Approximation and tractability of periodic Sobolev embeddings with increasing smoothness on high-dimensional domains

Thomas Kühn

Universität Leipzig, Germany

Workshop "On the frontiers of high-dimensional computation"
MATRIX, Creswick (Australia), 4–15 June 2018
11 June 2018

(Based on joint work with my diploma student Franziska Brückner)

High-dimensional approximation

- High-dimensional problems appear in many applications, examples will be presented in this workshop
- Quantum chemistry:
  N-particle systems modelled in Besov-type spaces
  $\curvearrowright$ approximation problem in dimension $d = 3N$, with huge N
- Financial mathematics:
  Stochastic PDEs, require measurements every day
  $\curvearrowright$ integration problem in dimension $d = 365n$ (n years)
- Often: Dimension not clear a priori (more particles, longer period)
- In this talk:
  Approximation of Sobolev functions on high-dimensional domains
- Aim: Tractability issues, with special emphasis on the dependence of the hidden constants on the dimension

Motivation

- Well-known fact: The approximation problem for isotropic Sobolev embeddings of fixed smoothness $s > 0$

$$I_d : H^s(\mathbb{T}^d) \to L_2(\mathbb{T}^d) \qquad (d \in \mathbb{N})$$

is at most weakly tractable, depending on the chosen norm.

several recent papers: K./Sickel/Ullrich (JoC 2014), Siedlecki/Weimar (JAT 2015), Chen/Wang (JoC 2017), Werschulz/Woźniakowski (JoC 2017),...

- Question: How can one improve the level of tractability?
- Natural option: Consider 'better' spaces, for instance...

Motivation

- Sobolev spaces of dominating mixed smoothness
- weighted Sobolev spaces
- spaces of C^∞- or analytic functions

  Many authors used such spaces for integration and approximation problems in IBC, here is an incomplete list:

  Chen, Dick, Fasshauer, Gnewuch, Hickernell, Irrgeher, Kritzer, Kühn, Kuo, Larcher, Laimer, Lifshits, Mayer, Novak, Papageorgiou, Petras, Pillichshammer, Sloan, T. Ullrich, Wang, Wasilkowski, Werschulz, Woźniakowski

  Most of these spaces are of tensor type.

- Our choice: isotropic Sobolev spaces with increasing smoothness

Approximation numbers

- **Approximation numbers** (also called linear widths) of a (bounded linear operator) $T : X \to Y$ between Banach spaces

$$a_n(T : X \to Y) := \inf\{\|T - A\| : \operatorname{rank} A < n\}$$

- **Many applications**

  Functional Analysis, Approximation Theory, Numerical Analysis,...

- **Useful properties**, in particular

  (1) Additivity $\quad a_{n+k-1}(S + T) \leq a_n(S) + a_k(T)$

  (2) Multiplicativity $\quad a_{n+k-1}(S \circ T) \leq a_n(S) \cdot a_k(T)$

  (3) Rank property $\quad \operatorname{rank} T < n \Longrightarrow a_n(T) = 0$

Interpretation in terms of algorithms

- Every operator $A : X \to Y$ of finite rank n can be written as

$$Ax = \sum_{j=1}^{n} L_j(x)\, y_j \qquad \text{for all } x \in X$$

with linear functionals $L_j \in X^*$ and vectors $y_j \in Y$.
$\curvearrowright$ A is a linear algorithm using arbitrary linear information

- worst-case error of the algorithm A

$$err^{wor}(A) := \sup_{\|x\| \le 1} \|Tx - Ax\| = \|T - A\|$$

- n-th minimal worst-case error of the approximation problem for T (w.r.t. linear algorithms and arbitrary linear information)

$$err_n^{wor}(T) := \inf_{\text{rank}\, A \le n} err^{wor}(A) = a_{n+1}(T)$$

Hilbert space setting

- Let $T : H \to F$ be a compact linear operator between Hilbert spaces.
- **Singular numbers** (= singular values, known from SVD)

$$s_n(T) := \sqrt{\lambda_n(T^* T)}$$

- **Schmidt representation.** $\exists$ ONS $(e_k) \subset H$ and $(f_k) \subset F$ s.t.

$$Tx = \sum_{k=1}^{\infty} s_k(T) \langle x, e_k \rangle f_k \quad \text{for all } x \in H\,.$$

- **Approximation numbers = singular numbers**

$$a_n(T) = \inf_{\operatorname{rank} A < n} \| T - A \| = s_n(T)$$

Best approximations - optimal algorithms

- **Truncated Schmidt representation** of $T : H \to F$

$$A_n x := \sum_{k=1}^{n} s_k(T)\langle x, e_k\rangle f_k \quad \curvearrowright \quad \|T - A_n\| = a_{n+1}(T) = err_n^{wor}(T)\,.$$

- **Input:** Linear information on an element of $x \in H$,
  n Fourier coefficients of x w.r.t the ONS (e_k)

  Output: $A_n x =$ best approximation of Tx,
  realizing the n-th minimal worst-case error,
  measured in the norm of the target space F.

