Order-optimal point configurations for function approximation

David Krieg

based on joint work with M. Sonnleitner

August 28th, 2020

JOHANNES KEPLER UNIVERSITÄT LINZ

FWF Der Wissenschaftsfonds.

Function approximation

Let F be a normed space of bounded functions on a bounded and convex domain $\Omega \subset \mathbb{R}^d$ with unit ball $B(F)$.

Let $P = \{x_1, \dots, x_n\} \subset \Omega$ be finite. The elements of P are called the sampling points. We want to approximate a function $f \in F$ based on the knowledge of $f(x_1), \dots, f(x_n)$, i.e., by algorithms

$$S_P(f) = \varphi(f(x_1), \dots, f(x_n)),$$

where $\varphi\colon \mathbb{R}^n \to L_\infty(\Omega)$ is arbitrary (or linear).

The error is measured in the $L_q(\Omega)$-norm for some $1 \leq q \leq \infty$. We consider the minimal worst case error

$$e^{\mathrm{lin}}(P, F \hookrightarrow L_q(\Omega)) := \inf_{S_P} \sup_{f \in B(F)} \left\| f - S_P(f) \right\|_{L_q(\Omega)}.$$

Sobolev spaces

The Sobolev space $W_p^s(\Omega)$ of smoothness $s \in \mathbb{N}$ and integrability $1 \le p \le \infty$ is the space

$$W_p^s(\Omega) = \left\{ f \in L_p(\Omega) \colon \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \in L_p(\Omega) \text{ for all } \alpha \in I(s,d) \right\},$$

where $I(s,d) = \{\alpha \in \mathbb{N}_0^d \colon |\alpha| = s\}$, equipped with the norm

$$\|f\|_{W_p^s(\Omega)} = \|f\|_{L_p(\Omega)} + |f|_{W_p^s(\Omega)},$$
$$|f|_{W_p^s(\Omega)} = \left\| \left\| \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \right\|_{L_p(\Omega)} \right\|_{\ell_p(I(s,d))}.$$

We will assume that $s > d/p$ such that the space consists of continuous and bounded functions.

Optimal sampling points

Theorem 1 (Novak/Triebel, 2006)

$$\inf_{\#P \le n} e^{\mathrm{lin}}(P, W_p^s(\Omega) \hookrightarrow L_q(\Omega)) \asymp n^{-s/d+(1/p-1/q)_+}.$$

- The optimal order is achieved for point sets with covering radius of the optimal order

$$h_{P,\Omega} := \sup_{y \in \Omega} \mathrm{dist}(y, P) \asymp n^{-1/d}.$$

- Novak and Triebel prove the result for bounded Lipschitz domains. The proof is based on a result of Wendland (2001).
- The result is much older for special domains like the cube.

Questions

Suppose that we cannot choose the sampling points P.

- Is there a simple way to determine their quality?
- Find a characterization of all those (sequences of) point sets that achieve the optimal order of convergence.
- How good are random sampling points?

Many authors use the covering radius to bound the error of sampling based algorithms. But this does not seem right: If we measure the error in L_q and q is small, then a few large gaps in the point set should be OK as long as most gaps are small.

Main result

Theorem 2

We have for any nonempty and finite point set $P \subset \Omega$ *that*

$$e^{\mathrm{lin}}(P, W_p^s(\Omega) \hookrightarrow L_q(\Omega)) \asymp \begin{cases} \|\operatorname{dist}(\cdot, P)\|_{L_\infty(\Omega)}^{s-d(1/p-1/q)} & \textit{if } q \geq p, \\ \|\operatorname{dist}(\cdot, P)\|_{L_\gamma(\Omega)}^{s} & \textit{if } q < p, \end{cases}$$

where $\gamma = s(1/q - 1/p)^{-1}$ *and the implied constants are independent of* P.

Main result

Theorem 2

We have for any nonempty and finite point set $P \subset \Omega$ that

$$e^{\mathrm{lin}}(P, W_p^s(\Omega) \hookrightarrow L_q(\Omega)) \asymp \begin{cases} \|\operatorname{dist}(\cdot, P)\|_{L_\infty(\Omega)}^{s-d(1/p-1/q)} & \text{if } q \geq p, \\ \|\operatorname{dist}(\cdot, P)\|_{L_\gamma(\Omega)}^{s} & \text{if } q < p, \end{cases}$$

where $\gamma = s(1/q - 1/p)^{-1}$ and the implied constants are independent of P.

