Recovering Sobolev Functions: Optimal versus Given Samples

David Krieg

based on joint work with Mathias Sonnleitner
(and work in progress with Erich Novak)

Online workshop
Sampling recovery and related problems
Moscow, May 2020

JOHANNES KEPLER UNIVERSITÄT LINZ

Der Wissenschaftsfonds.

Function approximation

- Ω bounded, convex set in $\mathbb{R}^d$
- F normed space of functions on Ω with $F \subset L_q(\Omega)$
- P set of sampling points $x_1, \ldots, x_n \in \Omega$

Recovery algorithms:

$$S_P(f) = \varphi(f(x_1), \ldots, f(x_n)) \quad \text{with} \quad \varphi\colon \mathbb{R}^n \to L_q(\Omega).$$

Minimal worst case error:

$$\operatorname{err}(P, F, L_q) := \inf_{\varphi} \sup_{\|f\|_F \le 1} \big\| f - S_P(f) \big\|_{L_q(\Omega)}.$$

Function approximation

- Ω bounded, convex set in $\mathbb{R}^d$
- F normed space of functions on Ω with $F \subset L_q(\Omega)$
- P set of sampling points $x_1, \ldots, x_n \in \Omega$

Recovery algorithms:

$$S_P(f) = \varphi(f(x_1), \ldots, f(x_n)) \quad \text{with} \quad \varphi\colon \mathbb{R}^n \to L_q(\Omega).$$

Minimal worst case error:

$$\mathrm{err}^{\mathrm{lin}}(P, F, L_q) := \inf_{\varphi \text{ linear}} \sup_{\|f\|_F \le 1} \big\| f - S_P(f) \big\|_{L_q(\Omega)}.$$

(Sometimes linear algorithms are preferable.)

Sobolev spaces

Sobolev space with smoothness $s \in \mathbb{N}$ and integrability $1 \leq p \leq \infty$:

$$W_p^s(\Omega) := \left\{ f \in L_p(\Omega) \;\middle|\; \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \in L_p(\Omega) \text{ if } |\alpha| \leq s \right\}$$

Here $|\alpha| = \sum_{i=1}^{d} \alpha_i$. This is a Banach space with norm

$$\|f\|_{W_p^s(\Omega)} := \|f\|_{L_p(\Omega)} + \underbrace{\sum_{|\alpha|=s} \left\| \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \right\|_{L_p(\Omega)}}_{|f|_{W_p^s(\Omega)}}.$$

We will always assume that $s > d/p$ such that the space consists of continuous and bounded functions (Sobolev embedding theorem).

Optimal sampling points

Classically, we look for good (optimal) sampling point sets.

$$\mathrm{err}\left(n, W_p^s, L_q\right) := \inf_{|P| \leq n} \mathrm{err}\left(P, W_p^s, L_q\right).$$

Theorem 1 (Narcowich/Ward/Wendland, Novak/Triebel)

$$\mathrm{err}\left(n, W_p^s, L_q\right) \asymp n^{-s/d+(1/p-1/q)_+}.$$

- The hidden constants depend on everything but n.
- The bound is achieved by Wendland's polynomial reproducing map and equispaced points.
- Novak/Triebel prove the result for bounded Lipschitz domains, the result is much older for special domains like the cube.

Optimal sampling points

Classically, we look for good (optimal) sampling point sets.

$$\mathrm{err}\left(n, W_p^s, L_q\right) := \inf_{|P|\leq n} \mathrm{err}\left(P, W_p^s, L_q\right).$$

Theorem 1 (Narcowich/Ward/Wendland, Novak/Triebel)

$$\mathrm{err}^{\mathrm{lin}}\left(n, W_p^s, L_q\right) \asymp n^{-s/d+(1/p-1/q)_+}.$$

- The hidden constants depend on everything but n.
- The bound is achieved by Wendland's polynomial reproducing map and equispaced points. This algorithm is linear.
- Novak/Triebel prove the result for bounded Lipschitz domains, the result is much older for special domains like the cube.

Questions

Suppose now, that we cannot choose the sampling points.