- **Best approximation:** given by the concrete algorithm A_n.

Information complexity

- Let $\boxed{S_d : F_d \to G_d \, , \; d \in \mathbb{N}}$ be an approximation problem.
- Approximation numbers:

  Fixed number n of information $\Longrightarrow$ optimal error $a_{n+1}(S_d)$
- From a practical point of view it is more reasonable to

  fix an error level $\varepsilon > 0$ and ask how many pieces of information an optimal algorithm requires, i.e. to consider the 'inverse' function, the

  information complexity $\quad n(\varepsilon, d) := \min\{n \in \mathbb{N} : a_{n+1}(S_d) \leq \varepsilon\}$
- $\curvearrowright$ hierarchy of tractability notions, which describe the behaviour of $n(\varepsilon, d)$ as $\varepsilon \to 0$ and/or $d \to \infty$

Polynomial tractability notions

For an approximation problem

$$S_d : F_d \to G_d \qquad (d \in \mathbb{N})$$

we consider the following levels of tractability:

- **SPT** – strong polynomial tractability

$$n(\varepsilon, d) \le C(1/\varepsilon)^p \quad \text{for some } C, p > 0$$

- **PT** – polynomial tractability

$$n(\varepsilon, d) \le C(1/\varepsilon)^p d^q \quad \text{for some } C, p, q > 0$$

- **QPT** – quasi-polynomial tractability

$$n(\varepsilon, d) \le C \exp(t \cdot (1 + \log \tfrac{1}{\varepsilon})(1 + \log d)) \quad \text{for some } C, t > 0$$

Weak tractability notions

- (α,β)-**WT** – (α,β)-weak tractability $(\alpha,\beta>0)$

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

- **WT** – weak tractability $= (1,1)$-WT

  intractable $=$ not weakly tractable

- **UWT** – uniform weak tractability

  $=$ (α,β)-WT for all $\alpha,\beta>0$

- **curse of dimension** – There are $C>1$ and $\varepsilon_0>0$ s.t.

  $n(\varepsilon_0,d)\geq C^d$ for infinitely many d

- SPT $\Longrightarrow$ PT $\Longrightarrow$ QPT $\Longrightarrow$ UWT $\Longrightarrow$ (α,β)-WT $\Longrightarrow$ no curse

Isotropic Sobolev spaces $H^s(\mathbb{T}^d)$

- Torus $\mathbb{T} = [0, 2\pi]$, equipped with normalized Lebesgue measure
- Fourier coefficients of $f \in L_2(\mathbb{T}^d)$

$$\widehat{f}(k) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} f(x) e^{-ikx} dx \quad , \quad k \in \mathbb{Z}^d$$

- $H^s(\mathbb{T}^d)$ consists of all $f \in L_2(\mathbb{T})$ such that

$$\|f\| = \underbrace{\Big(\sum_{k \in \mathbb{Z}^d} \Big(1 + \sum_{j=1}^{d} |k_j|^p \Big)^{2s/p} |\widehat{f}(k)|^2 \Big)^{1/2}}_{\text{weighted } \ell_2\text{-sum of Fourier coefficients}} < \infty \,.$$

- Here $0 < p < \infty$ is an arbitrary parameter. But for fixed $s > 0$ and $d \in \mathbb{N}$, all these norms are equivalent, with equivalence constants depending on s and d. We will always work with $p = 1$.

Polynomial tractability

Theorem (Brückner, K. 2018)

For the approximation problem of isotropic Sobolev spaces

$$I_d : H^{s(d)}(\mathbb{T}^d) \to L_2(\mathbb{T}^d) \qquad (d \in \mathbb{N})$$

with increasing smoothness $0 < s(1) \le s(2) \le \ldots \le s(d) \le \ldots$ *we have*

- *SPT* $\iff$ *PT* $\iff \inf\limits_{d\in\mathbb{N}} \frac{s(d)}{d} > 0$
- *QPT* $\iff \inf\limits_{d\in\mathbb{N}} \frac{s(d)(1+\log d)}{d} > 0$

Remark:
At a first glance, the equivalence SPT $\iff$ PT is quite surprising,
but this effect appeared already in several other results in the literature.

