Journal of Complexity 25 (2009) 398–404

Contents lists available at ScienceDirect

Journal of Complexity

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

Approximation of infinitely differentiable multivariate functions is intractable

Erich Novak [a], Henryk Woźniakowski [b,c,*]

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

[b] Department of Computer Science, Columbia University, New York, NY 10027, USA

[c] Institute of Applied Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland

ARTICLE INFO

Article history:
Received 9 September 2008
Received in revised form
13 November 2008
Accepted 18 November 2008
Available online 24 November 2008

Keywords:
Tractability
Curse of dimensionality
Rate of convergence
Approximation of smooth functions

ABSTRACT

We prove that L_∞-approximation of C^∞-functions defined on $[0, 1]^d$ is intractable and suffers from the curse of dimensionality. This is done by showing that the minimal number of linear functionals needed to obtain an algorithm with worst case error at most $\varepsilon \in (0, 1)$ is exponential in d. This holds despite the fact that the rate of convergence is infinite.

1. Introduction and result

The rate (order) of convergence is an important concept of numerical analysis and approximation theory. The rate of convergence measures how fast the minimal error $e(n)$ of algorithms using n function values or linear functionals goes to zero. Roughly speaking, if $e(n) = \Theta(n^{-\alpha})$, then the rate is α. To guarantee that the error is ε, we must take $n = \Theta(\varepsilon^{-1/\alpha})$ as ε tends to 0. Hence, asymptotically in ε, the larger the rate of convergence the easier the problem. However, it is not clear what this means for a fixed positive ε and how long we have to wait for the asymptotic behavior.

In this paper we assume that F_d is a normed linear space of functions $f : [0, 1]^d \to \mathbb{R}$ that are infinitely differentiable with respect to all variables and the norm of $f \in F_d$ is given as the largest

* Corresponding author at: Institute of Applied Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland.
E-mail addresses: novak@mathematik.uni-jena.de (E. Novak), henryk@cs.columbia.edu, onk62@hydra.mimuw.edu.pl (H. Woźniakowski).

0885-064X/$ – see front matter © 2009 Published by Elsevier Inc.
doi:10.1016/j.jco.2008.11.002

absolute value of all derivatives of f, i.e.,

$$\|f\|_{F_d} = \sup_{\alpha} \|D^\alpha f\|_\infty < \infty. \tag{1}$$

We approximate such functions with respect to the L_∞-norm. We consider the worst case setting (for the unit ball of F_d) and algorithms using arbitrary linear functionals as information operations on f. Here, d can be arbitrarily large. To stress the importance of d, we denote the minimal error $e(n)$ by $e(n, d)$. Finally, let $n(\varepsilon, d)$ denote the smallest number of linear functionals that is needed to find an algorithm with worst case error at most ε.

The optimal rate of convergence of this multivariate approximation problem is infinite since the functions have unbounded smoothness. That is, for any d and arbitrarily large r we have

$$e(n, d) = \mathcal{O}(n^{-r}) \quad \text{as } n \to \infty.$$

This implies that

$$n(\varepsilon, d) = \mathcal{O}(\varepsilon^{-1/r}) \quad \text{as } \varepsilon \to 0.$$

Hence for all d, we have an excellent asymptotic speed of convergence with respect to n. Furthermore, L_∞-approximation is asymptotically easy with respect to ε, since $n(\varepsilon, d)$ grows at most sub-linearly in ε^{-1}.

Obviously, the factors in the last two $\mathcal{O}$ bounds may depend on d. Despite the positive asymptotic results, it is not clear if the dependence on d is polynomial or exponential. This leads us to the notion of *tractability*.

Tractability means that $n(\varepsilon, d)$ does not depend exponentially on ε^{-1} and d. More precisely, a problem is *weakly tractable* if

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln n(\varepsilon, d)}{\varepsilon^{-1} + d} = 0,$$

and *intractable* if this relation does not hold. If $n(\varepsilon, d)$ depends exponentially on d, we say that a problem suffers from the *curse of dimensionality*.

