Journal of Computational Physics 354 (2018) 370–386

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Sequential function approximation on arbitrarily distributed point sets ☆

Kailiang Wu, Dongbin Xiu *

Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 17 April 2017
Received in revised form 12 September 2017
Accepted 16 October 2017
Available online xxxx

Keywords:
Approximation theory
Sequential approximation
Randomized algorithm

ABSTRACT

We present a randomized iterative method for approximating unknown function sequentially on arbitrary point set. The method is based on a recently developed sequential approximation (SA) method, which approximates a target function using one data point at each step and avoids matrix operations. The focus of this paper is on data sets with highly irregular distribution of the points. We present a nearest neighbor replacement (NNR) algorithm, which allows one to sample the irregular data sets in a near optimal manner. We provide mathematical justification and error estimates for the NNR algorithm. Extensive numerical examples are also presented to demonstrate that the NNR algorithm can deliver satisfactory convergence for the SA method on data sets with high irregularity in their point distributions.

1. Introduction

Approximating an unknown function using its samples is a standard task in approximation theory and numerical analysis. It has been extensively studied, see, for example, a few earlier books [3,4,12,13]. The common approaches include methods such as least squares, interpolation, compressed sensing, etc. Typically, these methods require operation and inversion of a certain model matrix of size $M \times N$, where M is the number of samples and N is the number of basis functions used in the approximation. For problems with very large number of samples, implementation of these methods can become challenging, as the size of the model matrix becomes exceedingly large.

In this paper, we focus on a different approach – sequential approximation (SA), based on the recent work of [15,19]. The method seeks to construct an approximation iteratively using one data point at each step. By doing so, it avoids the use of matrices and only requires vector operations. Therefore, the implementation of the method is straightforward. It was shown that the method converges, in expectation, when one uses randomly selected samples in the approximation domain. The theoretical convergence, particularly the optimal convergence using certain optimal sampling measure, was established in [15] for noiseless data and in [14] for noisy data. The method is particularly suitable for exceedingly large data sets, as it does not need to store the entire data set. In [19], the method was used in conjunction with tensorized Gauss quadrature points for multivariate function approximation, which is not feasible in high dimensions by other methods. It should also be remarked that the SA method was motivated by the randomized Kaczmarz (RK) method ([16]), which iteratively solves

☆ This work was partially supported by AFOSR FA95501410022 and NSF DMS 1418771.
* Corresponding author.
E-mail addresses: wu.3423@osu.edu (K. Wu), xiu.16@osu.edu (D. Xiu).

https://doi.org/10.1016/j.jcp.2017.10.020
0021-9991/

overdetermined linear system of equations one random row at a step. In term of linear systems of equations, RK method has been studied and extended in a series of work, cf., [10,20,2,5,9,1,17].

The focus of this paper is to study SA method for function approximation subject to arbitrarily given data sets. In [15, 19,14], the convergence of SA method was established when data samples are collected randomly and an optimal sampling measure to ensure faster convergence was derived. In practice, however, sample data are often collected in a vastly different manner and do not follow certain desirable patterns. Practical constraints in experimentation and/or measurements often prevent the samples from being collected at "mathematically optimal" locations. Moreover, samples can be highly clustered and distribute irregularly in the domain. Function approximation using such irregularly distributed samples can be challenging. In this paper we develop a variation of SA method to cope with irregularly distributed data sets. The method starts with random sampling using a probability measure, preferably the optimal measure developed in [15]. The resulting samples in general do not coincide with the existing points in the given data set. We then conduct a nearest neighbor search and replace the randomly generated sample by its nearest neighbor in the given data set. The SA method is then performed using the selected data sample. The procedure, termed "nearest neighbor replacement" (NNR) algorithm, effectively creates a discrete probability measure to sample the given data set. We demonstrate that this discrete measure is a weak approximation to the continuous measure from which samples are drawn in the first place. Theoretical analysis is provided to quantify the error in this weak approximation of the measure, and the convergence of the NNR-SA algorithm is established. Although the nearest neighbor search adds computational cost to the SA method, its implementation requires only vector operations and a simple linear search. It significantly improves the convergence of SA methods, when the samples are highly clustered. The NNR-SA algorithm is thus suitable for highly irregularly distributed data sets. Several numerical examples are provided to demonstrate its effectiveness.

2. Preliminaries

We seek to approximate an unknown target function $f : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^d$, $d \geq 1$, and $\mathbf{x} = (x_1, \ldots, x_d)$ the coordinate. We shall equip the domain D with a measure ω, which is assumed to be absolutely continuous and (without loss of generality) satisfy $\int_D d\omega = 1$. We then consider the standard Hilbert space with inner product

$$(g, h)_{L^2_\omega} := \int_D g(\mathbf{x}) h(\mathbf{x}) d\omega(\mathbf{x}),$$

and the corresponding induced norm $\|\cdot\|_{L^2_\omega}$. Throughout this paper we will assume $f \in L^2_\omega(D)$.

2.1. Setup

We shall approximate the target function f via its samples. Let

$$\Theta := \{\mathbf{x}_i\}_{i=1}^M \subset D \tag{2.1}$$

be a set of sample points. Throughout this paper we shall focus on the case of large data sets, i.e., $M \gg 1$. We also assume the sample points are arbitrarily distributed, in the sense that the configuration of the points in the set Θ does not follow any predetermined pattern such as tensor structure, quadrature/cubature rules, certain probability distributions, etc. Let

$$f_i = f(\mathbf{x}_i) + \epsilon_i, \qquad i = 1, \cdots, M, \tag{2.2}$$

be sample data, where $\{\epsilon_i\}_{i=1}^M$ are independent random noises with zero mean value and bounded variance. The case of $\epsilon_i \equiv 0$ corresponds to noiseless case.

We seek an approximation $p \approx f$ from a finite dimensional linear subspace V, i.e., $p \in V$, where $V \subset L^2_\omega(D)$ with $\dim V = N \geq 1$. Let $\{\psi_j(\mathbf{x})\}_{j=1}^N$ be a basis of V. Without loss of generality, we assume the basis functions are orthogonal with respect to the measure ω, i.e.,

$$\left(\psi_i, \psi_j\right)_{L^2_\omega} = \delta_{ij}, \qquad 1 \leq i, j \leq N. \tag{2.3}$$

(Note that non-orthogonal basis can always be orthogonalized by, for example, the Gram–Schmidt procedure.) The approximation can then be expressed as

$$p(\mathbf{x}) = \sum_{j=1}^N c_j \psi_j(\mathbf{x}). \tag{2.4}$$

Upon adopting the following vector notations

$$\mathbf{\Psi}(\mathbf{x}) = \left(\psi_1(\mathbf{x}), \ldots, \psi_N(\mathbf{x})\right)^\top, \qquad \mathbf{c} = (c_1, \ldots, c_N)^\top, \tag{2.5}$$

we write