Journal of Complexity 34 (2016) 129–151

Contents lists available at ScienceDirect

Journal of Complexity

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

Optimal sampling points in reproducing kernel Hilbert spaces☆

Rui Wang [a], Haizhang Zhang [b,c,*]

[a] School of Mathematics, Jilin University, Changchun 130012, PR China

[b] School of Data and Computer Science, Sun Yat-Sen University, Guangzhou 510275, PR China

[c] Guangdong Province Key Laboratory of Computational Science, Sun Yat-Sen University, Guangzhou 510275, PR China

ARTICLE INFO

Article history:
Received 10 November 2015
Accepted 20 November 2015
Available online 28 November 2015

Keywords:
Sampling points
Optimal distribution
Reproducing kernels
The Karhunen–Loève transform

ABSTRACT

The recent development of compressed sensing seeks to extract information from as few samples as possible. In such applications, since the number of samples is restricted, one should deploy the sampling points wisely. We are motivated to study the optimal distribution of finite sampling points in reproducing kernel Hilbert spaces, the natural background function spaces for sampling. Formulation under the framework of optimal reconstruction yields a minimization problem. In the discrete measure case, we estimate the distance between the optimal subspace resulting from a general Karhunen–Loève transform and the kernel space to obtain another algorithm that is computationally favorable. Numerical experiments are then presented to illustrate the effectiveness of the algorithms for the searching of optimal sampling points.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Functions describing natural phenomenon or social activities need to be converted into discrete data that can be handled by modern computers. From this viewpoint, sampling is the foundation

☆ Supported by Guangdong Provincial Government of China through the "Computational Science Innovative Research Team" program.

* Corresponding author.
E-mail addresses: rwang11@jlu.edu.cn (R. Wang), zhhaizh2@sysu.edu.cn (H. Zhang).

http://dx.doi.org/10.1016/j.jco.2015.11.010
0885-064X/© 2015 Elsevier Inc. All rights reserved.

for information processing and communication. The subject originated from the celebrated Shannon sampling theorem [26], which guarantees the complete reconstruction of a band-limited function from its values on some equally-spaced points. The elegant result motivates many follow-up studies, making sampling an important research subject in applied mathematics. We shall give a brief and partial introduction to the history and progresses.

Mathematically, sampling means to evaluate a function. To ensure the stability, it is arguable that sampling should only take place in function spaces where point evaluations are continuous. Such spaces when endowed with an inner product structure arise in many other areas of mathematics. They are termed as the reproducing kernel Hilbert spaces (RKHS), as by the Riesz's lemma there exists a function that is able to reproduce the function values through the inner product. In Shannon's theorem, the space of functions that are band-limited to $[-\pi, \pi]$ and are equipped with the inner product of $L^2(\mathbb{R})$ is an RKHS with the sinc function as its reproducing kernel. This interpretation gives the hope of searching for Shannon-type complete reconstruction formula for other RKHSs. It was found in [18] that as long as one has a frame or a Riesz basis formed by the reproducing kernel, then a Shannon-type sampling formula is immediately available by the general theory of frames. They showed that many past sampling formulae can be obtained in this manner. Recently, the approach has been generalized to reproducing kernel Banach spaces [34] by frames for Banach spaces via semi-inner-products [35].

Shannon type formulae enable us to have lossless representation of a function that is usually defined on an uncountable continuous domain using countable data. Going from uncountable to countable is a remarkable progress. However, countable is still infinite and computers cannot store or handle infinitely many data. This raises the question of how to reconstruct a function from its finite samples. For the crucial band-limited functions, two modified Shannon series have been proposed in the literature [10,23,31], where it was shown that over-sampling can lead to exponentially decaying approximation errors. The sampling process often comes with some cost. When sampled data are available, one is inclined to use as accurate reconstruction methods as possible. It has long been known that in the maximum sense, the best way of reconstruction in an RKHS is via the minimal norm interpolation [14]. The approximation error for optimal reconstruction from over-sampling in the Paley–Wiener space of band-limited functions was estimated in [16].

In this paper, we focus on another important question in sampling. Usually the number of sampling points in a practical application is limited. When that number is fixed, we ask what is the best strategy of deploying the sampling points, under the condition that the best reconstruction method is engaged. The study is also motivated by the recent development in basis pursuit [5] and compressed sensing [4], which seek to extract information from as few samples as possible. Since the number of samples is restricted, one should of course try to distribute the sampling points wisely. We shall stress in the next section that reproducing kernel Hilbert spaces should be the ideal background space for sampling. Thus, we will work under the assumption that the functions under sampling are from a given reproducing kernel Hilbert space associated with a fixed reproducing kernel. In other words, with a fixed reproducing kernel we are trying to find the optimal finite sampling points. We remark that this is different from the approach in machine learning of optimizing the reproducing kernel for regularization problems (see, for example, [13,19]), where the sampling points are fixed instead.

Optimal approximation of a function f from a general RKHS by the algorithms using only finitely many linear functional evaluations of f has been widely studied in the area of information-based complexity [9,11,20–22,27,29,30]. The optimal information is usually considered in two classes of linear functionals. The class of linear information considers all continuous linear functionals while the class of standard information only considers point evaluations. The errors of the algorithms are measured in three settings, including the worst case, the average case and the randomized case settings. The corresponding nth minimal error is defined as the smallest error among all algorithms that use at most n functional evaluations. Most of the researches focused on characterizing the nth minimal errors and the optimal order of convergence for both the linear information and the standard information in various settings [9,11,29,30]. The optimal approximation by the linear information has been well understood. Specifically, the nth minimal errors for this case can be characterized by the eigenvalues of some integral operator and the corresponding eigenfunctions can provide us the desired optimal algorithm. However, it is usually difficult to analyze the optimal standard

Download English Version:

https://daneshyari.com/en/article/4608540

Download Persian Version:

https://daneshyari.com/article/4608540

Daneshyari.com