1. How can we determine the quality of a given point set?
2. When do we get the optimal order of convergence?
3. How good are random sampling points?

Many authors use the radius of the largest hole in the point set (i.e. the covering radius) to bound the error of sampling based algorithms. If q is large, then this is indeed the right quantity (see below). If q is small, the covering radius is too restrictive: A few large holes should be OK as long as most holes are small.

Question 1:

How can we determine the quality of a given point set?

Notation

We write

$$\operatorname{dist}(x, P) := \min_{y \in P} \|x - y\|_2 .$$

This is the radius of the largest ball around x that does not intersect P. We call $B(x, \operatorname{dist}(x, P))$ a hole. The radius of the largest hole,

$$h_\Omega(P) := \| \operatorname{dist}(\cdot, P)\|_{L_\infty(\Omega)},$$

is called the covering radius of P in Ω. The radius of an average hole,

$$\mathcal{D}_{\Omega,\gamma}(P) := \| \operatorname{dist}(\cdot, P)\|_{L_\gamma(\Omega)},$$

where $0 < \gamma < \infty$, is called the distortion of P in Ω (or of the quantizer related to P).

General point sets: $q \geq p$

The error is determined by the size of the largest hole.

Theorem 2 (Narcowich/Ward/Wendland, Novak/Triebel)

Let $q \geq p$. Then we have for any finite point set $P \subset \Omega$ that

$$\operatorname{err}\left(P, W_p^s, L_q\right) \asymp h_\Omega(P)^{s-d(1/p-1/q)}$$

- The hidden constants depend on everything but P.
- The upper bound is achieved by Wendland's polynomial reproducing map.
- The result is not stated this way in these papers, but no further ideas are required.

General point sets: $q \geq p$

The error is determined by the size of the largest hole.

Theorem 2 (Narcowich/Ward/Wendland, Novak/Triebel)

Let $q \geq p$. Then we have for any finite point set $P \subset \Omega$ that

$$\mathrm{err}^{\mathrm{lin}}\left(P, W_p^s, L_q\right) \asymp h_\Omega(P)^{s-d(1/p-1/q)}$$

- The hidden constants depend on everything but P.
- The upper bound is achieved by Wendland's polynomial reproducing map. The algorithm is linear.
- The result is not stated this way in these papers, but no further ideas are required.

General point sets: $q < p$

The error is determined by the average hole size.

Theorem 3 (with M. Sonnleitner)

Let $q < p$. We put $\gamma = s(1/q - 1/p)^{-1}$. Then we have for any finite point set $P \subset \Omega$ that

$$\operatorname{err}\left(P, W_p^s, L_q\right) \asymp \mathcal{D}_{\Omega,\gamma}(P)^s \, .$$

- The hidden constants depend on everything but P.
- The upper bound is achieved if we apply Wendland's polynomial reproducing maps to a carefully chosen partition of the domain.

General point sets: $q < p$

The error is determined by the average hole size.

Theorem 3 (with M. Sonnleitner)

Let $q < p$. We put $\gamma = s(1/q - 1/p)^{-1}$. Then we have for any finite point set $P \subset \Omega$ that

$$\mathrm{err}^{\mathrm{lin}}\left(P, W_p^s, L_q\right) \asymp \mathcal{D}_{\Omega,\gamma}(P)^s \,.$$

- The hidden constants depend on everything but P.
- The upper bound is achieved if we apply Wendland's polynomial reproducing maps to a carefully chosen partition of the domain. The algorithm is linear.

On the proof of Theorem 3 (11 slides)

To avoid annoying boundary issues we sketch the proof for the modified class

$$\dot{W}_p^s(\Omega) := \Big\{ f \in W_p^s\big(\mathbb{R}^d\big) \;\big|\; f = 0 \text{ on } \mathbb{R}^d \setminus \Omega \Big\}$$

with zero boundary condition.

Remark: For this modified class, we do not need the convexity of the domain Ω. We may consider bounded open sets $\Omega \subset \mathbb{R}^d$.