Weak tractability

Theorem (Brückner, K. 2018)

Let $0 < s(1) \leq s(2) \leq \ldots$ *and set* $s := \lim_{d\to\infty} s(d)$.

Then the approximation problem

$$I_d : H^{s(d)}(\mathbb{T}^d) \to L_2(\mathbb{T}^d) \qquad (d \in \mathbb{N})$$

satisfies

- (α, β)- *WT* $\iff$ $\max(\alpha\, s\,, \beta) > 1 \qquad (\alpha, \beta > 0)$
- *WT* $\iff$ $s > 1$
- *UWT* $\iff$ $s = \infty$

Note: There is never curse of dimensionality.

Example - comparison with mixed spaces

- **Embeddings** $\boxed{H^{sd}(\mathbb{T}^d) \hookrightarrow H^s_{mix}(\mathbb{T}^d) \hookrightarrow H^s(\mathbb{T}^d)}$
- **Asymptotic behaviour of approximation numbers**

  For fixed d and s and $n \to \infty$, we have the following weak equivalences, with hidden constants depending on d and s.

  $a_n(I_d : H^{sd}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \sim n^{-s}$

  $a_n(I_d : H^s_{mix}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \sim n^{-s}(\log n)^{(d-1)s}$

  $a_n(I_d : H^s(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \sim n^{-s/d}$

- Remark: The spaces $H^{sd}(\mathbb{T}^d)$ and $H^s_{mix}(\mathbb{T}^d)$ are

  – very similar in the sense of approximation
  – but totally different concerning tractability

Optimal asymptotic constants

- For fixed s and d, the following limit exists.

$$\lim_{n\to\infty} \frac{a_n(I_d : H^{sd}(\mathbb{T}^d) \to L_2(\mathbb{T}^d))}{n^{-s}} = \left(\frac{2^d}{d!}\right)^s =: \lambda_s(d)$$

$\lambda = \lambda_s(d)$ is the optimal asymptotic constant in the following sense: For every $\varepsilon > 0$ there is $N_\varepsilon \in \mathbb{N}$ such that

$$\frac{1}{1+\varepsilon} \cdot \frac{\lambda}{n^s} \le a_n(I_d) \le (1+\varepsilon) \cdot \frac{\lambda}{n^s} \qquad \text{for all } n \ge N_\varepsilon \,.$$

- Compare with the corresponding result for mixed spaces

$$\lim_{n\to\infty} \frac{a_n(I_d : H^s_{mix}(\mathbb{T}^d) \to L_2(\mathbb{T}^d))}{n^{-s}(\log n)^{(d-1)s}} = \left(\frac{2^d}{(d-1)!}\right)^s$$

- For all $s > 0$ we have super-exponential decay of $\lambda_s(d)$ as $d \to \infty$.

Tractability

- Approximation problem $\boxed{I_d : F_d \to L_2(\mathbb{T}^d), \quad (d \in \mathbb{N})}$
- $F_d = H^{sd}(\mathbb{T}^d)$ isotropic, increasing smoothness

  strong polynomial tractability
- $F_d = H^s_{mix}(\mathbb{T}^d)$ dominating mixed smoothness

  quasi-polynomial tractability
- $F_d = H^s(\mathbb{T}^d)$ isotropic, fixed smoothness

  – **weakly tractable**, if $s > 1$
  – **intractable**, if $0 < s \leq 1$

Some ideas of the proofs

- Reduction of the problem in function spaces to a simpler problem for diagonal operators in sequence spaces (indexed by $\mathbb{Z}^d$)
- E.g., for $r \in \mathbb{N}$ and

$$n := \#\{k \in \mathbb{Z}^d : |k_1| + \ldots + |k_d| \leq r\},$$

  this implies

$$a_n(I_d : H^{s(d)}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) = (r+1)^{-s(d)}.$$

- Combinatorics, volume estimates in high-dimensional sequence spaces
- Rephrasing this is terms of information complexity $n(\varepsilon, d)$ plus some calculus proves the characterization of the different tractability levels.

Thank you for your attention!