Function space emdeddings for non-tensor product spaces and application to high-dimensional approximation

Michael Gnewuch

University of Osnabrück, Germany

August 23, 2024

Based on ongoing joint work with

Peter Kritzer (RICAM Linz) & Klaus Ritter (RPTU Kaiserslautern)

Motivation: Tractability Analysis

Of interest: Approximation problems, defined on scale of function spaces $(H_d)_{d\in\mathbb{N}}$, where d refers to number of variables.

Important goal: To find algorithms that break curse of dimensionality, preferably for large classes of scales of function spaces.

An approach to achieve this goal:

- Exploit structure of suitable (scales of) function spaces for analysis of promising algorithms.
- Establish embedding theorems for scales of function spaces to transfer results from scales with favourable structure to other scales of interest.

Already known: Several general embedding results if H_d is d-fold tensor product of reproducing kernel Hilbert space (RKHS) H_1; examples include weighted RKHSs based on product weights and spaces of increasing smoothness, see Hefter & Ritter'14; G., Hefter, Hinrichs, Ritter'17; G., Hefter, Hinrichs, Ritter, Wasilkowski'19; G., Hefter, Hinrichs, Ritter'22; G., Hinrichs, Ritter, Rüßmann'24,...

Outline of this talk: We discuss new general embedding approach for non-tensor product case and apply it to prove new results for L^∞-approximation on RKHS of functions that depend on infinitely many variables ($=$ limiting case of tractability analysis).

Function Spaces of Infinitely Many Variables

Spaces of univariate functions: D domain, $k : D \times D \to \mathbb{R}$ bounded reproducing kernel (RK) with $1 \notin H(k)$. Then $H(1+k) = H(1) \oplus H(k)$.

$\kappa := \sup_{x \in D} \sqrt{k(x,x)} < \infty$.

Spaces of multivariate functions: For $u \in \mathcal{U} := \{u \subset \mathbb{N} \, : \, |u| < \infty\}$ put

$$k_u(\mathbf{x}_u, \mathbf{y}_u) := \prod_{j \in u} k(x_j, y_j), \qquad \mathbf{x}_u = (x_j)_{j \in u}, \mathbf{y}_u = (y_j)_{j \in u} \in D^u.$$

Then $H(k_u) = \bigotimes_{j \in u} H(k)$.

Spaces of ∞-variate functions: $(\gamma_u)_{u \in \mathcal{U}}$ weights satisfying

$$\sum_{u \in \mathcal{U}} \gamma_u \kappa^{2|u|} < \infty \qquad \text{(summability condition)},$$

$$K_{\boldsymbol{\gamma}}(\mathbf{x}, \mathbf{y}) := \sum_{u \in \mathcal{U}} \gamma_u k_u(\mathbf{x}_u, \mathbf{y}_u), \qquad \mathbf{x}, \mathbf{y} \in D^{\mathbb{N}}.$$

Then $f \in H(K_{\boldsymbol{\gamma}})$ has orthogonal function decomposition

$$f(\mathbf{x}) = \sum_{u \in \mathcal{U}} f_u(\mathbf{x}_u), \qquad f_u \in H(k_u), \qquad \text{with} \qquad \|f\|^2_{K_{\boldsymbol{\gamma}}} = \sum_{u \in \mathcal{U}} \gamma_u^{-1} \|f_u\|^2_{k_u}.$$

Examples of Classes of Weights

$(\gamma_j)_{j\in\mathbb{N}}$ non-increasing sequence in $]0,\infty[$.

