Integration and L_2-Approximation of Functions of Infinitely Many Variables

Klaus Ritter

TU Kaiserslautern

Joint work with

M. Gnewuch, M. Hefter, A. Hinrichs, G. W. Wasilkowski

Introduction

Computational problem: For a class F of functions

$$f : E \to \mathbb{K},$$

integrate/approximate f based on a finite number of function values.

In this talk

- $E = D^{\mathbb{N}}$ with $D \subseteq \mathbb{R}$, i.e., f depends on the variables $y_1, y_2, \dots \in D$,
- increasing smoothness w.r.t. these variables,
- worst case analysis in a Hilbert space setting.

Introduction

Computational problem: For a class F of functions

$$f : E \to \mathbb{K},$$

integrate/approximate f based on a finite number of function values.

In this talk

- $E = D^{\mathbb{N}}$ with $D \subseteq \mathbb{R}$, i.e., f depends on the variables $y_1, y_2, \dots \in D$,
- increasing smoothness w.r.t. these variables,
- worst case analysis in a Hilbert space setting.

Integration and approximation of functions on $D^{\mathbb{N}}$

Hickernell, Müller-Gronbach, Niu, R (2010),
Kuo, Sloan, Wasilkowski, Woźniakowski (2010),
Baldeaux, Dick, Dũng, Gilbert, Gnewuch, Griebel, Hefter,
Hinrichs, Nuyens, Oswald, Plaskota, **. . .**

Random (parametric) PDEs

. . .

Introduction

Computational problem: For a class F of functions

$$f : E \to \mathbb{K},$$

integrate/approximate f based on a finite number of function values.

In this talk

- $E = D^{\mathbb{N}}$ with $D \subseteq \mathbb{R}$, i.e., f depends on the variables $y_1, y_2, \dots \in D$,
- increasing smoothness w.r.t. these variables,
- worst case analysis in a Hilbert space setting.

OUTLINE

I. The Function Spaces
II. Algorithms, Error, and Cost
III. Results and Remarks
IV. Embeddings and $\underbrace{\text{Optimal Algorithms}}_{\rightsquigarrow \textit{Dirk's talk this afternoon}}$

I. The Function Spaces

Consider

- the trigonometric basis $(e_\nu)_{\nu\in\mathbb{Z}}$ of $L_2([0,1])$,
- **smoothness parameters** $0 < r_1 \le r_2 \le \dots$.

For $j \in \mathbb{N}$ and $\nu \in \mathbb{Z}$ we define **Fourier weights**

$$\alpha_{\nu,j} = (1 + |\nu|)^{r_j}$$

and the corresponding scale of **Korobov spaces**

$$H_j = \{f \in L_2([0,1]) : \sum_{\nu\in\mathbb{Z}} \alpha_{\nu,j} \cdot \left|\langle f, e_\nu\rangle_{L_2([0,1])}\right|^2 < \infty\},$$

$$\langle f, g\rangle_{H_j} = \sum_{\nu\in\mathbb{Z}} \alpha_{\nu,j} \cdot \langle f, e_\nu\rangle_{L_2([0,1])} \cdot \langle e_\nu, g\rangle_{L_2([0,1])}.$$

I. The Function Spaces

Consider

- the trigonometric basis $(e_\nu)_{\nu \in \mathbb{Z}}$ of $L_2([0,1])$,
- **smoothness parameters** $0 < r_1 \leq r_2 \leq \dots$.

For $j \in \mathbb{N}$ and $\nu \in \mathbb{Z}$ we define **Fourier weights**

$$\alpha_{\nu,j} = (1 + |\nu|)^{r_j}$$

and the corresponding scale of **Korobov spaces**

$$H_j = \{f \in L_2([0,1]) : \sum_{\nu \in \mathbb{Z}} \alpha_{\nu,j} \cdot \left|\langle f, e_\nu \rangle_{L_2([0,1])}\right|^2 < \infty\},$$

$$\langle f, g \rangle_{H_j} = \sum_{\nu \in \mathbb{Z}} \alpha_{\nu,j} \cdot \langle f, e_\nu \rangle_{L_2([0,1])} \cdot \langle e_\nu, g \rangle_{L_2([0,1])}.$$

Remark $H_j \hookleftarrow H_{j+1}$ compact $\Leftrightarrow$ $r_j < r_{j+1}$.

I. The Function Spaces

Consider

- the trigonometric basis $(e_\nu)_{\nu \in \mathbb{Z}}$ of $L_2([0,1])$,
- **smoothness parameters** $0 < r_1 \leq r_2 \leq \dots$.

For $j \in \mathbb{N}$ and $\nu \in \mathbb{Z}$ we define **Fourier weights**

$$\alpha_{\nu,j} = (1 + |\nu|)^{r_j}$$

and the corresponding scale of **Korobov spaces**

$$H_j = \{f \in L_2([0,1]) : \sum_{\nu \in \mathbb{Z}} \alpha_{\nu,j} \cdot \left| \langle f, e_\nu \rangle_{L_2([0,1])} \right|^2 < \infty\},$$

$$\langle f, g \rangle_{H_j} = \sum_{\nu \in \mathbb{Z}} \alpha_{\nu,j} \cdot \langle f, e_\nu \rangle_{L_2([0,1])} \cdot \langle e_\nu, g \rangle_{L_2([0,1])}.$$

Remark $H_j \hookleftarrow H_{j+1}$ compact $\Leftrightarrow$ $r_j < r_{j+1}$.

