computational efficiency since k-fold cross-validation tests involving repeated applications of the algorithm to partitions of the training dataset are no longer needed to tune the shape parameters. Secondly, since we use a separate shape parameter for each basis function, our approach provides exceptional modeling flexibility and is better able to approximate unstructured datasets. A similar idea was used in the greedy algorithm of Schaback and Wendland [22], wherein a sequence of RBF models are constructed via repeated residual fitting in an inner–outer loop and each model in the sequence is permitted to use a different value of the shape parameter. This approach, while exceptionally efficient, suffers from slow training error convergence. Another distinguishing feature of the proposed algorithm is the use of an incremental thin QR factorization to update the RBF approximation. This strategy ensures numerical stability while providing significant reductions in computational complexity and memory requirements compared to greedy algorithms that require the full Gram matrix to be computed *a priori* and stored [26,30]. Additional benefits of the thin QR factorization update scheme include the ability to efficiently compute the weights of all the basis functions selected so far at each iteration and the reciprocal condition number of the coefficient matrix.

The remainder of this paper is organized as follows. In the next section, a more detailed discussion of the motivation behind the proposed algorithm is provided. This includes background information on classical RBF approximations and their various drawbacks. The proposed algorithm is then presented in Section 3 to address these drawbacks and highlight a number of additional benefits. Numerical studies are presented in Section 4 and the main conclusions are highlighted in Section 5.

2. Classical RBF approximation

Function approximation problems begin with a dataset of m observations $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\} \subset \Omega$ and their associated real function values or experiment results $y_i, i = 1, \ldots, m$. Typically Ω is a bounded domain in $\mathbb{R}^n$, while the output values are usually a result of evaluating a computationally-expensive function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ or conducting a costly experiment. An algorithm is then responsible for constructing $\hat{f}$ – a cheap approximation of f – which can be used to predict y at any new $\mathbf{x}$ as $\hat{y} = \hat{f}(\mathbf{x})$.

An RBF interpolation model in terms of positive definite basis functions (also known as Mercer kernels [1]) can be written in the form

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^{m} \alpha_i \phi_i(\mathbf{x}), \tag{1}$$

where $\phi_i(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is an RBF, and the undetermined weights $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}^T \in \mathbb{R}^m$ can be calculated by solving the following linear algebraic system of equations

$$\mathbf{K}\boldsymbol{\alpha} = \mathbf{y}, \tag{2}$$