(i) Product weights [Sloan & Woźniakowski'98]:

$$\gamma_u := \prod_{j\in u} \gamma_j \,.$$

(ii) Product and order dependent (POD) weights [Kuo, Sloan & Schwab'12] :

$$\gamma_u := \Gamma_{|u|} \prod_{j\in u} \gamma_j \,,$$

where $1 \leq \Gamma_k \leq c \cdot (k!)^b$ for $c > 0$ and $b < \mathrm{dec}((\gamma_j)_{j\in\mathbb{N}})$, where

$$\mathrm{dec}((\gamma_j)_{j\in\mathbb{N}}) := \sup\left\{ t \geq 0 \,:\, \sum_{j\in\mathbb{N}} \gamma_j^{1/t} < \infty \right\} \quad \text{("polynomial decay rate")}$$

(Example: $\gamma_j \asymp j^{-\alpha}\ln(j)^\beta \implies \mathrm{dec}((\gamma_j)_{j\in\mathbb{N}}) = \alpha\,.$)

Lemma 1a. γ *product or POD weights,* $C > 0$*. Then*

$$\sum_{u\in\mathcal{U}} \gamma_u C^{2|u|} < \infty \iff \sum_{u\in\mathcal{U}} \gamma_u < \infty \iff \sum_{j\in\mathbb{N}} \gamma_j < \infty\,.$$

Examples of Classes of Weights

(iii) Finite-order weights of order $\omega \in \mathbb{N}$ [Dick, Sloan, Wang & Woź.'06]:

$$\gamma_u = 0 \quad \text{for all } u \in \mathcal{U} \text{ with } |u| > \omega.$$

Lemma 1b. $\boldsymbol{\gamma}$ *finite order weights of order* ω, $C > 0$. *Then*

$$\sum_{u \in \mathcal{U}} \gamma_u C^{2|u|} < \infty \quad \Longleftrightarrow \quad \sum_{u \in \mathcal{U}} \gamma_u < \infty.$$

Remark. For summable (general) weights $\boldsymbol{\gamma}$:

$$H(K_{\boldsymbol{\gamma}}) = \bigotimes_{j \in \mathbb{N}} H(1 + \gamma_j k) \quad \Longleftrightarrow \quad \boldsymbol{\gamma} \quad \text{are product weights.}$$

Approximation Problem: Solution Operator, Algorithms & Cost

Solution operator: $B(D^{\mathbb{N}})$ space of bounded fcts. on $D^{\mathbb{N}}$ with norm $\|\cdot\|_\infty$.

$$S_{\boldsymbol{\gamma}} : H(K_{\boldsymbol{\gamma}}) \to B(D^{\mathbb{N}})\,,\ f \mapsto f\,,$$

$$\|S_{\boldsymbol{\gamma}}\| = \sup_{\mathbf{x}\in D^{\mathbb{N}}} \sqrt{K_{\boldsymbol{\gamma}}(\mathbf{x},\mathbf{x})} = \sum_{u\in\mathcal{U}} \gamma_u \kappa^{2|u|} < \infty\,.$$

Admissable algorithms: Linear algorithms

$$A_n(f) = \sum_{i=1}^{n} f(\mathbf{x}^{(i)}) g_i\,, \quad \mathbf{x}^{(i)} \in D^{\mathbb{N}}\,,\ g_i \in B(D^{\mathbb{N}})\,.$$

Cost: Fix default value $a \in D$.

Unrestricted Subspace Sampling [Kuo, Sloan, Wasilkowski, Woźniakowski'10a]

$$\mathrm{cost}(A_n) := \sum_{i=1}^{n} |\{j \in \mathbb{N} : x_j^{(i)} \neq a\}|$$

Approximation Problem: Error Criterion

Error criterion: Worst-case error

$$e(A_n; S_{\boldsymbol{\gamma}}, K_{\boldsymbol{\gamma}}) := \sup_{\|f\|_{K_{\boldsymbol{\gamma}}} \leq 1} \|A(f) - f\|_\infty$$

n-th minimal randomized error:

$$e(n, K_{\boldsymbol{\gamma}}) := \inf\{e(A; S_{\boldsymbol{\gamma}}, K_{\boldsymbol{\gamma}}) \,:\, A \text{ adm. alg.., } \operatorname{cost}(A) \leq n\}$$

"(Polynomial) Convergence rate" of $e(n, K_{\boldsymbol{\gamma}})$:

$$\operatorname{dec}(K_{\boldsymbol{\gamma}}) := \operatorname{dec}\left((e(n, K_{\boldsymbol{\gamma}}))_{n\in\mathbb{N}}\right)$$

Analysis of Algorithms

Task: Analyze the error of a promising adm. algorithm A on $H(K_{\boldsymbol{\gamma}})$.