Finally, based on the unit vectors $e_0 = 1$,

$$H = \bigotimes_{j \in \mathbb{N}} H_j.$$

Consider $L_2([0,1]^{\mathbb{N}})$ w.r.t the product μ of the uniform distribution μ_0 on $[0,1]$. Put

$$\varrho = \liminf_{j\to\infty} \frac{r_j}{\ln(j)} \in [0,\infty].$$

Lemma $\varrho > 0 \;\Rightarrow\; H \hookrightarrow L_2([0,1]^{\mathbb{N}})$ compact.

Consider $L_2([0,1]^{\mathbb{N}})$ w.r.t the product μ of the uniform distribution μ_0 on $[0,1]$. Put

$$\varrho = \liminf_{j\to\infty} \frac{r_j}{\ln(j)} \in [0,\infty].$$

Lemma $\varrho > 0 \;\Rightarrow\; H \hookrightarrow L_2([0,1]^{\mathbb{N}})$ compact.

A Hilbert space F is a RKHS on a domain $E \neq \emptyset$ if

- $F \subseteq \mathbb{K}^E$,
- $F \to \mathbb{K}$: $f \mapsto f(\mathbf{x})$ is continuous for every $\mathbf{x} \in E$.

Consider $L_2([0,1]^{\mathbb{N}})$ w.r.t the product μ of the uniform distribution μ_0 on $[0,1]$. Put

$$\varrho = \liminf_{j\to\infty} \frac{r_j}{\ln(j)} \in [0,\infty].$$

Lemma $\varrho > 0 \;\Rightarrow\; H \hookrightarrow L_2([0,1]^{\mathbb{N}})$ compact.

A Hilbert space F is a RKHS on a domain $E \neq \emptyset$ if

- $F \subseteq \mathbb{K}^E$,
- $F \to \mathbb{K}$: $f \mapsto f(\mathbf{x})$ is continuous for every $\mathbf{x} \in E$.

Lemma H_j RKHS on $[0,1] \;\Leftrightarrow\; r_j > 1$.

Lemma

$$(r_1 > 1 \wedge \varrho \ln(2) > 1) \;\Rightarrow\; H \text{ RKHS on } [0,1]^{\mathbb{N}} \;\Rightarrow\; (r_1 > 1 \wedge \varrho \ln(2) \geq 1).$$

II. Algorithms, Error, and Cost

Briefly consider

$$H^d = \bigotimes_{j=1}^{d} H_j.$$

Common assumption, if H^d is a RKHS:

- Evaluation of $f \in H^d$ at any $\mathbf{x} \in [0,1]^d$ at cost one.

II. Algorithms, Error, and Cost

Briefly consider

$$H^d = \bigotimes_{j=1}^{d} H_j.$$

Common assumption, if H^d is a RKHS:

- Evaluation of $f \in H^d$ at any $\mathbf{x} \in [0,1]^d$ at cost one.

Actually, we consider

$$H = \bigotimes_{j \in \mathbb{N}} H_j.$$

Assumption, if H is a RKHS:

- Fix $a \in [0,1]$. Evaluation of $f \in H$ at any $\mathbf{x} \in [0,1]^{\mathbb{N}}$ with

  $$\mathrm{active}(\mathbf{x}) = \left|\{j \in \mathbb{N} : \mathbf{x}_j \neq a\}\right| < \infty$$

  at cost active($\mathbf{x}$).

II. Algorithms, Error, and Cost

Briefly consider

$$H^d = \bigotimes_{j=1}^{d} H_j.$$

Common assumption, if H^d is a RKHS:

- Evaluation of $f \in H^d$ at any $\mathbf{x} \in [0,1]^d$ at cost one.

Actually, we consider

$$H = \bigotimes_{j \in \mathbb{N}} H_j.$$

Assumption, if H is a RKHS:

- Fix $a \in [0,1]$. Evaluation of $f \in H$ at any $\mathbf{x} \in [0,1]^{\mathbb{N}}$ with

$$\text{active}(\mathbf{x}) = \left|\{j \in \mathbb{N} : \mathbf{x}_j \neq a\}\right| < \infty$$

at cost active($\mathbf{x}$). See *Kuo, Sloan, Wasilkowski, Woźniakowski (2010)*, and cf. *Creutzig, Dereich, Müller-Gronbach, R* (2009).

