Reconstruction of ridge functions from function values

Sebastian Mayer

(joint work with Benjamin Doerr, Daniel Rudolph, Tino Ullrich, Jan Vybíral)

We are interested in reconstructing an unknown multivariate ridge function $f : \Omega \subseteq \mathbb{R}^d \to \mathbb{R}, x \mapsto g(a \cdot x)$ from a limited number of function values. The univariate function g is called the *profile* and the vector $a \in \mathbb{R}^d$ the *ridge direction*.

In [2] we studied the reconstruction of ridge functions which are defined on the Euclidean unit ball $\Omega = B_2^d := \{x \in \mathbb{R}^d : \|x\|_2 \le 1\}$. For $\alpha = s + \beta$, $s \in \mathbb{N}$, $0 < \beta \le 1$ and $0 < p \le 2$, consider the class of ridge functions

$$\mathcal{R}_d^{\alpha,p} := \{x \in \Omega \mapsto g(a \cdot x) : g \in \mathrm{Lip}_\alpha([-1,1]), \|g\|_\alpha \le 1, \|a\|_p \le 1\},$$

where $\|g\|_\alpha := \max\{\|g\|_\infty, \|g^{(1)}\|_\infty, \ldots, \|g^{(s)}\|_\infty, |g^{(s)}|_\beta\}$ and $|\cdot|_\beta$ denotes the Hölder constant with exponent β. Let $\mathcal{S}_n^{\mathrm{det}}$ be the class of all deterministic, adaptive sampling algorithms using at most n function values. For the deterministic worst-case error $\mathrm{err}_{\alpha,p}(n,d) := \inf_{S \in \mathcal{S}_n^{\mathrm{det}}} \sup_{f \in \mathcal{R}_d^{\alpha,p}} \|f - Sf\|_\infty$ we established the following characterization in terms of the *entropy numbers*.

Theorem 1 ([2, Section 4, Prop. 4.1, 4.2])**.** *Let $\alpha > 0$, $0 < p \le 2$ and $p' = \frac{1}{1-1/\max\{1,p\}}$. Then we have*

$$\varepsilon_n(\mathbb{S}_p^{d-1}, \ell_2^d)^{2\alpha} \lesssim \mathrm{err}_{\alpha,p}(n,d) \lesssim \varepsilon_{n/\binom{d+s}{s}}(B_2^d, \ell_{p'}^d)^\alpha,$$

Let $\Omega = [-1,1]^d$ be the unit cube. In [1] we consider the class of ridge functions

$$\mathcal{R}_d^{\alpha,p} := \{x \in \Omega \mapsto g(a \cdot x) : g \in \mathrm{Lip}_\alpha([-1,1]), \|g\|_\alpha \le 1, \|a\|_p \le 1\}$$

and the *probabilistic worst-case error* ${\mathrm{err}^{\mathrm{prob}}}_{\alpha,p}(n,d) := \inf_{S \in \mathcal{S}_n} \sup_{f \in \mathcal{R}_d^{\alpha,p}} e^{\mathrm{prob}}(S,f)$ where $e^{\mathrm{prob}}(S,f) = \inf\{\varepsilon > 0 : \ P(\|f - Sf\| \le \varepsilon) \ge 1/2\}$ and $\mathcal{S}_n$ is the class of all randomized, adaptive algorithms using at most n function values.

Theorem 2. *Let $\alpha > 1$ and $0 < p \le 1$. For*

$${\mathrm{err}^{\mathrm{prob}}}_{p,\alpha}(n,d) \lesssim \begin{cases} \left[1/\log(n)\right]^{\alpha(1/p-1)} & : n \le 2^d, \\ (2^d/n)^\alpha & : n > 2^d. \end{cases}$$

Theorem 3. *Let $\alpha > 0$ and $0 < p \le 1$. Then we have*

$${\mathrm{err}^{\mathrm{prob}}}_{p,\alpha}(n,d) \gtrsim \left[1/\log(n)\right]^{\alpha(1/p-1)}$$

for $n \le 2^{d/8}/4$.

References

[1] B. Doerr, S. Mayer, and D. Rudolph. *Tractability of capturing ridge functions on the cube.* Work in progress.

[2] S. Mayer, T. Ullrich, and J. Vybiral. *Entropy and sampling numbers of classes of ridge functions.* Constr. Approx. (2014), 1–34.