Problem: Norm on $H(K_{\boldsymbol{\gamma}})$ may not be well-suited for error analysis.

Remedy: Find RK $\widehat{k}$ on D and weights $\boldsymbol{\gamma}^{\uparrow}$, s.t. corresponding kernel $\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}}$ on $D^{\mathbb{N}}$ induces norm on $H(\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}})$ appropriate for error analysis and

$$H(K_{\boldsymbol{\gamma}}) \hookrightarrow H(\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}}).$$

How to Find a Suitable Embedding Space $H(\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}})$?

Step I: Choose suitable kernel $\widehat{k}$ on D.

Take $\widehat{k}$ with $H(1) \cap H(\widehat{k}) = \{0\}$ and $H(1+k) = H(1+\widehat{k})$ as vector spaces that has features supporting the error analysis.

Example 1. For $a \in D$ exists *always* kernel $\widehat{k}$ s.t. $\widehat{k}(a,a) = 0$ ("anchor condition") and $H(1+k) = H(1+\widehat{k})$ as vector spaces.

$\boldsymbol{\gamma}^{\uparrow}$ weights with $\sum_{u \in \mathcal{U}} \gamma_u^{\uparrow} \widehat{\kappa}^{2|u|} < \infty$, where $\widehat{\kappa} := \sup_{x \in D} \sqrt{\widehat{k}(x,x)}$.
Then $\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}}(\mathbf{a},\mathbf{a}) = 0$ and decomposition

$$f = \sum_{u \subset_f \mathbb{N}} f_u \,, \quad f_u \in H(\widehat{k}_u),$$

on $H(\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}})$ is anchored decomposition (or cut HDMR) of f.

It can be calculated directly as

$$f_u(\mathbf{x}) := \sum_{v \subseteq u} (-1)^{|u \setminus v|} f(\mathbf{x}_v, \mathbf{a}_{\mathbb{N} \setminus v}),$$

cf. [Kuo, Sloan, Wasilkowski, Woźniakowski'10b].

Anchored decomposition is, e.g., helpful if function decomposition should be explicitly calculated or if one wants to use Taylor expansion.

How to Find a Suitable Embedding Space $H(\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}})$?

Example 2. Let μ probability measure on D and k measurable. There exists *always* RK $\widehat{k}$ s.t.

$$\int_D \widehat{k}(x,y)\, d\mu(y) = 0 \quad \text{for all } x \in D. \quad (\text{“ANOVA condition”})$$

and $H(1+k) = H(1+\widehat{k})$ as vector spaces.

$\boldsymbol{\gamma}^{\uparrow}$ weights with $\sum_{u\in\mathcal{U}} \gamma_u^{\uparrow} \widehat{\kappa}^{2|u|} < \infty$, where $\widehat{\kappa} := \sup_{x\in D} \sqrt{\widehat{k}(x,x)}$.
Then $H(\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}}) \subset L^2(D^{\mathbb{N}}, \mu^{\mathbb{N}})$ and decomposition

$$f = \sum_{u\subset_f \mathbb{N}} f_u\,, \quad f_u \in H(\widehat{k}_u),$$

on $H(\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}})$ is (∞-variate) ANOVA decomposition of f.

In particular,

$$\|f\|^2_{L^2(D^{\mathbb{N}},\mu^{\mathbb{N}})} = \sum_{u\in\mathcal{U}} \|f_u\|^2_{L^2(D^u,\mu^u)} \quad \text{and} \quad \mathrm{Var}(f) = \sum_{u\in\mathcal{U}} \mathrm{Var}(f_u).$$

ANOVA decomposition is, e.g., helpful for L^2-approximation or for analysis of unbiased randomized algorithms (error = variance).

How to Find a Suitable Embedding Space $H(\widehat{K}_{\boldsymbol{\gamma}^\uparrow})$?

Step II: Choose suitable weights $\boldsymbol{\gamma}^\uparrow$.