Let $G = \mathbb{R}$ or $G = L_2([0,1]^{\mathbb{N}})$, and let $S : H \to G$ be given by

$$S(f) = \int_{[0,1]^{\mathbb{N}}} f \, d\mu \qquad \text{or} \qquad S(f) = f.$$

Let $G = \mathbb{R}$ or $G = L_2([0,1]^{\mathbb{N}})$, and let $S : H \to G$ be given by

$$S(f) = \int_{[0,1]^{\mathbb{N}}} f \, d\mu \qquad \text{or} \qquad S(f) = f.$$

We study **deterministic linear algorithms**

$$A(f) = \sum_{i=1}^{m} f(\mathbf{x}_i) \cdot g_i$$

with $g_i \in G$ and $\mathbf{x}_i \in [0,1]^{\mathbb{N}}$.

Let $G = \mathbb{R}$ or $G = L_2([0,1]^{\mathbb{N}})$, and let $S : H \to G$ be given by

$$S(f) = \int_{[0,1]^{\mathbb{N}}} f \, d\mu \qquad \text{or} \qquad S(f) = f.$$

We study **deterministic linear algorithms**

$$A(f) = \sum_{i=1}^{m} f(\mathbf{x}_i) \cdot g_i$$

with $g_i \in G$ and $\mathbf{x}_i \in [0,1]^{\mathbb{N}}$. **Error** and **cost**

$$\text{error}(A) = \sup\{\|S(f) - A(f)\|_G : f \in H,\ \|f\|_H \leq 1\},$$

$$\text{cost}(A) = \sum_{i=1}^{m} \text{active}(\mathbf{x}_i).$$

Let $G = \mathbb{R}$ or $G = L_2([0,1]^{\mathbb{N}})$, and let $S : H \to G$ be given by

$$S(f) = \int_{[0,1]^{\mathbb{N}}} f \, d\mu \qquad \text{or} \qquad S(f) = f.$$

We study **deterministic linear algorithms**

$$A(f) = \sum_{i=1}^{m} f(\mathbf{x}_i) \cdot g_i$$

with $g_i \in G$ and $\mathbf{x}_i \in [0,1]^{\mathbb{N}}$. **Error** and **cost**

$$\mathrm{error}(A) = \sup\{\|S(f) - A(f)\|_G : f \in H,\ \|f\|_H \leq 1\},$$

$$\mathrm{cost}(A) = \sum_{i=1}^{m} \mathrm{active}(\mathbf{x}_i).$$

Key quantity: ***n*-th minimal error**

$$e_n = \inf\{\mathrm{error}(A) : \mathrm{cost}(A) \leq n\}.$$

Let $G = \mathbb{R}$ or $G = L_2([0,1]^{\mathbb{N}})$, and let $S : H \to G$ be given by

$$S(f) = \int_{[0,1]^{\mathbb{N}}} f \, d\mu \qquad \text{or} \qquad S(f) = f.$$

We study **deterministic linear algorithms**

$$A(f) = \sum_{i=1}^{m} f(\mathbf{x}_i) \cdot g_i$$

with $g_i \in G$ and $\mathbf{x}_i \in [0,1]^{\mathbb{N}}$. **Error** and **cost**

$$\text{error}(A) = \sup\{\|S(f) - A(f)\|_G : f \in H, \ \|f\|_H \leq 1\},$$

$$\text{cost}(A) = \sum_{i=1}^{m} \text{active}(\mathbf{x}_i).$$

Key quantity: ***n*-th minimal error**

$$e_n = \inf\{\text{error}(A) : \text{cost}(A) \leq n\}.$$

Questions

- Order of convergence of the minimal errors e_n?
- Optimal choice of the sampling points $\mathbf{x}_1, \ldots, \mathbf{x}_n$?

III. Results and Remarks

Recall that

$$\varrho = \liminf_{j\to\infty} \frac{r_j}{\ln(j)}.$$

Put

$$s = \tfrac{1}{2} \cdot \min\bigl(r_1,\ \varrho \ln(2) - 1\bigr).$$

III. Results and Remarks

Recall that

$$\varrho = \liminf_{j\to\infty} \frac{r_j}{\ln(j)}.$$

Put

$$s = \tfrac{1}{2} \cdot \min\bigl(r_1,\ \varrho \ln(2) - 1\bigr).$$

Theorem Assume that $r_1 > 1$ and $\varrho \ln(2) > 1$. Then the minimal errors e_n for integration and L_2-approximation satisfy

$$\forall\, \varepsilon > 0\ \exists\, c_1, c_2 > 0\ \forall\, n \in \mathbb{N}:$$

$$c_1\, n^{-(s+\varepsilon)} \le e_n \le c_2\, n^{-(s-\varepsilon)}.$$

III. Results and Remarks

Recall that

$$\varrho = \liminf_{j\to\infty} \frac{r_j}{\ln(j)}.$$

Put

$$s = \tfrac{1}{2} \cdot \min\bigl(r_1,\ \varrho \ln(2) - 1\bigr).$$

Theorem Assume that $r_1 > 1$ and $\varrho \ln(2) > 1$. Then the minimal errors e_n for integration and L_2-approximation satisfy

$$\forall\, \varepsilon > 0\ \exists\, c_1, c_2 > 0\ \forall\, n \in \mathbb{N}:$$

$$c_1\, n^{-(s+\varepsilon)} \le e_n \le c_2\, n^{-(s-\varepsilon)}.$$

Remark For L_2-approximation using linear functionals at cost one

$$s = \tfrac{1}{2} \cdot \min\bigl(r_1,\ \varrho \ln(2)\bigr).$$

See *Papageorgiou, Woźniakowski (2010)*, *Siedlecki (2014)*, *Dũng, Griebel (2016)*, *Dũng, Griebel, Huy, Rieger (2018)*.

