Optimal approximation of smooth functions on high-dimensional domains

Thomas Kühn

Universität Leipzig, Germany

NPFSA 2017, Będlewo, 17–23 September 2017
22 September 2017

The talk is based on results from the following papers:

- T. Kühn, W. Sickel and T. Ullrich, *Approximation numbers of Sobolev embeddings – Sharp constants and tractability*, J. Complexity 30 (2014), 95–116.
- T. Kühn, W. Sickel and T. Ullrich, *Approximation of mixed order Sobolev functions on the d-torus – Asymptotics, preasymptotics and d-dependence*, Constr. Approx. 42 (2015), 353–398.
- F. Cobos, T. Kühn and W. Sickel, *Optimal approximation of multivariate periodic Sobolev functions in the sup-norm*, J. Funct. Anal. 270 (2016), 4196–4112.
- T. Kühn, S. Mayer and T. Ullrich, *Counting via entropy: New preasymptotics for the approximation numbers of Sobolev embeddings*, SIAM J. Numer. Anal. 54 (2016), 3625–3647.
- T. Kühn and M. Petersen, *Approximation in periodic Gevrey spaces*, work in progress.

High-dimensional approximation

- High-dimensional problems appear in many applications
- 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 numbers of embeddings of function spaces on high-dimensional domains
- Special emphasis:
  Dependence of the hidden constants on the dimension

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**

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

Best approximations - optimal algorithms

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

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

- **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.

- Note: The best approximation is given by the concrete algorithm A_n.

An example - Sobolev embeddings

- **Well-known** for Sobolev spaces of dominating mixed smoothness

$$c_{s,d} \cdot \left[\frac{(\log n)^{d-1}}{n}\right]^s \leq a_n(I_d : H^s_{mix}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \leq C_{s,d} \cdot \left[\frac{(\log n)^{d-1}}{n}\right]^s$$

- **Almost nothing known:**

  How do the constants $c_{s,d}$ and $C_{s,d}$ depend on s and d ?

  This is essential for high-dimensional numerical problems, and also for tractability questions in information-based complexity!

- Clearly, the constants heavily depend on the chosen norms.
  $\curvearrowright$ First we have to fix (somehow natural) norms.
  For all our norms, we will have norm one embeddings into $L_2(\mathbb{T}^d)$.

Asymptotics vs. preasymptotics

- To "see" the asymptotic rate

$$\varphi(n) := \left[\frac{(\log n)^{d-1}}{n} \right]^s$$

  in high dimensions, one has to wait super-exponentially long.

- (Dimension $d+1$): φ is increasing on $[1, e^d]$

$$\varphi(e^d) = \left(\frac{d}{e} \right)^{sd} \quad , \quad \varphi(d^d) = (\log d)^{sd} \gg 1 \, .$$

  Since $a_n \leq 1$ for all n, the asymptotic rate $\varphi(n)$ is useless for $n \leq d^d$. But d^d is much too large for practical purposes, even for moderate d.

- $\curvearrowright$ We need – information on the constants (d-dependence)
  – preasymptotic estimates (for small n, say $n \leq 2^d$)

General periodic spaces

- Fourier coefficients of $f \in L_2(\mathbb{T}^d)$,

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

- Given weights $w(k) \geq 1$, $k \in \mathbb{Z}^d$, let

  $F_d(w)$ be the space of all $f \in L_2(\mathbb{T}^d)$ such that

$$\|f|F_d(w)\| := \Big(\sum_{k \in \mathbb{Z}^d} w(k)^2 |c_k(f)|^2 \Big)^{1/2} < \infty \,.$$

- We have compact embeddings

$$F_d(w) \hookrightarrow L_2(\mathbb{T}^d) \quad \Longleftrightarrow \quad \lim_{|k| \to \infty} 1/w(k) = 0$$

$$F_d(w) \hookrightarrow L_\infty(\mathbb{T}^d) \quad \Longleftrightarrow \quad \sum_{k \in \mathbb{Z}^d} 1/w(k)^2 < \infty \,.$$

Isotropic Sobolev spaces

- Let $s > 0$, $d \in \mathbb{N}$, $0 < p < \infty$ and $\quad w_{s,p}(k) := \Big(1 + \sum_{j=1}^{d} |k_j|^p\Big)^{s/p}$.

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

  $$\|f|H^{s,p}(\mathbb{T}^d)\| := \Big(\sum_{k\in\mathbb{Z}^d} w_{s,p}(k)^2 |c_k(f)|^2\Big)^{1/2} < \infty .$$

- For fixed $s > 0$ and $d \in \mathbb{N}$, all these norms are equivalent, with equivalence constants depending on d.
  Therefore, all $H^{s,p}(\mathbb{T}^d)$, $0 < p < \infty$, coincide as vector spaces.