The radius of information

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

$$\mathrm{rad}(P, F, L_q) := \sup\left\{ \|f\|_{L_q(\Omega)} \;\middle|\; f \in B(F),\; f|_P = 0 \right\}.$$

Lemma 1 (Folklore)

$$\mathrm{rad}(P, F, L_q) \,\le\, \mathrm{err}(P, F, L_q) \,\le\, 2\,\mathrm{rad}(P, F, L_q).$$

- See e.g. the book of Novak and Woźniakowski (2008).
- If F is a Hilbert space, we even have equality:

$$\mathrm{rad}(P, F, L_q) \,=\, \mathrm{err}(P, F, L_q) \,=\, \mathrm{err}^{\mathrm{lin}}(P, F, L_q).$$

Proof of Lemma 1

Lower bound. Let $f_* \in B(F)$ with $f_*|_P = 0$.

We call f_* a fooling function: Any algorithm S_P of the form

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

satisfies $S_P(f_*) = S_P(-f_*)$. We obtain that

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

Thus $\mathrm{err}(P, F, L_q) \geq \mathrm{rad}(P, F, L_q)$.

Proof of Lemma 1

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 $g|_P = f|_P$.

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

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

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

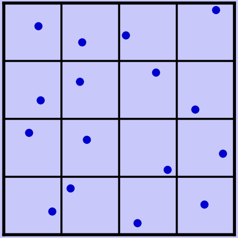
Thus $\operatorname{err}(P, F, L_q) \leq 2 \operatorname{rad}(P, F, L_q)$.

Thus we have to study the radius

$$\sup\Big\{\|f\|_{L_q(\Omega)} \;\Big|\; f \in W_p^s\big(\mathbb{R}^d\big),\ \|f\|_{W_p^s(\mathbb{R}^d)} \le 1, \quad f(x) = 0 \ \text{ if } \ x \in P \ \text{ or } \ x \in \mathbb{R}^d \setminus \Omega\Big\}.$$

Functions with scattered zeros

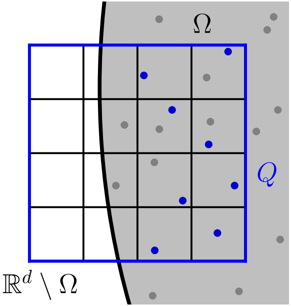
Lemma 2



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 say that $Q \subset \mathbb{R}^d$ is a good cube if each of the m^d subcubes of Q contains either an element of the sampling point set P or an element of $\mathbb{R}^d \setminus \Omega$.

Rescaling of the previous lemma:

Lemma 3

If $f \in \dot{W}_p^s(\Omega)$ with $f|_P = 0$ and Q is a good cube, then

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

For $x \in \Omega$, we define $r(x)$ to be the infimum of all $\varrho > 0$ such that $x + [-\varrho, \varrho]^d$ is a good cube. Then

$$Q(x) := x + [-r(x), r(x)]^d$$

is the smallest good cube centered at x.

- We can use Lemma 3 for $Q(x)$: If f vanishes on P then

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

- Since the cube with radius $r(x)/2$ is not a good cube, it contains a cube $Q^*(x) \subset \Omega$ of radius $r(x)/2m$ that does not intersect P.

There is an efficient covering of Ω by good cubes.

Lemma 4

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

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

Proof. The function $x \mapsto r(x)$ 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

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

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. We take any nonnegative smooth function ψ with compact support and $\psi(0) > 0$.

We let $\psi_i = \psi \circ T_i$ with a linear transformation T_i such that ψ_i is supported in Q_i^*. Recall that the Q_i^* are disjoint subsets of Ω and empty of P. 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 (yielding $\alpha_i \approx r_i^{s+\gamma/p}$), we obtain the lower bound

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

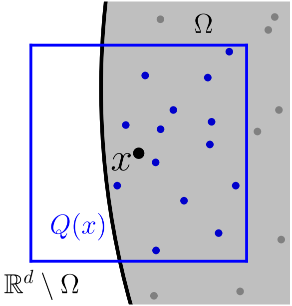
□

This finishes the proof of Theorem 3.