Up next:

- discussion of the proof,
- related results (including the integration problem),
- examples and open problems.

Main result – What we are going to prove

Theorem 2

We have for any nonempty and finite point set $P \subset \Omega$ that

$$e^{\mathrm{lin}}(P, \dot{W}_p^s(\Omega) \hookrightarrow L_q(\Omega)) \asymp \begin{cases} \| \operatorname{dist}(\cdot, P)\|_{L_\infty(\Omega)}^{s-d(1/p-1/q)} & \text{if } q \geq p, \\ \| \operatorname{dist}(\cdot, P)\|_{L_\gamma(\Omega)}^{s} & \text{if } q < p, \end{cases}$$

where $\gamma = s(1/q - 1/p)^{-1}$ and the implied constants are independent of P.

Here,

$$\dot{W}_p^s(\Omega) = \{f \in W_p^s(\mathbb{R}^d) \mid \operatorname{supp} f \subset \overline{\Omega}\}.$$

A useful result from information-based complexity

For any normed space F of bounded functions (here $F = \dot{W}_p^s(\Omega)$), define the radius of information

$$r(P) = \sup\left\{\|f\|_{L_q(\Omega)} \colon f \in B(F),\ f|_P = 0\right\}.$$

Lemma 1 (Folklore)

$$r(P) \le e(P) \le 2r(P).$$

- See e.g. the book of Novak and Woźniakowski (2008).
- If F is a Hilbert space, we even have equality, also for $e^{\mathrm{lin}}(P)$.

Proof of the lower bound. Let $f_* \in B(F)$ with $f_*|_P = 0$. Any algorithm S_P of the form

$$S_P(f) = \varphi(f(x_1), \ldots, f(x_n))$$

satisfies $S_P(f_*) = S_P(-f_*)$. We call f_* a fooling function. We obtain that

$$\begin{aligned} &\sup_{f \in B(F)} \|S_P(f) - f\|_{L_q(\Omega)} \\ &\qquad \geq \max\left\{\|S_P(f_*) - f_*\|_{L_q(\Omega)}, \|S_P(f_*) + f_*\|_{L_q(\Omega)}\right\} \\ &\qquad\qquad \geq \|f_*\|_{L_q(\Omega)}. \end{aligned}$$

Proof of the upper bound. Consider an interpolatory algorithm

$$S_P(f) = \varphi(f(x_1), \ldots, f(x_n))$$

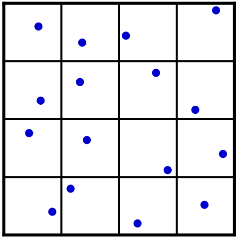
that maps $f \in B(F)$ to any function $g \in B(F)$ with $g\,|_P = f|_P$.

$$\sup_{f \in B(F)} \|f - S_P(f)\|_{L_q(\Omega)}$$

$$\leq \sup_{f,g \in B(F)\colon\, f|_P = g\,|_P} \| \underbrace{f - g}_{2h} \|_{L_q(\Omega)}$$

$$\leq \sup_{h \in B(F)\colon\, h\,|_P = 0} \|2h\|_{L_q(\Omega)}.$$

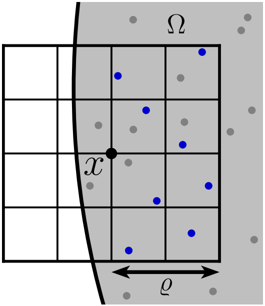
A useful result from approximation theory



There are constants m and C (depending on s, p, d) such that any function with a zero on each of the m^d subcubes of $[0,1]^d$ satisfies

$$\sup_{x \in [0,1]^d} |f(x)| \leq C \, |f|_{W_p^s([0,1]^d)}.$$

- "Small derivatives and enough zeros yield small functions."
- This can be proven using a result of Wendland (2001) on polynomial reproducing maps and results on best polynomial approximation on $W_p^s([0,1]^d)$, see the book of Maz'ya (1985).



We define $r(x)$ to be the infimum of all $\varrho > 0$ such that each of the m^d open subcubes of $x + [-\varrho, \varrho]^d$ contains either a point of P or a point of $\mathbb{R}^d \setminus \Omega$.