- The fine parameter p is motivated by comparison with classical norms. It also shows the very subtle dependence of approximation numbers on the chosen norms!

Comparison with classical norms

- Isotropic Sobolev spaces $H^m(\mathbb{T}^d)$ with integer smoothness $m \in \mathbb{N}$
- **Classical norm** (all partial derivatives)

$$\|f|H^m(\mathbb{T}^d)\| := \Big(\sum_{|\alpha|\le m} \| D^\alpha f \,|L_2(\mathbb{T}^d)\|^2 \Big)^{1/2} \sim \|f|H^{m,2}(\mathbb{T}^d)\|$$

with equivalence constants independent on d.

- **Modified classical norm** (only highest derivatives in each coordinate)

$$\Big(\| f \,|L_2(\mathbb{T}^d)\|^2 + \sum_{j=1}^{d} \Big\| \frac{\partial^m f}{\partial x_j^m} \Big| L_2(\mathbb{T}^d) \Big\|^2 \Big)^{1/2} = \|f|H^{m,m}(\mathbb{T}^d)\|$$

Sobolev spaces of dominating mixed smoothness

- Let $s > 0$, $d \in \mathbb{N}$, $0 < p < \infty$ and $\quad w_{s,p}^{mix}(k) := \prod_{j=1}^{d} (1 + |k_j|^p)^{s/p}$.

  $H_{mix}^{s,p}(\mathbb{T}^d)$ consists of all $f \in L_2(\mathbb{T}^d)$ such that

$$\|f|H_{mix}^{s,p}(\mathbb{T}^d)\| := \Big(\sum_{k \in \mathbb{Z}^d} w_{s,p}^{mix}(k)^2 |c_k(f)|^2 \Big)^{1/2} < \infty\,,$$

- For fixed $s > 0$ and $d \in \mathbb{N}$, all these norms are equivalent, hence all the spaces $H_{mix}^{s,p}(\mathbb{T}^d)$, $0 < p < \infty$, coincide as vector spaces.

Reduction to sequence spaces

Commutative diagram

$$
\begin{array}{ccc}
F_d(w) & \xrightarrow{\quad I_d \quad} & L_2(\mathbb{T}^d) \\
\Big\downarrow A & & \Big\uparrow B \\
\ell_2(\mathbb{Z}^d) & \xrightarrow{\quad D \quad} & \ell_2(\mathbb{Z}^d)
\end{array}
$$

$$
Af := (w(k)\, c_k(f))_{k \in \mathbb{Z}^d} \quad , \quad B\xi := \sum_{k \in \mathbb{Z}^d} \xi_k\, e^{ikx} \quad , \quad D(\xi_k) := (\xi_k / w(k))
$$

A and B are unitary operators $\curvearrowright$ $a_n(I_d) = a_n(D) = s_n(D) = \sigma_n$

where $(\sigma_n)_{n \in \mathbb{N}}$ is the non-increasing rearrangement of $(1/w(k))_{k \in \mathbb{Z}^d}$.

Combinatorics, isotropic case

- The weight "sequence" $(w_{s,p}(k))_{k\in\mathbb{Z}^d}$ is piecewise constant, it attains all values $(1+r^p)^{s/p}$, each of them many times, e.g. for $k=\pm re_j, j=1,...,d$.
- Let
$$N(r,d) := card\{k\in\mathbb{Z}^d : \sum_{j=1}^d |k_j|^p \le r^p\}, r\in\mathbb{N}.$$

Lemma

If $N(r-1,d) < n \le N(r,d)$, *then*

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

- In principle, this gives $a_n(I_d)$ for all n. But, unless $p=1$, the exact computation of the cardinalities $N(r,d)$ is impossible. The hard work is to find good estimates, using e.g. volume or entropy arguments.

Asymptotic constants

- Let B_p^d denote the unit ball in $(\mathbb{R}^d, \|.\|_p)$.

Theorem (KSU 2014 and KSU 2015)

Let $0 < s, p < \infty$ *and* $d \in \mathbb{N}$. *Then*

$$\lim_{n\to\infty} n^{s/d} a_n(I_d : H^{s,p}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) = vol(B_p^d)^{s/d} \sim d^{-s/p}$$

and

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

Some comments

- Isotropic case: The asymptotic constant is of order
  $d^{-s/2}$ for the natural norm ($p = 2$),
  $d^{-1/2}$ for the modified natural norm ($p = 2s$) .

  This gives the correct order $n^{-s/d}$ of a_n as $n \to \infty$ and the
  exact decay rate $d^{-s/p}$ of the constants as $d \to \infty$.

- Mixed case: It is interesting that the limit is independent on p.