Furthermore, a problem is *polynomially tractable* if there exist non-negative numbers C, p and q such that

$$n(\varepsilon, d) \le C\, \varepsilon^{-p} d^q \quad \text{for all } \varepsilon \in (0, 1) \text{ and } d \in \mathbb{N}.$$

If $q = 0$ above then a problem is *strongly polynomially tractable*. For a detailed discussion of tractability, the reader is referred to [1].

The L_∞-approximation problem presented here was studied by Huang and Zhang in [2]. They proved that this problem is *not* strongly tractable. Their main result says that

$$\lim_{d\to\infty} e(n, d) = 1$$

for each fixed n. These authors also conjectured that the problem is not polynomially tractable, see also Open Problem 2 of [1]. In this paper we prove the conjecture of Huang and Zhang. In fact, we prove that not only is the problem not polynomially tractable, but also that the problem is not weakly tractable. This also partially solves Open Problem 2 of [1] in the case of multivariate approximation.

Using the technique of [3], see also section 5.4 of [1], we prove the following theorem.

Theorem 1. *For L_∞-approximation defined over F_d we have*

$$e(n, d) = 1 \text{ for all } n = 0, 1, \ldots, 2^{\lfloor d/2 \rfloor} - 1.$$

Therefore

$$n(\varepsilon, d) \ \ge\ 2^{\lfloor d/2 \rfloor} \quad \text{for all } \varepsilon \in (0, 1) \text{ and } d \in \mathbb{N},$$

and L_∞-approximation is intractable and suffers from the curse of dimensionality.

This result illustrates that the rate of convergence does not tell us everything about the difficulty of solving the problem. We may have an excellent rate of convergence and exponential dependence on d. Equivalently, we must wait exponentially long to enjoy the excellent asymptotic behavior.

Similar results can be found in a few papers. For an approximation problem for C^∞-functions equipped with different norms than that considered here, it was shown in [4], see also Section 3.1.4 of [1], that the rate of convergence and tractability are not related. If the infinite smoothness is replaced by an arbitrarily large smoothness, then a similar result for multivariate approximation can be found in [3], see also section 5.4 of [4]. A similar result for multivariate integration can be found in [5].

We briefly compare the results of this paper to the results of [4]. The paper [4] is most relevant to the current paper. The norms studied in [4] are more general than here. In particular, functions are approximated with respect to a W_p^m-norm of the Sobolev space. Then intractability and the curse of dimensionality of approximation were established for $m \geq 1$ and all $p \in [1, \infty]$. For $m = 0$, the case $p = 2$ was only studied and weak tractability and the lack of polynomial tractability were established. It was mentioned as an open problem to verify weak tractability for, in particular, $p = \infty$. This open problem is solved here in the negative, i.e., weak tractability does not hold. This also partially solves Open Problem 5 of [4] in the case of $p = \infty$.

The choice of the domain $[0, 1]^d$ is not important. In fact, the curse of dimensionality is present for all domains of the form $[c_1, c_2]^d$ with $\ell = c_2 - c_1 > 0$. However, our bounds show that the curse of dimensionality may be delayed if ℓ is small, see Remark 2.

The choice of the L_∞-norm can be also replaced by the L_p-norm with $p \in [1, \infty)$. The curse of dimensionality is still present for all p and all domains $[c_1, c_2]^d$ with $\ell > 0$, see Remark 3.

We briefly comment on the proof technique used in this paper. We consider a subspace of polynomials that are linear in each variable. This subspace is obviously of dimension 2^d and has the property that the norms in the source and target spaces are the same if the length of the univariate domain interval is sufficient large. In fact, it is enough to assume that $\ell \geq 2(p+1)^{1/p}$. If this inequality holds then the nth minimal worst case errors are just 1 as long as $n < 2^d$. If ℓ is smaller than $2(p+1)^{1/p}$ then we group the variables to enlarge the domain, and show that the nth minimal worst case errors are still 1 for all $n < 2^{\lfloor d/k \rfloor}$ with $k = \lceil 2(p+1)^{1/p}/\ell \rceil$.