We call $Q(x) = x + [-r(x), r(x)]^d$ a good cube. Note that $Q(x)/2$ contains an empty cube $Q^*(x) \subset \Omega$ of radius $r(x)/2m$.

Lemma 2

Every $f \in \dot{W}_p^s(\Omega)$ with $f|_P = 0$ satisfies

$$\|f\|_{L_q(Q(x))} \lesssim r(x)^{s(1+d/\gamma)} |f|_{W_p^s(Q(x))}.$$

There is an efficient covering of Ω by good cubes.

Lemma 3

There are points $y_1, \ldots, y_N \in \overline{\Omega}$ such that

- *The cubes $Q_i = Q(y_i)$ cover $\overline{\Omega}$.*
- *The cubes $Q_i^* = Q^*(y_i)$ are pairwise disjoint.*
- *Every $y \in \mathbb{R}^d$ is contained in at most 2^d of the cubes Q_i.*

Proof. The function r is upper semi-continuous. Choose y_1 as a maximizer of r on $\overline{\Omega}$, and recursively y_k as a maximizer of r on $\overline{\Omega} \setminus \bigcup_{i<k} Q_i$. The rest of the proof is homework. □

Putting things together

Proof of the upper bound. Let f be from the unit ball of $\dot{W}_p^s(\mathbb{R}^d)$ such that $f|_P = 0$. Then

$$\|f\|_{L_q(\Omega)}^q \leq \sum \|f\|_{L_q(Q_i)}^q \lesssim \sum r_i^{sq(1+d/\gamma)} \, |f|_{W_p^s(Q_i)}^q$$

and Hölder's inequality gives us

$$\leq \Big(\sum \underbrace{r_i^{\gamma+d}}_{\int_{Q_i^*} \mathrm{dist}(y,P)^\gamma \, \mathrm{d}y} \Big)^{qs/\gamma} \Big(\sum \underbrace{|f|_{W_p^s(Q_i)}^p}_{\int_{Q_i} |\partial^\alpha f(y)|^p \, \mathrm{d}y} \Big)^{q/p}$$

and using the efficiency of the covering,

$$\lesssim \Big(\int_\Omega \mathrm{dist}(y, P)^\gamma \, \mathrm{d}y \Big)^{qs/\gamma}.$$

Putting things together

Proof of the lower bound. Recall that the Q_i^* are disjoint and empty subsets of Ω. Let T_i transform Q_i^* linearly into $[-1,1]^d$.

We take any nonnegative smooth function ψ with support in the cube $[-1,1]^d$ and $\psi(0) > 0$ and let $\psi_i = \psi \circ T_i$. Define the fooling function

$$f_* = \frac{\sum \alpha_i \psi_i}{\left\| \sum \alpha_i \psi_i \right\|_{W_p^s(\Omega)}}.$$

By optimizing the α_i, we obtain the lower bound

$$\|f_*\|_{L_q(\Omega)} \gtrsim \|\operatorname{dist}(\cdot, P)\|_{L_\gamma(\Omega)}^s.$$

Back to the theorem

Theorem 2

We have for any nonempty and finite point set $P \subset \Omega$ that

$$e^{\mathrm{lin}}(P, W_p^s(\Omega) \hookrightarrow L_q(\Omega)) \asymp \begin{cases} \|\operatorname{dist}(\cdot, P)\|_{L_\infty(\Omega)}^{s-d(1/p-1/q)} & \text{if } q \geq p, \\ \|\operatorname{dist}(\cdot, P)\|_{L_\gamma(\Omega)}^{s} & \text{if } q < p, \end{cases}$$

where $\gamma = s(1/p - 1/q)^{-1}$ and the implied constants are independent of P.

The result may be extended to ...

- ... a wider range of (isotropic) function spaces.
- ... more general domains.

Integration

The minimal worst-case error for the integration problem on a function space F is

$$e(P, F, \mathrm{INT}) := \inf_{a_i \in \mathbb{R}} \sup_{f \in B(F)} \Big| \int_\Omega f(x)\, \mathrm{d}x - \sum_{i=1}^n a_i\, f(x_i) \Big|.$$

- Smolyak/Bakhvalov (1971): The infimum does not change if we also allow nonlinear algorithms.