The algorithm (general case without boundary condition)

For every $x \in \Omega$, let $Q(x)$ be the smallest cube with center x that contains a constant number of well-distributed points:

$$h_{Q(x)\cap\Omega}(Q(x) \cap P) \leq c \cdot \mathrm{diam}(Q(x)).$$



The algorithm (general case without boundary condition)

For every $x \in \Omega$, let $Q(x)$ be the smallest cube with center x that contains a constant number of well-distributed points:

$$h_{Q(x)\cap\Omega}(Q(x) \cap P) \leq c \cdot \operatorname{diam}(Q(x)).$$

- Choose x_1 as a maximizer of $\operatorname{diam}(Q(x))$.
- Set $\Omega_1 = Q(x_1) \cap \Omega$ and $P_1 = P \cap \Omega_1$.
- There are functions $u_x \colon \Omega_1 \to \mathbb{R}$, $x \in P_1$, such that

  $$T_1 \colon L_\infty(\Omega_1) \to L_\infty(\Omega_1), \quad T_1(f) = \sum_{x \in P_1} f(x) u_x$$

  reproduces polynomials of degree s and its norm is at most a constant [Wendland]. For $x \in \Omega_1$, return $S_P f(x) = T_1 f(x)$.

The algorithm (general case without boundary condition)

For every $x \in \Omega$, let $Q(x)$ be the smallest cube with center x that contains a constant number of well-distributed points:

$$h_{Q(x)\cap\Omega}(Q(x) \cap P) \le c \cdot \operatorname{diam}(Q(x)).$$

- Choose x_2 as a maximizer of $\operatorname{diam}(Q(x))$ on $\Omega \setminus \Omega_1$.
- Set $\Omega_2 = Q(x_2) \cap \Omega$ and $P_2 = P \cap \Omega_2$.
- There are functions $u_x \colon \Omega_1 \to \mathbb{R}$, $x \in P_2$, such that

$$T_2 \colon L_\infty(\Omega_2) \to L_\infty(\Omega_2), \quad T_2(f) = \sum_{x \in P_2} f(x) u_x$$

  reproduces polynomials of degree s and its norm is at most a constant. For $x \in \Omega_2 \setminus \Omega_1$, return $S_P f(x) = T_2 f(x)$.

The algorithm (general case without boundary condition)

For every $x \in \Omega$, let $Q(x)$ be the smallest cube with center x that contains a constant number of well-distributed points:

$$h_{Q(x)\cap\Omega}(Q(x) \cap P) \leq c \cdot \mathrm{diam}(Q(x)).$$

- Choose x_3 as a maximizer of $\mathrm{diam}(Q(x))$ on $\Omega \setminus (\Omega_1 \cup \Omega_2)$.
- And so forth ...

Back to the result

Put $\gamma = s(1/q - 1/p)^{-1}$. Then

$$\operatorname{err}\left(P, W_p^s, L_q\right) \asymp \begin{cases} \left\| \operatorname{dist}(\cdot, P)\right\|_{L_\infty(\Omega)}^{s+d/q-d/p} & \text{if } q \geq p, \\ \left\| \operatorname{dist}(\cdot, P)\right\|_{L_\gamma(\Omega)}^{s} & \text{if } q < p. \end{cases}$$

This may be extended to ...

- ... non-integer smoothness $s > d/p$.
- ... the quasi-Banach case ($p < 1$ or $q < 1$).
- ... the wider range of (isotropic) Triebel-Lizorkin spaces.
- ... Sobolev spaces on compact Riemannian manifolds.

Integration

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

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

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

Theorem 4 (with M. Sonnleitner)

Integration on $W_p^s(\Omega)$ is as hard as $L_1(\Omega)$-approximation:

$$\mathrm{err}\left(P, W_p^s, \mathrm{INT}\right) \asymp \mathrm{err}\left(P, W_p^s, L_1\right)$$

(with constants independent of the point set 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 convexity is just a technical assumption and that the results hold for more general domains (including all bounded Lipschitz domains).
- Are there similar estimates for the quality of general point sets that take better care of tractability issues ?
- Do we have a similar result for function spaces of mixed smoothness if we replace the Euclidean distance by some kind of hyperbolic distance ???