Abstract approach, based on

- a probability measure μ_0 on any set $D \neq \emptyset$,
- an orthonormal system $(e_\nu)_{\nu \in \mathbb{N}_0}$ in $L_2(D)$ with $e_0 = 1$,
- Fourier weights $\alpha_{\nu,j}$ such that, for $\nu, j \in \mathbb{N}$,

$$\alpha_{1,1} > 1 = \alpha_{0,j}$$

and

$$\alpha_{\nu,j} \geq \max(\alpha_{\nu,1}, \alpha_{1,j}).$$

Abstract approach, based on

- a probability measure μ_0 on any set $D \neq \emptyset$,
- an orthonormal system $(e_\nu)_{\nu \in \mathbb{N}_0}$ in $L_2(D)$ with $e_0 = 1$,
- Fourier weights $\alpha_{\nu,j}$ such that, for $\nu, j \in \mathbb{N}$,

$$\alpha_{1,1} > 1 = \alpha_{0,j}$$

and

$$\alpha_{\nu,j} \geq \max(\alpha_{\nu,1}, \alpha_{1,j}).$$

Example $\alpha_{\nu,j} = \exp(\nu^{b_j})$ with $0 < b_1 \leq b_2 \leq \ldots$.
Cf. *Irrgeher, Kritzer, Pillichshammer, Woźniakowski (2016)*.

Abstract approach, based on

- a probability measure μ_0 on any set $D \neq \emptyset$,
- an orthonormal system $(e_\nu)_{\nu \in \mathbb{N}_0}$ in $L_2(D)$ with $e_0 = 1$,
- Fourier weights $\alpha_{\nu,j}$ such that, for $\nu, j \in \mathbb{N}$,

$$\alpha_{1,1} > 1 = \alpha_{0,j}$$

  and

$$\alpha_{\nu,j} \geq \max(\alpha_{\nu,1}, \alpha_{1,j}).$$

Example $\alpha_{\nu,j} = \exp(\nu^{b_j})$ with $0 < b_1 \leq b_2 \leq \dots$.
Cf. *Irrgeher, Kritzer, Pillichshammer, Woźniakowski (2016)*.

Example

- Walsh, Haar, Jacobi (Legendre) basis $\rightsquigarrow$ domain $[a, b]^{\mathbb{N}}$.

Abstract approach, based on

- a probability measure μ_0 on any set $D \neq \emptyset$,
- an orthonormal system $(e_\nu)_{\nu \in \mathbb{N}_0}$ in $L_2(D)$ with $e_0 = 1$,
- Fourier weights $\alpha_{\nu,j}$ such that, for $\nu, j \in \mathbb{N}$,

$$\alpha_{1,1} > 1 = \alpha_{0,j}$$

  and

$$\alpha_{\nu,j} \geq \max(\alpha_{\nu,1}, \alpha_{1,j}).$$

Example $\alpha_{\nu,j} = \exp(\nu^{b_j})$ with $0 < b_1 \leq b_2 \leq \dots$.
Cf. *Irrgeher, Kritzer, Pillichshammer, Woźniakowski (2016)*.

Example

- Walsh, Haar, Jacobi (Legendre) basis $\rightsquigarrow$ domain $[a, b]^{\mathbb{N}}$.
- Hermite basis $\rightsquigarrow$ domain $\mathfrak{X} \subsetneq \mathbb{R}^{\mathbb{N}}$.
  Cf. Gnewuch, Mayer, R (2014).

Abstract approach, based on

- a probability measure μ_0 on any set $D \neq \emptyset$,
- an orthonormal system $(e_\nu)_{\nu \in \mathbb{N}_0}$ in $L_2(D)$ with $e_0 = 1$,
- Fourier weights $\alpha_{\nu,j}$ such that, for $\nu, j \in \mathbb{N}$,

$$\alpha_{1,1} > 1 = \alpha_{0,j}$$

  and

$$\alpha_{\nu,j} \geq \max(\alpha_{\nu,1}, \alpha_{1,j}).$$

Example $\alpha_{\nu,j} = \exp(\nu^{b_j})$ with $0 < b_1 \leq b_2 \leq \dots$.
Cf. *Irrgeher, Kritzer, Pillichshammer, Woźniakowski (2016)*.

Example

- Walsh, Haar, Jacobi (Legendre) basis $\rightsquigarrow$ domain $[a,b]^{\mathbb{N}}$.
- Hermite basis $\rightsquigarrow$ domain $\mathfrak{X} \subsetneq \mathbb{R}^{\mathbb{N}}$.
  Cf. Gnewuch, Mayer, R (2014).
- Gaussian kernels with increasing shape parameters.