Theorem 3

We have for any nonempty and finite point set $P \subset \Omega$ *that*

$$e(P, W_p^s(\Omega), \mathrm{INT}) \asymp \big\| \operatorname{dist}(\cdot, P) \big\|_{L_\gamma(\Omega)}^s,$$

where $\gamma = s(1 - 1/p)^{-1}$.

Characterization of optimal sampling points

Corollary 1

A sequence of n-point sets P_n is optimal for $L_q(\Omega)$-approximation or integration on $W_p^s(\Omega)$ in the sense that the minimal worst case error has the optimal order of convergence if and only if

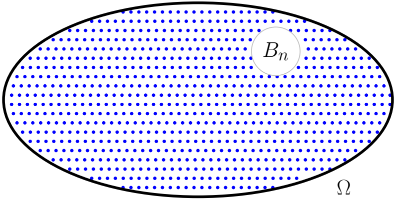
$$\|\operatorname{dist}(\cdot, P)\|_{L_\gamma(\Omega)} \asymp n^{-1/d},$$

where $\gamma = s(1/q - 1/p)^{-1}$ for $q < p$ (with $q = 1$ for the integration problem) and $\gamma = \infty$ for $q \geq p$.

- The quantity $\|\operatorname{dist}(\cdot, P)\|_{L_\gamma(\Omega)}$ is also called the distortion of the point set P (or of the quantizer that "rounds" elements from Ω to elements of P).

Example: Data with a big hole

For $n \in \mathbb{N}$, let P_n be an n-point set and let B_n be a ball of radius r_n. Assume that the covering radius of P_n in $\Omega \setminus B_n$ is of optimal order $n^{-1/d}$.



- The point sets are order-optimal if and only if

$$r_n \lesssim n^{-1/d+1/(\gamma+d)}.$$

- Compare with the size $n^{-1/d}$ of the other holes.

Example: Random sampling points

Let P_n be a set of n independent, uniformly distributed points. Then we have in various ways

$$\big\| \operatorname{dist}(\cdot, P) \big\|_{L_\gamma(\Omega)} \asymp \begin{cases} n^{-1/d} & \text{if } \gamma < \infty, \\ (n/\log n)^{-1/d} & \text{if } \gamma = \infty. \end{cases}$$

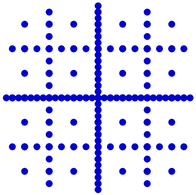
Precise statements may be found in Cohort (2004).

Corollary 2

Random data is optimal for $L_q(\Omega)$-approximation (integration) on $W_p^s(\Omega)$ if and only if $q < p$.

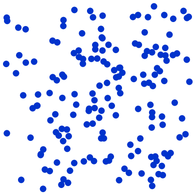
Some open questions

- We did not manage to prove the result for the integration problem for general Triebel Lizorkin spaces.
- We believe that the results hold for more general domains (including all bounded Lipschitz domains) and manifolds.
- We can only guess how optimal point sets look like for function spaces of mixed smoothness (i.e., the derivatives up to order $\frac{\partial^{sd}}{\partial x_1^s \dots \partial x_d^s}$ are bounded). So far, the best known point sets in the Hilbert-case are:



$\leftarrow$ Sparse grids for small dimensions [T. Ullrich/Sickel]

Random points for large dimensions [M. Ullrich/K.] $\rightarrow$



References I

P. Cohort.
Limit theorems for random normalized distortion.
Ann. Appl. Probab., 14(1):118–143, 2004.

D. Krieg and M. Ullrich.
Function values are enough for L_2-approximation.
arXiv:1905.02516, 2019.

V. G. Maz'ja.
Sobolev spaces.
Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, 1985.
Translated from the Russian by T. O. Shaposhnikova.

References II

E. Novak and H. Triebel.
Function Spaces in Lipschitz Domains and Optimal Rates of Convergence for Sampling.
Constr. Approx., 23:325–350, 2006.

E. Novak and H. Woźniakowski.
Tractability of multivariate problems. Vol. 1: Linear information, volume 6 of *EMS Tracts in Mathematics*.
European Mathematical Society (EMS), Zürich, 2008.

W. Sickel and T. Ullrich.
The smolyak algorithm, sampling on sparse grids and function spaces of dominating mixed smoothness.
East J. Approx., 13(4):387–425, 2007.

References III

H. Wendland.
Scattered Data Approximation.
Cambridge University Press, 2004.