Question 2:

For which point sets do we have the optimal order of convergence?

Observation

For any bounded set $\Omega \subset \mathbb{R}^d$ with positive measure and all $\gamma > 0$ we have

$$\inf_{P \subset \Omega:\, |P|=n} \big\| \operatorname{dist}(\cdot, P) \big\|_{L_\gamma(\Omega)} \asymp n^{-1/d}.$$

- Upper bound: We can cover Ω with n cubes with side-length $\asymp n^{-1/d}$. Choosing one point in each cube yields a covering radius of order $n^{-1/d}$.
- Lower bound: The balls of radius $cn^{-1/d}$ (with a small constant c) around the points of P cover at most half of Ω. In the other half, we have $\operatorname{dist}(\cdot, P) \geq cn^{-1/d}$.

Characterization of optimal sampling points

We put

$$\gamma := \begin{cases} \infty & \text{if } q \geq p, \\ s(1/q - 1/p)^{-1} & \text{if } q < p. \end{cases}$$

Corollary 1 (with M. Sonnleitner)

A sequence of n-point sets $P_n \subset \Omega$ is order-optimal for the $L_q(\Omega)$-approximation on $W_p^s(\Omega)$, i.e.,

$$\mathrm{err}\left(P_n, W_p^s, L_q\right) \asymp \mathrm{err}\left(n, W_p^s, L_q\right)$$

if and only if

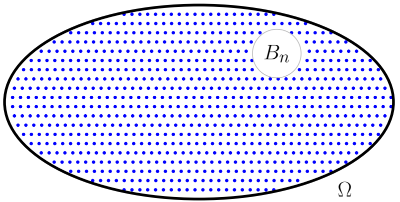
$$\left\| \mathrm{dist}(\cdot, P_n) \right\|_{L_\gamma(\Omega)} \asymp n^{-1/d}.$$

A simple example

Consider integration or L_1-approximation on $W_2^2(\Omega)$ with $\Omega \subset \mathbb{R}^2$.

The optimal covering radius is $n^{-1/2}$.

But: We can get an optimal approximation even if the point set has one or several holes of radius $n^{-1/3}$ since only the L_4-norm of the distance function counts.



Also random points are optimal here ...

Question 3:

How good are random sampling points?

We now consider the case that the sampling points $x_1, \ldots, x_n$ are independently and uniformly distributed (i.u.d.) on the domain Ω.

Covering radius of random points

Coupon collector problem: If we split Ω into m regions of equal volume, then we need roughly $n = m \log m$ independent uniformly distributed points to hit all of them.

Conversely: The largest hole in a point set P_n of n independent uniformly distributed points is typically of volume $\asymp \log n / n$. We have

$$\left[\mathbb{E}\, h_\Omega(P_n)^\alpha\right]^{1/\alpha} \asymp \left(\frac{\log n}{n}\right)^{1/d} \qquad (0 < \alpha < \infty).$$

Reznikow and Saff (2015) prove this for quite general metric measure spaces.

Distortion of random points

However, most holes in the point set are of volume $\lesssim 1/n$,

$$\left[\mathbb{E}\ \mathcal{D}_{\Omega,\gamma}(P_n)^\alpha\right]^{1/\alpha} \asymp \left(\frac{1}{n}\right)^{1/d} \qquad (0 < \alpha < \infty).$$

Upper bound for $\alpha = \gamma = 1$ (the other cases are similar):

$$\mathbb{E}\ \mathcal{D}_{\Omega,1}(P_n) = \int_\Omega \mathbb{E}\ \mathrm{dist}(x, P_n)\,\mathrm{d}x \lesssim n^{-1/d},$$

since for fixed x, the probability that $\mathrm{dist}(x, P_n)$ is larger than the radius of a ball with volume t/n is at most $(1 - t/n)^n \approx e^{-t}$.