IV. Embeddings and Optimal Algorithms

Notation

- $H(K)$ RKHS with reproducing kernel K,
- $\mathbf{u}$ any finite subset of $\mathbb{N}$,
- μ_0 uniform distribution on $[0, 1]$,
- μ product of μ_0 on $[0, 1]^{\mathbb{N}}$, $\mu_{\mathbf{u}}$ product of μ_0 on $[0, 1]^{\mathbf{u}}$.

IV. Embeddings and Optimal Algorithms

Notation

- $H(K)$ RKHS with reproducing kernel K,
- $\mathbf{u}$ any finite subset of $\mathbb{N}$,
- μ_0 uniform distribution on $[0,1]$,
- μ product of μ_0 on $[0,1]^{\mathbb{N}}$, $\mu_{\mathbf{u}}$ product of μ_0 on $[0,1]^{\mathbf{u}}$.

We have $H = H(K)$ with

$$K(\mathbf{x},\mathbf{y}) = \prod_{j\in\mathbb{N}} \Big(\underbrace{1 + \sum_{\nu\neq 0} \alpha_{\nu,j}^{-1} \cdot e_\nu(x_j) \cdot \overline{e_\nu(y_j)}}_{=k_j(x_j,y_j)}\Big)$$

IV. Embeddings and Optimal Algorithms

Notation

- $H(K)$ RKHS with reproducing kernel K,
- $\mathbf{u}$ any finite subset of $\mathbb{N}$,
- μ_0 uniform distribution on $[0,1]$,
- μ product of μ_0 on $[0,1]^{\mathbb{N}}$, $\mu_{\mathbf{u}}$ product of μ_0 on $[0,1]^{\mathbf{u}}$.

We have $H = H(K)$ with

$$K(\mathbf{x},\mathbf{y}) = \prod_{j\in\mathbb{N}} \Big(1 + \underbrace{\sum_{\nu\neq 0} \alpha_{\nu,j}^{-1} \cdot e_\nu(x_j) \cdot \overline{e_\nu(y_j)}}_{=k_j(x_j,y_j)}\Big) = \Big(\sum_{\mathbf{u}} \underbrace{\bigotimes_{j\in\mathbf{u}} k_j}_{=k_{\mathbf{u}}}\Big)(\mathbf{x},\mathbf{y}).$$

IV. Embeddings and Optimal Algorithms

Notation

- $H(K)$ RKHS with reproducing kernel K,
- $\mathbf{u}$ any finite subset of $\mathbb{N}$,
- μ_0 uniform distribution on $[0,1]$,
- μ product of μ_0 on $[0,1]^{\mathbb{N}}$, $\mu_{\mathbf{u}}$ product of μ_0 on $[0,1]^{\mathbf{u}}$.

We have $H = H(K)$ with

$$K(\mathbf{x},\mathbf{y}) = \prod_{j\in\mathbb{N}} \Big(1 + \underbrace{\sum_{\nu\neq 0} \alpha_{\nu,j}^{-1} \cdot e_\nu(x_j) \cdot \overline{e_\nu(y_j)}}_{=k_j(x_j,y_j)}\Big) = \Big(\sum_{\mathbf{u}} \underbrace{\bigotimes_{j\in\mathbf{u}} k_j}_{=k_{\mathbf{u}}}\Big)(\mathbf{x},\mathbf{y}).$$

Since $H(1) \perp H(k_j)$ in $H(1+k_j)$, we have the orthogonal decomposition

$$H(K) = \bigoplus_{\mathbf{u}} H(k_{\mathbf{u}}).$$

IV. Embeddings and Optimal Algorithms

Notation

- $H(K)$ RKHS with reproducing kernel K,
- $\mathbf{u}$ any finite subset of $\mathbb{N}$,
- μ_0 uniform distribution on $[0,1]$,
- μ product of μ_0 on $[0,1]^{\mathbb{N}}$, $\mu_{\mathbf{u}}$ product of μ_0 on $[0,1]^{\mathbf{u}}$.

We have $H = H(K)$ with

$$K(\mathbf{x},\mathbf{y}) = \prod_{j\in\mathbb{N}} \Big(1 + \underbrace{\sum_{\nu\neq 0} \alpha_{\nu,j}^{-1} \cdot e_\nu(x_j) \cdot \overline{e_\nu(y_j)}}_{=k_j(x_j,y_j)}\Big) = \Big(\sum_{\mathbf{u}} \underbrace{\bigotimes_{j\in\mathbf{u}} k_j}_{=k_{\mathbf{u}}}\Big)(\mathbf{x},\mathbf{y}).$$

Since $H(1) \perp H(k_j)$ in $H(1+k_j)$, we have the orthogonal decomposition

$$H(K) = \bigoplus_{\mathbf{u}} H(k_{\mathbf{u}}).$$

For the corresponding projections $f_{\mathbf{u}} \in H(k_{\mathbf{u}})$ of $f \in H(K)$

$$\int_{[0,1]^{\mathbb{N}}} f \, d\mu = \sum_{\mathbf{u}} \int_{[0,1]^{\mathbf{u}}} f_{\mathbf{u}} \, d\mu_{\mathbf{u}}.$$

Recall that

$$\int_{[0,1]^{\mathbb{N}}} f \, d\mu = \sum_{\mathbf{u}} \int_{[0,1]^{\mathbf{u}}} f_{\mathbf{u}} \, d\mu_{\mathbf{u}}.$$

Basic idea: approximate the most relevant finite-dimensional integrals based on function values of f.