Let C denote norm of embedding $H(k) \hookrightarrow H(1+\widehat{k})$.
Assume $\boldsymbol{\gamma}$ satisfies

$$\sum_{v \in \mathcal{U}} \gamma_v C^{2|v|} < \infty.$$

Define

$$\gamma_u^\uparrow := \sum_{u \subseteq v \in \mathcal{U}} \gamma_v C^{2|v|}$$

Theorem 1. *Let $\widehat{K}_{\boldsymbol{\gamma}^\uparrow}$ weighted RK on $D^{\mathbb{N}}$ corresponding to $\widehat{k}$ and $\boldsymbol{\gamma}^\uparrow$. Then*

$$H(K_{\boldsymbol{\gamma}}) \hookrightarrow H(\widehat{K}_{\boldsymbol{\gamma}^\uparrow})$$

with embedding norm at most 1.

Lemma 1. *If $\boldsymbol{\gamma}$ are product weights, POD weights, or finite-order weights then*

$$\mathrm{dec}(\boldsymbol{\gamma}) = \mathrm{dec}(\boldsymbol{\gamma}^\uparrow).$$

How to Attack the L^∞-Approximation Problem?

For given RK k and weights $\boldsymbol{\gamma}$ as above assume k anchored in default value $a \in D$. Resulting anchored space $H(K_{\boldsymbol{\gamma}})$ well-suited for analysis of

Multivariate decomposition method (MDM):

For $f \in H(K_{\boldsymbol{\gamma}})$ denote its anchored decomposition by

$$f(\mathbf{x}) = \sum_{u \in \mathcal{U}} f_{u,a}(\mathbf{x}_u).$$

1. For given error tolerance ε choose finite set $\mathcal{A}(\varepsilon) \subset \mathcal{U}$ of most important groups of variables.
2. Choose for each $u \in \mathcal{A}(\varepsilon)$ linear algorithm A_{u,n_u} using n_u samples to approximate $S(f_{u,a}) = f_{u,a}$.

Final algorithm A^{MDM} is of form

$$A^{\mathrm{MDM}}(f) = \sum_{u \in \mathcal{A}(\varepsilon)} A_{u,n_u}(f_{u,a})$$

Multivariate Decomposition Method

MDM *fka* changing dimension algorithm for ∞-variate setting introduced in

- Kuo, Sloan, Wasilkowski &Woźniakowski'10a; Plaskota & Wasilkowski'11 (∞-variate integration)
- Wasilkowski & Woźniakowski'11a, '11b (∞-variate approximation).

Further papers on MDM include: Wasilkowski'12; G.'13; Dick & G.'14a, 14b; Plaskota & Wasilkowski'14; Wasilkowski'14; Gilbert & Wasilkowski'17; G., Hefter, Hinrichs & Ritter'17; G., Hefter, Hinrichs, Ritter & Wasilkowski'17; Kuo, Nuyens, Plaskota, Sloan, Wasilkowski'17; Gilbert, Kuo, Nuyens & Wasilkowski'18; G., Hefter, Hinrichs, Ritter & Wasilkowski'19; G. & Wnuk'20; G., Hinrichs, Ritter & Rüßmann'24...

Similar idea used for multivariate integration in Griebel & Holtz'10 (dimension-wise quadrature methods).

Result for Anchored Kernels and POD Weights

"Clean result" in the case of POD weights (we have also results for "general weights", but they are a bit more technical...):

Theorem 2. *Assume $\boldsymbol{\gamma}$ are summable POD weights. Let k be bounded RK anchored in $a \in D$ and $K_{\boldsymbol{\gamma}}$ corresponding weighted RK. Then*

$$\mathrm{dec}(K_{\boldsymbol{\gamma}}) = \min\left(\mathrm{dec}(1+k), \frac{\mathrm{dec}(\boldsymbol{\gamma})-1}{2}\right),$$

where $\mathrm{dec}(1+k)$ denotes convergence rate of nth minimal errors of univariate L^∞-approximation on $H(1+k)$.