- The constants decay
  – polynomially in the isotropic case, and even
  – super-exponentially in the mixed case, roughly like $\left(\frac{2e}{d}\right)^{sd}$.

  This helps in error estimates!

Estimates for large n

Next step: Estimates of a_n for "large" n.

Theorem (KSU 2014, isotropic case, $p = 1$)

Let $s > 0$ and $n \geq 6^d/3$. Then

$$d^{-s} n^{-s/d} \leq a_n(I_d : H^{s,1}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \leq (4e)^s d^{-s} n^{-s/d}\,.$$

- Remember: The asymptotic constant is of order d^{-s} as $d \to \infty$. This d-dependence d^{-s} is reflected in the above estimates!
- Proof: via combinatorial estimates of the cardinalities $N(r, d)$
- We have similar estimates for all other $0 < p < \infty$,

  and also in the mixed case $I_d : H^{s,p}_{mix}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)$

Preasymptotic estimates – isotropic case

Theorem (KSU 2014)

Let $s > 0$, $p = 1$ *and* $2 \leq n \leq 2^d$. *Then*

$$\left(\frac{1}{2 + \log_2 n}\right)^s \leq a_n(I_d : H^{s,1}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \leq \left(\frac{\log_2(1+2d)}{\log_2 n}\right)^s .$$

- Proof by combinatorial arguments, which only work for $p = 1$. Using a relation to entropy numbers, we could close the gap between lower and upper bounds and treat arbitrary p's.

Theorem (KMU 2015)

Let $0 < s, p < \infty$ *and* $2 \leq n \leq 2^d$. *Then*

$$a_n(I_d : H^{s,p}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \sim_{s,p} \left(\frac{\log_2(1 + d/\log_2 n))}{\log_2 n}\right)^{s/p} .$$

Preasymptotic estimates – mixed case

Theorem (KSU 2015)

Let $s > 0$, $d \geq 2$ and $8 \leq n \leq d\,2^{2d-1}$. Then

$$a_n(I_d : H^{s,1}_{mix}(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \leq \left(\frac{e^2}{n}\right)^{\frac{s}{2+\log_2 d}}.$$

- The bound is non-trivial (i.e. < 1) for all n in the given range, since $\frac{e^2}{n} \leq \frac{e^2}{8} = 0.9236... < 1$.
- We have similar lower estimates. They are not matching but they show, that one has to wait exponentially long until one can "see" the correct asymptotic rate n^{-s}, ignoring the log-terms.

Gevrey classes

- Introduced already in 1918, since then many applications in PDEs.

Definition (Maurice Gevrey, 1918)

Let $\sigma > 1$. A function $f \in C^\infty(\mathbb{R}^d)$ belongs to the Gevrey class $\mathbf{G}^\sigma(\mathbb{R}^d)$, if for every compact subset $K \subset \mathbb{R}^d$ there are constants $C = C(K) > 0$ and $R = R(K) > 0$ such that for all multi-indices $\alpha \in \mathbb{N}_0^d$

$$\sup_{x \in K} |D^\alpha f(x)| \le C \cdot R^{\alpha_1 + \ldots + \alpha_d} \cdot (\alpha_1! \cdots \alpha_d!)^\sigma .$$

- The Gevrey classes $\mathbf{G}^\sigma(\mathbb{R}^d)$ are linear spaces, but not normed spaces.
- $\sigma = 1$: All functions in $\mathbf{G}^\sigma(\mathbb{R}^d)$ are analytic.
- $\sigma > 1$: $\mathbf{G}^\sigma(\mathbb{R}^d)$ contains non-analytic functions.
- For periodic $f : \mathbb{R}^d \to \mathbb{C}$, the growth conditions on the derivatives can be rephrased it terms of Fourier coefficients.

Periodic Gevrey spaces

Definition

Let $0 < s < 1$ and $c > 0$. The periodic Gevrey space $G^{s,c}(\mathbb{T}^d)$ consists of all C^∞-functions $f : \mathbb{T}^d \to \mathbb{C}$ such that

$$\|f|G^{s,c}(\mathbb{T}^d)\| := \Big(\sum_{k\in\mathbb{Z}^d} \underbrace{\exp(c\,\|k\|_1^s)^2}_{=w(k)} |c_k(f)|^2 \Big)^{1/2} < \infty\,.$$

- For periodic f: $\quad f \in \mathbf{G}^\sigma(\mathbb{R}^d) \iff \exists\, c > 0 : f \in G^{1/\sigma,c}(\mathbb{T}^d)\,.$

Theorem (KP2015)

Let $0 < s < 1$, $c > 0$ *and* $d \in \mathbb{N}$. *Then*

$$\lim_{n\to\infty} \frac{a_n(I_d : G^{s,c}(\mathbb{T}^d) \to L_2(\mathbb{T}^d))}{\exp\big(-2^{-s} c\,(d!\,n)^{s/d}\big)} = 1$$

- We also have preasymptotic estimates, as well as estimates for large n.

