The article is devoted to ridge functions, i.e., functions $f$ of several variables, such that there is a direction vector $a$ along which $f$ may vary; along lines perpendicular to $a$, the function is constant. In other words,

$$f(x) = g(a \cdot x), \quad x \in \mathbb{R}^d,$$

where $g$ is a univariate function called the profile.

Despite the fact that being a ridge function is a serious restriction, this is a frequent assumption in statistics. Furthermore, approximation by ridge function has also been investigated and had applications.

The problem of tractability is one of the main points of interest throughout the article.

The authors investigate entropy numbers to show how “large” the ridge function classes are, to quantify the compactness of the ridge function classes in $L_\infty$. It is shown that they are essentially as compact as the class of univariate Lipschitz functions.

The authors pay special attention to determination of the complexity of approximating ridge functions when the only available information is a limited amount of function values. The assumptions are the following: The domain of the ridge functions is $d$-dimensional Euclidean unit ball; the profiles are Lipschitz of order $\alpha > 0$ (including infinite smoothness $\alpha = \infty$); the ridge vectors are in an $l_p^d$-ball with $0 < p \leq 2$.

In addition, the case when $|g'(0)| \geq k$ for all admissible profiles $g$ and some prescribed $0 < k \leq 1$ is investigated. In this case, the complexity of sampling ridge functions reduces drastically to the complexity of sampling univariate Lipschitz functions. The authors obtained lower and upper bounds for the deterministic worst-case error with regard to standard information. It is observed that the sampling problem’s degree of difficulty varies essentially, depending on $\alpha$ and $p$. As the authors note, surprisingly, they could see almost the entire hierarchy of tractability levels as introduced in [E. Novak and H. Woźniakowski, Tractability of multivariate problems. Volume I: Linear information. Zürich: European Mathematical Society (EMS) (2008; Zbl 1156.65001)] and [E. Novak and H. Woźniakowski, Tractability of multivariate problems. Volume II: Standard information for functionals. Zürich: European Mathematical Society (EMS) (2010; Zbl 1241.65025)].

The article is organized as follows. Section 1 is an introduction. Precise definitions and several interesting auxiliary results are given in Section 2. Section 3 is devoted to entropy numbers for the ridge function classes. Lower and upper bounds for sampling numbers are found in Section 4. Section 5 is devoted specifically to tractability, in particular, to polynomial tractability. Here, the authors make connections of their findings on sampling numbers to information-based complexity.

The article should be interesting for specialists in approximation theory, data analysis, algorithms, and many areas of applied mathematics.

Reviewer: Alexander Tovstolis (Orlando)
Keywords:
ridge functions; sampling numbers; entropy numbers; rate of convergence; information-based complexity; curse of dimensionality; sampling approximation; tractability; polynomial tractability; weak tractability

Full Text: DOI arXiv

References:

[1] Bühlmann, P., van de Geer, S.: Statistics for High-Dimensional Data. Springer, Heidelberg (2011) - Zbl 1273.62015 · doi:10.1007/978-3-642-20192-9

[2] Buhmann, MD; Pinkus, A, Identifying linear combinations of ridge functions, Adv. Appl. Math., 22, 103-118, (1999) - Zbl 0931.41011 · doi:10.1006/aama.1998.0623

[3] Candès, EJ, Harmonic analysis of neural networks, Appl. Comput. Harmon. Anal., 6, 197-218, (1999) - Zbl 0931.68104 · doi:10.1006/acha.1998.0248

[4] Candès, EJ; Donoho, DL, Ridgelets: a key to higher-dimensional intermittency?, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 357, 2405-2509, (1999) - Zbl 1082.42503 · doi:10.1098/rsta.1999.0444

[5] Carl, B., Stefani, I.: Entropy, Compactness and the Approximation of Operators. Cambridge Tracts in Mathematics, vol. 98.

[6] Cohen, A; Daubechies, I; DeVore, RA; Kerkyacharian, G; Picard, D, Capturing ridge functions in high dimensions from point queries, Constr. Approx., 35, 229-243, (2012) - Zbl 1218.62286 · doi:10.1007/s00365-011-9117-6

[7] Creutzig, J; Dereich, S; Müller-Kronbach, T; Ritter, K, Infinite-dimensional quadrature and approximation of distributions, Found. Comput. Math., 9, 391-429, (2009) - Zbl 1177.65011 · doi:10.1007/s10208-008-9029-x

[8] Cucker, F., Zhou, D.-X.: Learning theory: an approximation theory viewpoint. Cambridge Monographs on Applied and Computational Mathematics, vol. 24. Cambridge University Press, Cambridge (2007) - Zbl 1274.41001

[9] DeVore, R.A., Lorentz, G.G.: Constructive Approximation. Springer, Berlin (1993) - Zbl 0797.62007 · doi:10.1007/978-3-662-09699-8

[10] Edmunds, D.E., Triebel, H.: Function Spaces, Entropy Numbers, Differential Operators. Cambridge Tracts in Mathematics, vol. 120. Cambridge University Press, Cambridge (1996) - Zbl 0885.46020 · doi:10.1007/978-0-387-98199-4

[11] Flad, HJ; Hackbusch, W; Khomroushij, BN; Schneider, R; Olshesky, V (ed.); Tsyritshnyikov, E (ed.), Concepts of data-sparse tensor-product approximation in many-particle modeling, (2010), Singapore

[12] Fornasier, M; Schnass, K; Vybíral, J, Learning functions of few arbitrary linear parameters in high dimensions, Found. Comput. Math., 12, 229-262, (2012) - Zbl 1252.65036 · doi:10.1007/s10208-012-9115-y