Recall that

$$\int_{[0,1]^{\mathbb{N}}} f \, d\mu = \sum_{\mathbf{u}} \int_{[0,1]^{\mathbf{u}}} f_{\mathbf{u}} \, d\mu_{\mathbf{u}}.$$

Basic idea: approximate the most relevant finite-dimensional integrals based on function values of f.

More precisely, embed $H(K)$ into another RKHS on the domain $[0,1]^{\mathbb{N}}$, where this makes sense.

Proof of the Upper Bound

Step 1: Weighted kernels instead of increasing smoothness

Put

$$\gamma_j = \sup_{\nu \neq 0} \frac{\alpha_{\nu,1}}{\alpha_{\nu,j}} = 2^{r_1 - r_j}.$$

We have

- $H_1 = H(1 + \gamma_j k_1)$ with equivalent norms for every $j \in \mathbb{N}$, but
- $H_j \hookrightarrow H(1 + \gamma_j k_1)$ with norm one, and compactly if $r_j > r_1$.

Proof of the Upper Bound

Step 1: Weighted kernels instead of increasing smoothness

Put

$$\gamma_j = \sup_{\nu \neq 0} \frac{\alpha_{\nu,1}}{\alpha_{\nu,j}} = 2^{r_1 - r_j}.$$

We have

- $H_1 = H(1 + \gamma_j k_1)$ with equivalent norms for every $j \in \mathbb{N}$, but
- $H_j \hookrightarrow H(1 + \gamma_j k_1)$ with norm one, and compactly if $r_j > r_1$.

Furthermore, for $f \in H_1$ and $I(f) = \int_{[0,1]} f \, d\mu_0$,

$$\|f\|^2_{H(1+\gamma_j k_1)} = \big(I(f)\big)^2 + \frac{1}{\gamma_j} \cdot \|f - I(f)\|^2_{H(k_1)}.$$

Proof of the Upper Bound

Step 1: Weighted kernels instead of increasing smoothness

Put

$$\gamma_j = \sup_{\nu \neq 0} \frac{\alpha_{\nu,1}}{\alpha_{\nu,j}} = 2^{r_1 - r_j}.$$

We have

- $H_1 = H(1 + \gamma_j k_1)$ with equivalent norms for every $j \in \mathbb{N}$, but
- $H_j \hookrightarrow H(1 + \gamma_j k_1)$ with norm one, and compactly if $r_j > r_1$.

Furthermore, for $f \in H_1$ and $I(f) = \int_{[0,1]} f \, d\mu_0$,

$$\|f\|^2_{H(1+\gamma_j k_1)} = \big(I(f)\big)^2 + \frac{1}{\gamma_j} \cdot \|f - I(f)\|^2_{H(k_1)}.$$

Lemma For

$$L = \bigotimes_{j \in \mathbb{N}} (1 + \gamma_j k_1) = \sum_{\mathbf{u}} \underbrace{\prod_{j \in \mathbf{u}} \gamma_j}_{= \gamma_{\mathbf{u}}} \cdot \bigotimes_{j \in \mathbf{u}} k_1$$

we have $H(K) \hookrightarrow H(L)$ with norm one.

Step 2: Anchored kernels instead of ANOVA kernels

Lemma *Gnewuch, Hefter, Hinrichs, R (2017)*
For every $a \in [0, 1]$ there exists

- a reproducing kernel $m : [0, 1] \times [0, 1] \to \mathbb{C}$ and
- a constant $c > 1$

such that

$$m(a, a) = 0, \qquad H(1 + m) = H_1,$$

Step 2: Anchored kernels instead of ANOVA kernels

Lemma *Gnewuch, Hefter, Hinrichs, R (2017)*
For every $a \in [0,1]$ there exists

- a reproducing kernel $m : [0,1] \times [0,1] \to \mathbb{C}$ and
- a constant $c > 1$

such that

$$m(a,a) = 0, \qquad H(1+m) = H_1,$$

and for

$$M = \bigotimes_{j \in \mathbb{N}} (1 + c\gamma_j m) = \sum_{\mathbf{u}} c^{|\mathbf{u}|} \gamma_{\mathbf{u}} \underbrace{\bigotimes_{j \in \mathbf{u}} m}_{= m_{\mathbf{u}}}$$

we have

$$H(L) \hookrightarrow H(M) \quad \text{continuously.}$$

Step 2: Anchored kernels instead of ANOVA kernels

Lemma *Gnewuch, Hefter, Hinrichs, R (2017)*
For every $a \in [0,1]$ there exists

- a reproducing kernel $m : [0,1] \times [0,1] \to \mathbb{C}$ and
- a constant $c > 1$

such that

$$m(a,a) = 0, \qquad H(1+m) = H_1,$$

and for

$$M = \bigotimes_{j \in \mathbb{N}} (1 + c\gamma_j m) = \sum_{\mathbf{u}} c^{|\mathbf{u}|} \gamma_{\mathbf{u}} \underbrace{\bigotimes_{j \in \mathbf{u}} m}_{= m_{\mathbf{u}}}$$

we have

$$H(L) \hookrightarrow H(M) \quad \text{continuously.}$$

Key property: M is a superposition of weighted tensor products of an anchored kernel.