Finally, we add a few words about multivariate integration defined for the class F_d. It is conjectured in [6] that multivariate integration is not polynomially tractable. Wojtaszczyk [7] proved that this problem is *not strongly* tractable. Both polynomial and weak tractability of multivariate integration are still open. However, we show the curse of dimensionality for a related space V_d if quadratures with only non-negative coefficients are used, see Remark 4.

2. Proof

First, we precisely define how we approximate functions f from F_d and what we mean by the nth minimal error and the minimal number of information evaluations.

We approximate a function f from F_d by algorithms $A_{n,d}$ that may use arbitrary linear functionals, i.e.,

$$A_{n,d}(f) = \varphi_n \left(L_1(f), L_2(f), \ldots, L_n(f) \right), \tag{2}$$

where $\varphi_n : \mathbb{R}^n \to L_\infty([0, 1]^d)$ is some linear or non-linear mapping, and L_j is an arbitrary continuous linear functional whose choice may adaptively depend on the previously computed values $L_1(f), L_2(f), \ldots, L_{j-1}(f)$. The worst case error of $A_{n,d}$ is defined by

$$e^{\text{wor}}(A_{n,d}) = \sup_{\|f\|_{F_d} \leq 1} \|f - A_{n,d}(f)\|_\infty,$$

and the nth minimal error by

$$e(n, d) = \inf_{A_{n,d}} e^{\text{wor}}(A_{n,d}).$$

The minimal number of information operations needed to solve the problem to within ε is given by

$$n(\varepsilon, d) = \min \left\{ n : e(n, d) \leq \varepsilon \right\}.$$

We use the technique of [3], see also section 5.4 of [1]. For this technique it is enough to identify a linear space $V_d \subseteq C^\infty([0,1]^d)$ with $\dim V_d = k$ and

$$\|f\|_{F_d} = \|f\|_\infty \quad \text{for all } f \in V_d$$

to conclude that

$$e(k-1,d) = 1. \tag{3}$$

We now define an appropriate space V_d. For $d = 1$, we start with the elementary fact that for all $g : [-1,1] \to \mathbb{R}$ of the form $g(x) = ax + b$ we have

$$\sup_\alpha \|D^\alpha g\|_\infty = \|g\|_\infty = |a| + |b|,$$

where $D^\alpha g = g^{(\alpha)}$ for $\alpha \in \mathbb{N}_0$. It is useful to observe that the same equality

$$\sup_\alpha \|D^\alpha g\|_\infty = \|g\|_\infty \tag{4}$$

holds for $g(x) = ax + b$ on any interval $I = [c_1, c_2] \subseteq \mathbb{R}$ with length $c_2 - c_1 \geq 2$. Indeed, (4) is equivalent to the following inequality. For arbitrary real a, b, c_1, c_2 with $c_2 - c_1 \geq 2$ we have $|a| \leq \max(|ac_1 + b|, |ac_2 + b|)$. This holds for $a = 0$, and for $a \neq 0$, we can divide both sides by $|a|$ and we need to show that $1 \leq \max(|c_1 - t|, |c_2 - t|)$ for $t = -b/a$. Obviously, t that minimizes the maximum is $t = (c_2 - c_1)/2$ and then we need to have $1 \leq (c_2 - c_1)/2$ which holds due to the assumption. Observe that the condition $c_2 - c_1 \geq 2$ is generally necessary.

Let $d \geq 1$. Assume that $g : [-1,1]^d \to \mathbb{R}$ is of the form

$$g(x) = \sum_{i \in \{0,1\}^d} a_i x^i. \tag{5}$$

Here, $i = [i_1, i_2, \ldots, i_d]$ with $i_j \in \{0,1\}$ and $x^i = \prod_{j=1}^d x_j^{i_j}$.

