Journal of Complexity 28 (2012) 59–75

Contents lists available at SciVerse ScienceDirect

Journal of Complexity

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

Tractability results for weighted Banach spaces of smooth functions

Markus Weimar

Mathematisches Institut, Universität Jena, Ernst-Abbe-Platz 2, 07740 Jena, Germany

ARTICLE INFO

Article history:
Received 9 December 2010
Accepted 10 March 2011
Available online 8 April 2011

Keywords:
Tractability
Curse of dimensionality
L_∞-approximation
Worst case error
Banach spaces
Approximation of smooth functions

ABSTRACT

We study the L_∞-approximation problem for weighted Banach spaces of smooth d-variate functions, where d can be arbitrarily large. We consider the worst case error for algorithms that use finitely many pieces of information from different classes. Adaptive algorithms are also allowed. For a scale of Banach spaces we prove necessary and sufficient conditions for tractability in the case of product weights. Furthermore, we show the equivalence of weak tractability with the fact that the problem does not suffer from the curse of dimensionality.

1. Introduction

The so-called *curse of dimensionality* can often be observed for multivariate approximation problems. That is, the minimal number of information operations needed to compute an ε-approximation of a d-variate problem depends exponentially on the dimension d. The phrase *curse of dimensionality* was already coined by Bellman in 1957. Since the late 1980s there has been a considerable interest in finding optimal algorithms, also concerning the optimal dependence on d and a theory called *information-based complexity* (IBC) has been created; see, e.g., [10]. Since there are different ways to measure the lack of exponential behavior, several kinds of tractabilities were introduced. A brief history of the studies of multivariate problems, as well as general tractability results and many concrete examples can be found in, e.g., [5,6,8].

In this paper we especially consider the L_∞-approximation problem defined on some Banach spaces $\mathcal{F}_d$ of real-valued d-variate functions. In Section 2 we formulate the problem exactly and recall the usual error definitions, as well as notions of tractability. Afterward, in Section 3, we illustrate the hardness of the problem with an example studied by Novak and Woźniakowski [7] and show how

E-mail address: markus.weimar@uni-jena.de.

0885-064X/$ – see front matter
doi:10.1016/j.jco.2011.03.004

weighted spaces can help to improve this negative result. Thereby, we especially concentrate on so-called *product weights*. While there exists a well-developed concept to handle problems defined on Hilbert spaces, we need an essentially new approach to conclude the results in the general Banach space setting. These new ideas are presented in Section 4. Using this technique we prove a lower error bound for a very small class of functions, i.e., we consider the space $\mathcal{P}_d^\gamma$ of d-variate polynomials of degree at most one in each coordinate, equipped with some weighted norm. In Section 5 we recall a known result of Kuo et al. [3] about upper error bounds on a certain weighted reproducing kernel Hilbert space $\mathcal{H}_d^\gamma$. Next, in Section 6, we prove the three main theorems of this paper. That is, we show necessary and sufficient conditions for several kinds of tractabilities for a whole scale of weighted Banach function spaces $\mathcal{F}_d^\gamma$, where $\mathcal{P}_d^\gamma \hookrightarrow \mathcal{F}_d^\gamma \hookrightarrow \mathcal{H}_d^\gamma$, in terms of the weights γ. In particular, we provide a characterization of weak tractability and the curse of dimensionality. It is shown that for these kinds of tractability results we can restrict ourselves to linear non-adaptive algorithms. We illustrate our results by applying them to selected examples and discuss a typical case of product weights. Finally, in Section 7, we add some remarks about possible extensions of the result to other domains. In addition, we briefly consider the L_p-approximation problem for $1 \leq p < \infty$ and correct a small mistake stated in [7].

2. The approximation problem

We investigate the tractability properties of the approximation problem defined on some Banach spaces $\mathcal{F}_d$ of bounded functions $f\colon [0,1]^d \to \mathbb{R}$. We want to minimize the *worst case error*

$$\mathrm{e}^{\mathrm{wor}}(A_{n,d};\mathcal{F}_d) = \sup_{f\in\mathcal{B}(\mathcal{F}_d)} \left\| f - A_{n,d}(f) \mid L_\infty([0,1]^d) \right\|$$

with respect to all algorithms $A_{n,d} \in \mathcal{A}_n$ that use n pieces of information in d dimensions from a certain class Λ. Here $\mathcal{B}(\mathcal{F}_d) = \{f \in \mathcal{F}_d \mid \|f \mid \mathcal{F}_d\| \leq 1\}$ denotes the unit ball of $\mathcal{F}_d$. Hence, we study the *nth minimal error*

$$e(n,d;\mathcal{F}_d) = \inf_{A_{n,d}\in\mathcal{A}_n} \mathrm{e}^{\mathrm{wor}}(A_{n,d};\mathcal{F}_d)$$

of L_∞-approximation on $\mathcal{F}_d$. An algorithm $A_{n,d} \in \mathcal{A}_n$ is modeled as a mapping $\phi\colon \mathbb{R}^n \to L_\infty([0,1]^d)$ and a function $N\colon \mathcal{F}_d \to \mathbb{R}^n$ such that $A_{n,d} = \phi \circ N$. In detail, the information map N is given by

$$N(f) = (L_1(f), L_2(f), \ldots, L_n(f)), \quad f \in \mathcal{F}_d, \tag{1}$$

where $L_j \in \Lambda$. Here we distinguish certain classes of information operations Λ. In one case we assume that we can compute arbitrary continuous linear functionals. Then $\Lambda = \Lambda^{\mathrm{all}}$ coincides with $\mathcal{F}_d^*$, the dual space of $\mathcal{F}_d$. Often only function evaluations are permitted, i.e., $L_j(f) = f\left(t^{(j)}\right)$ for a certain fixed $t^{(j)} \in [0,1]^d$. In this case $\Lambda = \Lambda^{\mathrm{std}}$ is called *standard information*. If function evaluation is continuous for all $t \in [0,1]^d$ we have $\Lambda^{\mathrm{std}} \subset \Lambda^{\mathrm{all}}$. If L_j depends continuously on f but is not necessarily linear the class is denoted by Λ^{cont}. Note that in this case also N is continuous and we obviously have $\Lambda^{\mathrm{all}} \subset \Lambda^{\mathrm{cont}}$.

Furthermore, we distinguish between *adaptive* and *non-adaptive* algorithms. The latter case is described above in formula (1), where L_j does not depend on the previously computed values $L_1(f), \ldots, L_{j-1}(f)$. In contrast, we also discuss algorithms of the form $A_{n,d} = \phi \circ N$ with

$$N(f) = (L_1(f), L_2(f;y_1), \ldots, L_n(f;y_1,\ldots,y_{n-1})), \quad f \in \mathcal{F}_d, \tag{2}$$

where $y_1 = L_1(f)$ and $y_j = L_j(f;y_1,\ldots,y_{j-1})$ for $j = 2,3,\ldots,n$. If N is adaptive we restrict ourselves to the case where L_j depends linearly on f, i.e., $L_j(\cdot;y_1,\ldots,y_{j-1}) \in \Lambda^{\mathrm{all}}$.

In all cases of information maps, the mapping ϕ can be chosen arbitrarily and is not necessarily linear or continuous. The smallest class of algorithms under consideration is the class of linear, non-adaptive algorithms of the form

$$(A_{n,d}f)(x) = \sum_{j=1}^n L_j(f) \cdot g_j(x), \quad x \in [0,1]^d,$$