Step 3: The multivariate decomposition method

We have the orthogonal decomposition

$$H(M) = \bigoplus_{\mathbf{u}} H(\gamma_{\mathbf{u}} m_{\mathbf{u}}).$$

Step 3: The multivariate decomposition method

We have the orthogonal decomposition

$$H(M) = \bigoplus_{\mathbf{u}} H(\gamma_{\mathbf{u}} m_{\mathbf{u}}).$$

The corresponding projections $f_{\mathbf{u}} \in H(\gamma_{\mathbf{u}} k_{\mathbf{u}})$ of $f \in H(M)$ are given by

$$f_\emptyset = f(a, a, \dots) \quad \text{and} \quad f_{\mathbf{u}}((x_j)_{j\in\mathbf{u}}) = \Big(f - \sum_{\mathbf{v}\subsetneq\mathbf{u}} f_{\mathbf{v}}\Big)(\mathbf{y})$$

for $\mathbf{u} \neq \emptyset$, where

$$y_j = \begin{cases} x_j, & \text{if } j \in \mathbf{u}, \\ a, & \text{otherwise.} \end{cases}$$

Step 3: The multivariate decomposition method

We have the orthogonal decomposition

$$H(M) = \bigoplus_{\mathbf{u}} H(\gamma_{\mathbf{u}} m_{\mathbf{u}}).$$

The corresponding projections $f_{\mathbf{u}} \in H(\gamma_{\mathbf{u}} k_{\mathbf{u}})$ of $f \in H(M)$ are given by

$$f_{\emptyset} = f(a, a, \dots) \quad \text{and} \quad f_{\mathbf{u}}((x_j)_{j\in\mathbf{u}}) = \Big(f - \sum_{\mathbf{v}\subsetneq\mathbf{u}} f_{\mathbf{v}}\Big)(\mathbf{y})$$

for $\mathbf{u} \neq \emptyset$, where

$$y_j = \begin{cases} x_j, & \text{if } j \in \mathbf{u}, \\ a, & \text{otherwise.} \end{cases}$$

Lemma *Kuo, Sloan, Wasilkowski, Woźniakowski* (2010)
Values of $f_{\mathbf{u}}$ can be efficiently computed from a 'small' number of values of f using only a 'few' active variables.

Step 3: The multivariate decomposition method

We have the orthogonal decomposition

$$H(M) = \bigoplus_{\mathbf{u}} H(\gamma_{\mathbf{u}} m_{\mathbf{u}}).$$

The corresponding projections $f_{\mathbf{u}} \in H(\gamma_{\mathbf{u}} k_{\mathbf{u}})$ of $f \in H(M)$ are given by

$$f_\emptyset = f(a, a, \dots) \quad \text{and} \quad f_{\mathbf{u}}((x_j)_{j\in\mathbf{u}}) = \Big(f - \sum_{\mathbf{v}\subsetneq\mathbf{u}} f_{\mathbf{v}}\Big)(\mathbf{y})$$

for $\mathbf{u} \neq \emptyset$, where

$$y_j = \begin{cases} x_j, & \text{if } j \in \mathbf{u}, \\ a, & \text{otherwise.} \end{cases}$$

Lemma *Kuo, Sloan, Wasilkowski, Woźniakowski* (2010)
Values of $f_{\mathbf{u}}$ can be efficiently computed from a 'small' number of values of f using only a 'few' active variables.

Theorem *Plaskota, Wasilkowski (2011), Wasilkowski (2012)*
The minimal errors e_n for int/app on $H(M)$ satisfy

$$\forall\, \varepsilon > 0\ \exists\, c > 0\ \forall\, n \in \mathbb{N}: \qquad e_n \le c\, n^{-(s-\varepsilon)}.$$

Step 3: The multivariate decomposition method

We have the orthogonal decomposition

$$H(M) = \bigoplus_{\mathbf{u}} H(\gamma_{\mathbf{u}} m_{\mathbf{u}}).$$

The corresponding projections $f_{\mathbf{u}} \in H(\gamma_{\mathbf{u}} k_{\mathbf{u}})$ of $f \in H(M)$ are given by

$$f_{\emptyset} = f(a, a, \dots) \quad \text{and} \quad f_{\mathbf{u}}((x_j)_{j \in \mathbf{u}}) = \Big(f - \sum_{\mathbf{v} \subsetneq \mathbf{u}} f_{\mathbf{v}}\Big)(\mathbf{y})$$

for $\mathbf{u} \neq \emptyset$, where

$$y_j = \begin{cases} x_j, & \text{if } j \in \mathbf{u}, \\ a, & \text{otherwise.} \end{cases}$$

Lemma *Kuo, Sloan, Wasilkowski, Woźniakowski* (2010)
Values of $f_{\mathbf{u}}$ can be efficiently computed from a 'small' number of values of f using only a 'few' active variables.