Then g is linear in each variable, i.e., if all variables but x_j are fixed then g is linear in x_j. Therefore we can conclude from (4) that again

$$\sup_\alpha \|D^\alpha g\|_\infty = \|g\|_\infty, \tag{6}$$

where now $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_d]$ with $\alpha_j \in \mathbb{N}_0$, and

$$D^\alpha g = \frac{\partial^{|\alpha_1| + \cdots + |\alpha_d|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}\, g.$$

We stress that for the last conclusion we used the domain $[-1,1]^d$ instead of $[0,1]^d$ and again that (6) also holds for any cube $[c_1, c_2]^d$ with $c_2 - c_1 \geq 2$.

To consider the domain $[0,1]^d$ which is the common domain of functions from F_d, we take $s = \lfloor d/2 \rfloor$, and consider functions $f : [0,1]^{2s} \to \mathbb{R}$ of the form

$$f(x) = \sum_{i \in \{0,1\}^s} a_i (x_1 + x_2)^{i_1} (x_3 + x_4)^{i_2} \cdots (x_{2s-1} + x_{2s})^{i_s}. \tag{7}$$

Since $2s \leq d$, we have $f \in F_{2s} \subseteq F_d$. The last inclusion is understood in the following sense. Let $d_1 < d_2$. If $f \in F_{d_1}$ then f can be also regarded as a function of d_2 variables that is independent of $x_{d_1+1}, x_{d_1+2}, \ldots, x_{d_2}$. Note that in this case we have $\|f\|_{F_{d_1}} = \|f\|_{F_{d_2}}$.

We are ready to define the linear space V_d as the set of functions of the form (7) with arbitrary coefficients a_i. Clearly, $\dim(V_d) = 2^s$ and $V_d \subseteq F_{2s} \subseteq F_d$. We claim that

$$\|f\|_{F_d} = \|f\|_\infty \quad \text{for all } f \in V_d.$$

Indeed, let $z_j = x_{2j-1} + x_{2j} \in [0,2]$ for $j = 1, 2, \ldots, s$. For $f \in V_d$ of the form (7) define

$$g_f(z) = \sum_{i \in \{0,1\}^s} a_i z_1^{i_1} z_2^{i_2} \cdots z_s^{i_s}$$

which is of the form (5).

Note that for $f \in V_d$ and $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_{2s}]$ we have $D^\alpha f = 0$ if $\alpha_{2j-1} + \alpha_{2j} = 2$ for some $j \in [1, s]$. Furthermore for all α such that $\alpha_{2j-1} + \alpha_{2j} \le 1$ for all $j \in [1, s]$, we have

$$D^\alpha f(x) = D^\beta g_f(z),$$

where $\beta = [\beta_1, \beta_2, \dots, \beta_s]$ with $\beta_j = \alpha_{2j-1}$ if $\alpha_{2j-1} = 1$ or $\beta_j = \alpha_{2j}$ if $\alpha_{2j} = 1$, or $\beta_j = 0$ if $\alpha_{2j-1} = \alpha_{2j} = 0$. This yields that

$$\|f\|_{F_d} = \sup_\alpha \|D^\alpha f\|_\infty = \sup_\beta \|D^\beta g_f\|_\infty = \|g_f\|_\infty = \|f\|_\infty,$$

as claimed.

Hence, we can use (3) with $k = 2^s = 2^{\lfloor d/2 \rfloor}$ and the proof is completed. $\square$

Remark 2 (*More General Domains*). Similarly, we can obtain an intractability result for the space F_d of functions defined as before, except that the domain of functions is now an arbitrary cube $[c_1, c_2]^d$ with $\ell = c_2 - c_1 > 0$.

Choose $k = \lceil 2/\ell \rceil$ such that $k\ell \ge 2$. Then we can use functions of the form

$$g(x) = \sum_{i \in \{0,1\}^s} a_i (x_1 + x_2 + \cdots + x_k)^{i_1} (x_{k+1} + x_{k+2} + \cdots x_{2k})^{i_2} \cdots (x_{(k-1)s+1} + x_{(k-1)s+2} + x_{ks})^{i_s}$$

to conclude that for $s = \lfloor d/k \rfloor$ we obtain

$$e(n, d) = 1 \quad \text{for } n < 2^{\lfloor d/k \rfloor}.$$

Hence we get intractability for an arbitrary cube, i.e, for arbitrary ℓ. However, for small ℓ the curse of dimensionality is "delayed". For $\ell \ge 2$ we obtain