Cohort (2004) proves more: The random variable $n^{1/d} D_{\Omega,\gamma}(P_n)$ converges in L_α and almost surely to a constant.

Together with Theorem 3 we obtain that

$$\mathbb{E}\,\mathrm{err}\left(P_n, W_p^s, L_q\right) \asymp \begin{cases} \mathrm{err}\left(\frac{n}{\log n}, W_p^s, L_q\right) & \text{if } q \geq p, \\ \mathrm{err}\left(n, W_p^s, L_q\right) & \text{if } q < p. \end{cases}$$

Conclusion (with M. Sonnleitner)

Independent and uniformly distributed points are optimal for $L_q(\Omega)$-approximation on $W_p^s(\Omega)$ if and only if $q < p$. They are optimal for integration on $W_p^s(\Omega)$ if and only if $p > 1$.

[See the survey with Hinrichs/Novak/Prochno/M. Ullrich for a predecessor in the case $q \geq p$.]

A different error criterion

Given an algorithm $S_P(f) = \varphi(f(x_1), \ldots, f(x_n))$ with a random point set $P = \{x_1, \ldots, x_n\}$ and a random recovery map $\varphi\colon \mathbb{R}^n \to L_q$, we may also consider the randomized error

$$\mathrm{err}^{\mathrm{ran}}(S_P, F, L_q) \;=\; \sup_{\|f\|_F \le 1} \mathbb{E}\, \big\| f - S_P(f) \big\|_{L_q(\Omega)}.$$

Difference: Previously, we considered $\mathbb{E} \sup$ instead of $\sup \mathbb{E}$. A small error meant that with high probability the recovery error is small for all functions $f \in F$ at once. Now, a small error means that for each individual function $f \in F$, the recovery error is small with high probability. The new error criterion is weaker, we might get better upper bounds.

Optimal points for the randomized error

For a given set P of random points, we define

$$\mathrm{err}^{\mathrm{ran}}(P, F, L_q) \;:=\; \inf_{S_P} \mathrm{err}^{\mathrm{ran}}(S_P, F, L_q).$$

Taking the infimum over all random n-point sets:

$$\mathrm{err}^{\mathrm{ran}}(n, F, L_q) \;:=\; \inf_{|P| \leq n} \mathrm{err}^{\mathrm{ran}}(P, F, L_q).$$

Theorem 5 (Mathé)

$$\mathrm{err}^{\mathrm{ran}}\left(n, W_p^s, L_q\right) \;\asymp\; \mathrm{err}\left(n, W_p^s, L_q\right)$$

We do not gain anything compared to the deterministic error. The optimal order is already achieved with deterministic points.

Exception: Now we might also allow $s \leq d/p$.

How good are i.u.d. samples?

What do we loose now, if we cannot choose the randomness, but are stuck with independent uniformly distributed samples P_n?

It turns out (work in progress) that

$$\mathrm{err}^{\mathrm{ran}}\left(P_n, W_p^s, L_q\right) \asymp \mathrm{err}^{\mathrm{ran}}\left(n, W_p^s, L_q\right) \quad \text{for all } q < \infty.$$

Conclusion (with E. Novak and M. Sonnleitner)

For $q < \infty$, independent uniformly distributed samples are optimal for $L_q(\Omega)$-approximation on $W_p^s(\Omega)$ in the randomized setting.

- In the deterministic setting, we needed $q < p$.
- We think that the optimality also holds for $q = \infty$, if $p \neq \infty$. It does not hold for $p = q = \infty$.

What's the difference?

We only need to consider $q \geq p$ (no difference for $q < p$).

Deterministic error: We draw the sampling points first. There will be a large hole of volume $\log n/n$. If we consider a bump function with support in this hole, no algorithm will be able to see it.

Randomized error: If we fix the bump function first (supported in a ball of volume $\log n/n$), it is very improbable that the random point set will miss the bump.

The case $p = \infty$ is an exception: We can cover the hole domain with such bumps, since additional bumps do not increase the norm of the function. The random point set will miss one of the bumps.