See *Gilbert, Kuo, Nuyens, Wasilkowski (2017) for the implementation.*

Proof of the Lower Bound for Integration

Since $H_1 \hookrightarrow H(K)$ with norm one, we get

$$\exists\, c > 0\ \forall\, n \in \mathbb{N}: \qquad e_n \geq c\, n^{-r_1/2}.$$

Proof of the Lower Bound for Integration

Since $H_1 \hookrightarrow H(K)$ with norm one, we get

$$\exists\, c > 0\ \forall\, n \in \mathbb{N}: \qquad e_n \geq c\, n^{-r_1/2}.$$

To verify

$$\forall\, \varepsilon > 0\ \exists\, c > 0\ \forall\, n \in \mathbb{N}: \qquad e_n \geq c\, n^{-(\varrho \ln(2) - 1 + \varepsilon)/2}$$

Proof of the Lower Bound for Integration

Since $H_1 \hookrightarrow H(K)$ with norm one, we get

$$\exists\, c > 0\ \forall\, n \in \mathbb{N}: \qquad e_n \geq c\, n^{-r_1/2}.$$

To verify

$$\forall\, \varepsilon > 0\ \exists\, c > 0\ \forall\, n \in \mathbb{N}: \qquad e_n \geq c\, n^{-(\varrho \ln(2) - 1 + \varepsilon)/2}$$

we

- start with $\operatorname{span}\{e_0, e_1\} \subset H(1 + k_j)$,
- apply the two embedding steps reversely,
- employ the lower bound from *Plaskota, Wasilkowski (2011)* for superpositions of weighted tensor products of anchored kernels.

Does randomization help?

Put

$$s^{\mathrm{ran}} = \frac{1}{2} \cdot \min\bigl(r_1 + 1,\ \varrho \ln(2) - 1\bigr).$$

Does randomization help?

Put

$$s^{\mathrm{ran}} = \tfrac{1}{2} \cdot \min\bigl(r_1 + 1,\ \varrho \ln(2) - 1\bigr).$$

Theorem Assume that $r_1 > 1$ and $\varrho \ln(2) > 1$. Then the minimal errors e_n^{ran} of randomized algorithms for integration satisfy

$$\forall\, \varepsilon > 0\ \exists\, c_1, c_2 > 0\ \forall\, n \in \mathbb{N}:$$

$$c_1\, n^{-(s^{\mathrm{ran}}+\varepsilon)} \le e_n^{\mathrm{ran}} \le c_2\, n^{-(s^{\mathrm{ran}}-\varepsilon)}.$$

Does randomization help?

Put

$$s^{\mathrm{ran}} = \frac{1}{2} \cdot \min\big(r_1 + 1,\ \varrho \ln(2) - 1\big).$$

Theorem Assume that $r_1 > 1$ and $\varrho \ln(2) > 1$. Then the minimal errors e_n^{ran} of randomized algorithms for integration satisfy

$$\forall \varepsilon > 0 \ \exists c_1, c_2 > 0 \ \forall n \in \mathbb{N}:$$

$$c_1 \, n^{-(s^{\mathrm{ran}}+\varepsilon)} \leq e_n^{\mathrm{ran}} \leq c_2 \, n^{-(s^{\mathrm{ran}}-\varepsilon)}.$$

Proof: Use embeddings and *Dick, Gnewuch (2014)*.

Summary

- Function spaces: countable tensor products of (Korobov) spaces of increasing smoothness.
- n-th minimal error for integration and L_2-approximation of order

$$s = \tfrac{1 \pm \varepsilon}{2} \cdot \min\big(r_1,\ \varrho \ln(2) - 1\big).$$

Summary

- Function spaces: countable tensor products of (Korobov) spaces of increasing smoothness.
- n-th minimal error for integration and L_2-approximation of order

$$s = \frac{1 \pm \varepsilon}{2} \cdot \min\big(r_1,\ \varrho \ln(2) - 1\big).$$

- Main tool: embeddings, see *Hefter, R (2015)*, **. . .** .
- Here: weights and anchored kernels instead of increasing smoothness.

Summary

- Function spaces: countable tensor products of (Korobov) spaces of increasing smoothness.
- n-th minimal error for integration and L_2-approximation of order

$$s = \tfrac{1 \pm \varepsilon}{2} \cdot \min\big(r_1,\ \varrho \ln(2) - 1\big).$$

- Main tool: embeddings, see *Hefter, R (2015)*, **. . . .**
- Here: weights and anchored kernels instead of increasing smoothness.
- Form a complexity point of view: excessive amount of smoothness in the tensor products of (Korobov) spaces of increasing smoothness.