[13] Foucart, S; Pajor, A; Rahul, H; Ulrich, T, The Gelfand widths of lp-balls for 0 < p ≤ 1, Adv. Comput. Math., 32, 239-253, (2010) - Zbl 1204.41018 · doi:10.1007/s10208-010-9108-7

[14] Friedman, JH; Stuetzle, W, Projection pursuit regression, J. Am. Stat. Assoc., 76, 817-823, (1981) - doi:10.1080/01621459.1981.10477729

[15] Golubev, GK, Asymptotically minimax estimation of a regression function in an additive model, Problemy Peredachi Informatsii, 28, 101-112, (1992)

[16] Graham, R; Sloane, N, Lower bounds for constant weight codes, IEEE Trans. Inform. Theory, 26, 37-43, (1980) - Zbl 0441.94012 · doi:10.1109/TIT.1980.1056141

[17] Hastie, T., Tibshirani, R., Friedman, J., The elements of statistical learning. Springer, New York (2001) - Zbl 0973.62007 · doi:10.1007/978-0-387-21606-5

[18] Heinrichs, A, Mayer, S.: Entropy numbers of spheres in Banach and quasi-Banach spaces. University of Bonn, preprint - Zbl 1326.41011

[19] Heinrichs, A; Novak, E; Ulrich, M; Woźniakowski, H, The curse of dimensionality for numerical integration of smooth functions II, J. Complex., 30, 117-143, (2014) - Zbl 1288.65040 · doi:10.1016/j.jco.2013.10.007

[20] Hristache, M; Juditsky, A; Spokoiny, V, Direct estimation of the index coefficient in a single-index model, Ann. Stat., 29, 595-623, (2001) - Zbl 1012.62043 · doi:10.1214/aos/1009210681

[21] Kühn, T, A lower estimate for entropy numbers, J. Approx. Theory, 110, 120-124, (2001) - Zbl 0995.47014 · doi:10.1006/jath.2000.3554

[22] Logan, BP; Shepp, LA, Optimal reconstruction of a function from its projections, Duke Math. J., 42, 645-659, (1975) - Zbl 0343.41020 · doi:10.1215/S0012-7094-75-04256-8

[23] Lorentz, G., von Golitschek, M., Makovoz, Y.: Constructive Approximation: Advanced Problems. Volume 304 of Grundlehren der Mathematischen Wissenschaften, Springer, Berlin (1996) - Zbl 0893.41007

[24] Maiorov, V, Geometric properties of the ridge manifold, Adv. Comput. Math., 32, 239-253, (2010) - Zbl 1188.41016 · doi:10.1007/s10444-008-9106-3

[25] Novak, E; Triebel, H, Function spaces in Lipschitz domains and optimal rates of convergence for sampling, Constr. Approx., 23, 325-350, (2006) - Zbl 1166.41014 · doi:10.1007/s00365-005-0612-y

[26] Novak, E., Woźniakowski, H.: Tractability of Multivariate Problems, Volume I: Linear Information. EMS Tracts in Mathematics, vol. 6, Eur. Math. Soc. Publ. House, Zürich (2008)

[27] Novak, E; Woźniakowski, H, Approximation of infinitely differentiable multivariate functions is intractable, J. Complex., 25, 398-404, (2009) - Zbl 1180.41004 · doi:10.1016/j.jco.2008.11.002
[28] Novak, E., Woźniakowski, H.: Tractability of Multivariate Problems, Volume II: Standard Information for Functionals. EMS Tracts in Mathematics, vol. 12, Eur. Math. Soc. Publ. House, Zürich (2010) · Zbl 0931.68104

[29] Paskov, S; Traub, J, Faster evaluation of financial derivatives, J. Portf. Manag., 22, 113-120, (1995) · doi:10.3905/jpm.1995.409541

[30] Pinkus, A; Méhauté, A (ed.); Rabut, C (ed.); Schumaker, LL (ed.), Approximating by ridge functions, 279-292, (1997), Nashville · Zbl 0937.65016

[31] Pinkus, A, Approximation theory of the MLP model in neural networks, Acta Numerica, 8, 143-195, (1999) · Zbl 0959.68109 · doi:10.1017/S0962492900002919

[32] Raskutti, G; Wainwright, MJ; Yu, B, Minimax-optimal rates for sparse additive models over kernel classes via convex programming. J. Mach. Learn. Res., 13, 389-427, (2012) · Zbl 1283.62071

[33] Schütt, C, Entropy numbers of diagonal operators between symmetric Banach spaces, J. Approx. Theory, 40, 121-128, (1984) · Zbl 0497.41017 · doi:10.1016/0021-9045(84)90021-2

[34] Schwab, C; Gittelson, CJ, Sparse tensor discretizations of high-dimensional parametric and stochastic pdes, Acta Numerica, 20, 291-467, (2011) · Zbl 1269.65010 · doi:10.1017/S0962492911000055

[35] Traub, J., Wasilkowski, G., Woźniakowski, H.: Information-Based Complexity. Academic Press, New York (1988)

[36] Triebel, H.: Fractals and Spectra. Birkhäuser, Basel (1997) · Zbl 0898.46030 · doi:10.1007/978-3-0348-0034-1

[37] Tyagi, H., Cevher, V.: Active learning of multi-index function models. In: Bartlett, P., Pereira, F.C.N., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 25, pp. 1475-1483. Curran Associates, Red Hook (2012) · Zbl 1244.41019

[38] Vybíral, J, Weak and quasi-polynomial tractability of approximation of infinitely differentiable functions, J. Complex., 30, 48-55, (2014) · Zbl 1308.46034 · doi:10.1016/j.jco.2013.04.003

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.