Proof of the non-optimality for $p = q = \infty$

Let $m \approx n/2\log n$. There are bump functions $f_1, \ldots, f_m$ with support in disjoint balls of volume $\asymp 1/m$ such that $\|f_i\|_{W^s_\infty} = 1$ and $\|f_i\|_\infty \gtrsim m^{-s/d}$ (simple scaling).

Let

$$F := \left\{ \sum_{i=1}^m \varepsilon_i f_i \,\middle|\, \varepsilon_i \in \{-1, 1\} \right\}.$$

Then $\|f\|_{W^s_\infty} = 1$ for all $f \in F$. For every algorithm S_{P_n} that uses the random sampling point set P_n, we have

$$\begin{aligned} \mathrm{err}^{\mathrm{ran}}(S_{P_n}, W^s_\infty, L_\infty) &= \sup_{\|f\|_{W^s_\infty} \le 1} \mathbb{E}\, \|f - S_{P_n}(f)\|_\infty \\ &\ge \sup_{f \in F} \mathbb{E}\, \|f - S_{P_n}(f)\|_\infty \ge \mathbb{E}\, \frac{1}{2^m} \sum_{f \in F} \|f - S_{P_n}(f)\|_\infty. \end{aligned}$$

Proof of the non-optimality for $p = q = \infty$

With constant probability, P_n misses one of the balls (recall the coupon collector's problem) and S_{P_n} cannot determine the sign of the corresponding bump. Thus

$$\|f - S_{P_n}(f)\|_\infty \geq \|f_i\|_\infty$$

for at least half of the functions $f \in F$ and

$$\frac{1}{2^m} \sum_{f \in F} \|f - S_{P_n}(f)\|_\infty \geq \frac{1}{2} \|f_i\|_\infty.$$

This yields

$$\mathrm{err}^{\mathrm{ran}}(S_{P_n}, W_\infty^s, L_\infty) \gtrsim m^{-s/d}.$$

□

Integration

For the integration problem, Bakhvalov (1962) stated

$$\mathrm{err}^{\mathrm{ran}}\left(n, W_p^s, \mathrm{INT}\right) \asymp n^{-s/d+(1/p-1/2)_+-1/2}.$$

- Lower bound: See Novak (1988).
- Upper bound: Use $n/2$ deterministic points as in Theorem 1 for L_2-approximation and $n/2$ i.u.d. points to estimate the integral of the remainder by standard Monte Carlo (variance reduction technique, see e.g. Heinrich).

By the previous result, the first $n/2$ points of this algorithm may be replaced by i.u.d. points as well.

Conclusion (with E. Novak and M. Sonnleitner)

Independent uniformly distributed samples are optimal for integration on $W_p^s(\Omega)$ with respect to the randomized error.

A glimpse on mixed smoothness

The Sobolev space of mixed smoothness $s \in \mathbb{N}$ and integrability $1 \le p \le \infty$ is given by

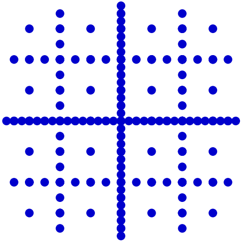
$$\mathbf{W}_p^s(\Omega) := \left\{ f \in L_p(\Omega) \;\middle|\; \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \in L_p(\Omega) \text{ if all } \alpha_i \le s \right\}$$

This is a Banach space with norm

$$\|f\|_{\mathbf{W}_p^s(\Omega)} := \sum_{\|\alpha\|_\infty \le s} \left\| \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \right\|_{L_p(\Omega)}.$$

For simplicity, we only consider $L_q(\Omega)$-approximation on the torus $\Omega = \mathbb{T}^d$ (modeled by $[0,1]^d$ where opposite faces are identified) and in the Hilbert case $p = 2$ (where linear algorithms are optimal).

Optimal point sets



$q > 2$: Sparse grids are optimal.

This was proven by Temlyakov for $q = \infty$ and by Dũng and Byrenheid/T. Ullrich for $q < \infty$. See the book of Dũng/Temlyakov/T. Ullrich.