$$e(n, d) = 1 \quad \text{for } n < 2^d.$$

The last bound 2^d can be improved for larger ℓ. For example, if $\ell \ge 8$ then we can start for $d = 1$ with polynomials

$$g(x) = a + bx + cx^2.$$

We then obtain[1] $\|g\|_{F_1} = \|g\|_\infty$, and hence

$$e(n, d) = 1 \quad \text{for } n < 3^d.$$

Remark 3 (*The L_p-norm*). Using the same proof technique we can show the curse of dimensionality for a modified approximation problem defined as follows. For $p \in [1, \infty]$ and $d \in \mathbb{N}$, let $F_{d,p}$ be the class of functions $f : [c_1, c_2]^d \to \mathbb{R}$ that are infinitely differentiable and for which

$$\|f\|_{F_{d,p}} = \sup_\alpha \|D^\alpha f\|_{L_p} < \infty.$$

Obviously, we assume that $c_2 > c_1$.

---

[1] Assume that the domain is $[-4, 4]$. We need to show that

$$\max(\|g'\|_\infty, \|g''\|_\infty) \le \|g\|_\infty.$$

We have two cases. Case 1: $8|c| < |b|$. Then the last inequality is equivalent to $8|c| + |b| \le \max(|a + 16c|, 4|b|)$. Since $8|c| + |b| < 2|b|$ we are done. Case II: $8|c| \ge |b|$. Then we need to show that $|b| + 8|c| \le \max(|a + 16c| + 4|b|, |a - b^2/(4c)|)$. Dividing by $|c|$ we have $8 + |b/c| \le \max(|16 + a/c|, |a/c - (b/c)^2/4|)$. This is obvious if $a/c \ge 0$. If $a/c < 0$ and $|a/c| \le 8 + 3|b/c|$, then the first term of the maximum is at least $8 + |b/c|$; if $a/c < 0$ and $|a/c| > 8 + 3|b/c|$, then the second term does the job.

We want to approximate f from $F_{d,p}$ in the L_p-norm, i.e., the worst case error of an algorithm $A_{n,d}$ given by (2) is now given by

$$e_p^{\mathrm{wor}}(A_{n,d}) = \sup_{\|f\|_{F_{d,p}} \le 1} \|f - A_{n,d}(f)\|_{L_p},$$

and the nth minimal error by

$$e_p(n,d) = \inf_{A_{n,d}} e_p^{\mathrm{wor}}(A_{n,d}).$$

Note that for $p = \infty$, we have the case studied before.

It is easy to check that for the subspace V_d of linear (in each variable) polynomials g we have

$$\|g\|_{F_{d,p}} = \|g\|_{L_p}$$

whenever $\ell := c_2 - c_1 \ge 2(p+1)^{1/p}$. If this inequality holds then

$$e_p(n,d) = 1 \quad \text{for } n < 2^d.$$

If $\ell < 2(p+1)^{1/p}$ then we define $k = \lceil 2(p+1)^{1/p}/\ell \rceil$ such that $k\ell \ge 2(p+1)^{1/p}$. Using the same reasoning as in Remark 2 we conclude that

$$e_p(n,d) = 1 \quad \text{for } n < 2^{\lfloor d/k \rfloor}.$$

Hence, we have the curse of dimensionality for any value of p and for any domain.