Remark. (i) *Upper bound on $\mathrm{dec}(K^a_{\boldsymbol{\gamma}})$ can be derived from corresponding bound for integration problem for product weights from Kuo, Sloan, Wasilkowski, Woźiakowski'10a.*
(ii) *In special case where $H(K_{\boldsymbol{\gamma}})$ is Sobolev space of dominating mixed smoothness $r \in \mathbb{N}$ the lower bound on $\mathrm{dec}(K^a_{\boldsymbol{\gamma}})$was proved in Wasilkowski'14; in that case $\mathrm{dec}(1+k) = r - 1/2$. (He actually considered L^q-approximation for L^p-Sobolev spaces for $1 \le p, q \le \infty$.)*

Result for Anchored Kernels and POD Weights

For given bounded RK k and summable POD weights $\boldsymbol{\gamma}$ choose corresponding RK $\widehat{k}$ anchored in default value $a \in D$ and corresponding weights $\boldsymbol{\gamma}^{\uparrow}$.

Resulting space $H(K_{\boldsymbol{\gamma}^{\uparrow}})$ is anchored in $\mathbf{a}$, but weights $\boldsymbol{\gamma}^{\uparrow}$ are not necessarily POD weights anymore... but there exist POD weights η, ζ s.t.

$$\eta_u \leq \gamma_u^{\uparrow} \leq \zeta_u \quad \text{for all } u \in \mathcal{U} \quad \text{and} \quad \operatorname{dec}(\eta) = \operatorname{dec}(\boldsymbol{\gamma}^{\uparrow}) = \operatorname{dec}(\zeta).$$

Corollary 1. *Assume $\boldsymbol{\gamma}$ are POD weights. Let k be bounded RK and $K_{\boldsymbol{\gamma}}$ corresponding weighted RK. Then*

$$\operatorname{dec}(K_{\boldsymbol{\gamma}}) = \min\left(\operatorname{dec}(1+k), \frac{\operatorname{dec}(\boldsymbol{\gamma}) - 1}{2}\right),$$

where $\operatorname{dec}(1+k)$ denotes convergence rate of nth minimal errors of univariate L^{∞}-approximation on $H(1+k)$.

Proof: Theorem 2, applied to $H(\widehat{K}_{\zeta})$, and $H(K_{\boldsymbol{\gamma}}) \hookrightarrow H(\widehat{K}_{\boldsymbol{\gamma}^{\uparrow}}) \hookrightarrow H(\widehat{K}_{\zeta})$ yield the lower bound on $\operatorname{dec}(K_{\boldsymbol{\gamma}})$.
Upper bound on $\operatorname{dec}(K_{\boldsymbol{\gamma}})$ can be derived from corresponding upper bound for integration for corresponding product weights in G., Hefter, Hinrichs, Ritter'17.

Appendix: Some Properties of the Manipulated Weights $\boldsymbol{\gamma}^{\uparrow}$

For $u \in \mathcal{U}$ define Difference Operator Δ_u via

$$\Delta_u \boldsymbol{\gamma} := \gamma_v - \gamma_{u \cap v} \quad \text{for all weights } \boldsymbol{\gamma},\ v \in \mathcal{U}.$$

Weights $\boldsymbol{\gamma}$ are *completely monotone* if

$$(\Delta_{u_n} \Delta_{u_{n-1}} \cdots \Delta_{u_1} \boldsymbol{\gamma})_v \geq 0$$

for every $n \in \mathbb{N}$, all $u_1, \ldots, u_n \in \mathcal{U}$, and all $v \in \mathcal{U}$.

Assumption: $(\gamma_u)_{u \in \mathcal{U}}$ weights satisfying $\sum_{u \in \mathcal{U}} \gamma_u \kappa^{2|u|} < \infty$

Then $\boldsymbol{\gamma}^{\uparrow}$ are completely monotone.

Define $\boldsymbol{\gamma}^{\downarrow}$ via

$$\gamma_u^{\downarrow} := C^{-2|u|} \cdot \lim_{s \to \infty} \sum_{u \subseteq v \in \mathcal{U}} (-1)^{|v|-|u|} \gamma_v\,, \quad u \in \mathcal{U}.$$

Then

$$((\gamma^{\uparrow})^{\downarrow})_u = \gamma_u \quad \text{for all } u \in \mathcal{U}.$$