$q \leq 2$: We do not know optimal point sets. So far, the best known point set is obtained via subsampling from independent uniformly distributed points, see Nagel/Schäfer/T. Ullrich.

?

How good are i.u.d. points?

For L_2-approximation, we know the following:

With M. Ullrich: For the deterministic error, independent and uniformly distributed points P_n are optimal up to logarithms:

$$\mathrm{err}\left(n, \mathbf{W}_2^s, L_2\right) \lesssim \mathbb{E}\,\mathrm{err}\left(P_n, \mathbf{W}_2^s, L_2\right) \lesssim \mathrm{err}\left(\frac{n}{\log n}, \mathbf{W}_2^s, L_2\right).$$

Algorithm: A least squares method.

K'19: For the randomized error, independent and uniformly distributed points are optimal:

$$\mathrm{err}^{\mathrm{ran}}\left(P_n, \mathbf{W}_2^s, L_2\right) \asymp \mathrm{err}^{\mathrm{ran}}\left(n, \mathbf{W}_2^s, L_2\right).$$

Algorithm: A multilevel Monte Carlo method.

Could we have the same phenomena as for isotropic smoothness?

A wild conjecture:

$$\mathbb{E}\,\mathrm{err}\left(P_n, \mathbf{W}_2^s, L_q\right) \asymp \begin{cases} \mathrm{err}\left(\frac{n}{\log n}, \mathbf{W}_2^s, L_q\right) & \text{if } q \geq 2, \\ \mathrm{err}\left(n, \mathbf{W}_2^s, L_q\right) & \text{if } q < 2. \end{cases}$$

- This would mean that i.u.d. points are optimal for $q < 2$.
- It would yield the order of convergence of $\mathrm{err}\left(n, \mathbf{W}_2^s, L_2\right)$ ...

Another wild conjecture:

$$\mathrm{err}^{\mathrm{ran}}\left(P_n, \mathbf{W}_2^s, L_q\right) \asymp \mathrm{err}^{\mathrm{ran}}\left(n, \mathbf{W}_2^s, L_q\right) \quad \text{for all } 1 \leq q \leq \infty.$$

References I

N. S. Bakhvalov.

On a rate of convergence of indeterministic integration processes within the functional classes Wp(l) (in Russian).

Theory. Probab. Appl., 7:227, 1962.

G. Byrenheid and T. Ullrich.

Optimal sampling recovery of mixed order Sobolev embeddings via discrete Littlewood-Paley type characterizations.

Anal. Math., 43(2):133–191, 2017.

D. Dũng.

B-spline quasi-interpolation sampling representation and sampling recovery in Sobolev spaces of mixed smoothness.

Acta Math. Vietnamica, 43:83–110, 2018.

D. Dũng, V. N. Temlyakov, and T. Ullrich.

Hyperbolic Cross Approximation.

Advanced Courses in Mathematics - CRM Barcelona. Springer International Publishing, 2018.

References II

S. Heinrich.
Random approximation in numerical analysis.
In K. D. Bierstedt and et al., editors, *Functional Analysis*, pages 123–171. Dekker, New York, 1994.

A. Hinrichs, D. Krieg, E. Novak, J. Prochno, and M. Ullrich.
On the power of random information.
In F. J. Hickernell and P. Kritzer, editors, *Multivariate Algorithms and Information-Based Complexity*, pages 43–64. De Gruyter, Berlin/Boston, 1994.

P. Mathé.
Random approximation of Sobolev embeddings.
J. Complexity, 7(3):261–281, 1991.

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

References III

N. Nagel, M. Schäfer, and T. Ullrich.
A new upper bound for sampling numbers.
Found. Comput. Math., to appear.

F. J. Narcowich, J. D. Ward, and H. Wendland.
Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting.
Math. Comp., 74(250):743–763, 2004.

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.

References IV

A. Reznikov and E. B. Saff.
The covering radius of randomly distributed points on a manifold.
Int. Math. Res. Not. IMRN, 2016(19):6065–6094, 2016.

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