Remark 4 (*Integration*). We now consider multivariate integration $\int_{[0,1]^d} f(x)\,\mathrm{d}x$ for $f \in V_d$. Here, V_d is the same 3^d-dimensional space of quadratic polynomials over $[0,1]^d$ as in Remark 2 and as in [8]. We use a tensor product norm, and we start for $d = 1$ with a norm

$$\|f\|^2 = \sum_{j=0}^{k} \|D^j f\|_{L_2}^2,$$

where $k \ge 2$ is fixed. Of course, we can also take $k = \infty$. Observe that the unit ball of V_d contains a function with $\|f\|_\infty > 1$, hence it is not contained in the unit ball of F_d. For positive quadrature formulas $Q_n(f) = \sum_{i=1}^n a_i f(x_i)$ with non-negative a_i, it was proved in [8] that

$$e^{\mathrm{wor}}(Q_n)^2 \ge 1 - n \cdot c^d, \tag{8}$$

with $c = 0.9985$. Hence the integration problem is intractable on V_d for positive quadrature formulas. In particular, the problem is intractable for quasi-Monte Carlo methods for the space V_d. However, it is not known whether positive quadrature formulas are optimal for V_d and whether a lower bound of the form (8) also holds for general quadrature formulas, i.e., for quadrature formulas with some negative a_i. It is also not known whether (8) holds for F_d instead of V_d.

Remark 5 (*Borsuk–Ulam Theorem*). The lower error bounds of this paper hold for algorithms of the form (2) with linear functionals L_j. For the proof technique we identified a linear space

$$V_d \subseteq F_d \quad \text{with } \dim V_d = k \quad \text{and } \|f\|_{F_d} = \|f\|_{G_d} \tag{9}$$

to conclude that $e(k-1,d) \ge 1$.

Assuming (9), we claim that the same lower bound $e^{\mathrm{wor}}(A_{k-1,d}) \ge 1$ also holds for arbitrary approximations $A_{k-1,d}$ of the form

$$A_{k-1,d} = \varphi \circ \mathbb{N}, \quad \text{where } N : F_d \to \mathbb{R}^{k-1} \text{ is continuous} \tag{10}$$

(but otherwise arbitrary) and $\varphi : \mathbb{R}^{k-1} \to G_d$ is arbitrary.

This follows from the Borsuk–Ulam theorem, which states that for any continuous $N : V_d \to \mathbb{R}^{k-1}$ there is an $f \in V_d$ with $\|f\|_\infty = 1$ and $N(f) = N(-f)$. Hence $A_{k-1,d}(f) = A_{k-1,d}(-f)$, and so

$$\begin{aligned} e^{\mathrm{wor}}(A_{k-1,d}) &\geq \max(\|f - A_{k-1,d}(f)\|_\infty, \|-f - A_{k-1,d}(-f)\|_\infty) \\ &= \max(\|f - A_{k-1,d}(f)\|_\infty, \|f + A_{k-1,d}(f)\|_\infty) \\ &\geq \|f\|_\infty = 1, \end{aligned}$$

as claimed. Hence the lower error bound also holds for other approximations, such as n-term approximations, as long as they can be written in the form (10).

Acknowledgments

We thank Art Werschulz for his valuable comments. The second author was partially supported by the National Science Foundation under Grant DMS-0608727.

References

[1] E. Novak, H. Woźniakowski, Tractability of Multivariate Problems, Volume 1, Linear Information, European Math. Soc, Zürich, 2008.
[2] F.L. Huang, S. Zhang, Approximation of infinitely differentiable multivariate functions is not strongly tractable, J. Complexity 23 (2007) 73–81.
[3] A.G. Werschulz, H. Woźniakowski, Tractability of multivariate approximation over a weighted unanchored Sobolev space, Constr. Approx. (in press).
[4] E. Novak, H. Woźniakowski, Optimal order of convergence and (in)tractability of multivariate approximation of smooth functions, Constr. Approx. (in press).
[5] I.H. Sloan, H. Woźniakowski, An intractability result for multiple integration, Math. Comp. 66 (1997) 1119–1124.
[6] H. Woźniakowski, Open problems for tractability of multivariate integration, J. Complexity 19 (2003) 434–444.
[7] J.O. Wojtaszczyk, Multivariate integration in $C^\infty([0,1]^d)$ is not strongly tractable, J. Complexity 19 (2003) 638–643.
[8] E. Novak, Is there a curse of dimension for integration?, in: Choi-Hong Lai et al. (Eds.), 11th Int. Conf. on Domain Decomposition Methods DDM. org, Augsburg, London, 1998, pp. 89–96.