From L_2-approximation to L_∞-approximation

- Recall: $F_d(w) \hookrightarrow L_\infty(\mathbb{T}^d) \iff \sum_{k\in\mathbb{Z}^d} 1/w(k)^2 < \infty$ In this case:

Theorem (CKS2016)

$$a_n(I_d : F_d(w) \to L_\infty(\mathbb{T}^d)) = \left(\sum_{j=n}^{\infty} a_j(I_d : F_d(w) \to L_2(\mathbb{T}^d))^2 \right)^{1/2}.$$

- This can be applied to Sobolev and Gevrey embeddings

$$H^{s,p}(\mathbb{T}^d) \hookrightarrow L_\infty(\mathbb{T}^d) \iff s > d/2$$

$$H^{s,p}_{mix}(\mathbb{T}^d) \hookrightarrow L_\infty(\mathbb{T}^d) \iff s > 1/2$$

$$G^{s,c}(\mathbb{T}^d) \hookrightarrow L_\infty(\mathbb{T}^d) \quad \forall\, 0 < s < 1, c > 0$$

We get optimal asymptotic constants for these embeddings.

- Open problem: Preasymptotics for Sobolev embeddings?

Information-based complexity (IBC)

- Consider the approximation problem

$$I_d : F_d(\mathbb{T}^d) \to L_2(\mathbb{T}^d), \quad d \in \mathbb{N}.$$

- Optimal worst case error of linear algorithms
(using n pieces of arbitraray linear information)

$$err_n^{wor}(I_d) = a_{n+1}(I_d)$$

- Information complexity

$$n(\varepsilon, d) := \min\{n \in \mathbb{N} : a_{n+1}(I_d) \le \varepsilon\}$$

- What is the behaviour of $n(\varepsilon, d)$ as $d \to \infty$ and/or $\varepsilon \to 0$?
- Tractability notions in IBC classify this behaviour

Tractability notions

- polynomial $n(\varepsilon, d) \leq C\,\varepsilon^{-p} d^q$
- quasi-polynomial $\ln n(\varepsilon, d) \leq C(1 + \ln \frac{1}{\varepsilon})(1 + \ln d)$
- uniformly weakly tractable $\lim\limits_{1/\varepsilon + d \to \infty} \frac{\ln n(\varepsilon, d)}{(1/\varepsilon + d)^q} = 0$ for all $q > 0$
- weakly tractable $\lim\limits_{1/\varepsilon + d \to \infty} \frac{\ln n(\varepsilon, d)}{1/\varepsilon + d} = 0$
- intractable = not weakly tractable
- curse of dimension $n(\varepsilon_0, d) \geq C^d$
  for some $\varepsilon_0 > 0$ and $C > 1$
  and infinitely many $d \in \mathbb{N}$

Tractability of our approximation problems

Our estimates can be translated into tractability results.

Theorem

None of the approximation problems $I_d : F_d(\mathbb{T}^d) \to L_2(\mathbb{T}^d), \quad d \in \mathbb{N}$, *with*

$$F_d(\mathbb{T}^d) = H^{s,p}(\mathbb{T}^d) \quad \text{or} \quad H^{s,p}_{mix}(\mathbb{T}^d) \quad \text{or} \quad G^{s,c}(\mathbb{T}^d)$$

suffers from the curse of dimensionality.

Theorem (isotropic Sobolev spaces)

Let $0 < s, p < \infty$. *The approximation problem*

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

is weakly tractable, if $s > p$, *and intractable, if* $s \leq p$.

Tractability - continued

Theorem (mixed Sobolev spaces)

For all $0 < s, p < \infty$ *the approximation problem*

$$I_d : H^{s,p}_{mix}(\mathbb{T}^d) \to L_2(\mathbb{T}^d) \quad , d \in \mathbb{N},$$

is quasi-polynomially tractable, but not polynomially tractable.

Theorem (Gevrey spaces)

For all $0 < s < 1$ *and* $c > 0$ *the approximation problem*

$$I_d : G^{s,c}(\mathbb{T}^d) \to L_2(\mathbb{T}^d) \quad , d \in \mathbb{N},$$

is uniformly weakly tractable, but not quasi-polynomially tractable.

Final remarks

An interesting observation

- Gevrey spaces consist of C^∞-functions and are much smaller than the mixed spaces, which are only of finite smoothness.
- The decay rate of the approximation numbers a_n as $n \to \infty$ is
  – subexponential for Gevrey embeddings
  – polynomial for mixed-space embeddings
- But the tractability of the mixed-space embeddings is better. This is due to the preasymptotic behaviour.

Thank you for your attention!