# Uncertainty Quantification and Quasi-Monte Carlo

## Wintersemester 2022/23

Vesa Kaarnioja
vesa.kaarnioja@fu-berlin.de

FU Berlin, FB Mathematik und Informatik

Seventh lecture, November 28, 2022

Today's lecture follows the survey article

J. Dick, F. Y. Kuo, and I. H. Sloan. High-dimensional integration: The quasi-Monte Carlo way. *Acta Numer.* **22**:133–288, 2013.
https://doi.org/10.1017/S0962492913000044

# Notations

- $\{1:s\} := \{1,2,\ldots,s\}$ for $s \in \mathbb{N}$. We use fraktur letters to denote subsets $\mathfrak{u} \subseteq \{1:s\}$. We use $|\mathfrak{u}|$ to denote the cardinality of set $\mathfrak{u}$.
- For $x \geq 0$, we define the fractional part $\{x\} := x - \lfloor x \rfloor = \mathrm{mod}(x,1)$. For $x < 0$, $\{x\} := x + \lfloor |x| \rfloor$. For $\boldsymbol{x} \in \mathbb{R}^s$, we define

$$\{\boldsymbol{x}\} := (\{x_1\},\{x_2\},\ldots,\{x_s\}).$$

  For example, $\{(1.2, 0.5, 2.77)\} = (0.2, 0.5, 0.77)$.
- For $\mathfrak{u} \subseteq \{1:s\}$, we define $\boldsymbol{x}_{\mathfrak{u}} = (x_j)_{j\in\mathfrak{u}}$ and

$$\frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{x}_{\mathfrak{u}}} f(\boldsymbol{x}) := \prod_{j\in\mathfrak{u}} \frac{\partial}{\partial x_j} f(\boldsymbol{x}).$$

  For example, with $\mathfrak{u} = \{1,2,4\}$, we have $|\mathfrak{u}| = 3$, $\boldsymbol{x}_{\mathfrak{u}} = (x_1, x_2, x_4)$, and

$$\frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{x}_{\mathfrak{u}}} f(\boldsymbol{x}) = \frac{\partial^3}{\partial x_1 \partial x_2 \partial x_4} f(\boldsymbol{x}).$$

# Quasi-Monte Carlo methods

Let $f \in C([0,1]^s)$. We consider the problem of approximating the high-dimensional integral

$$I_s f = \int_{[0,1]^s} f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

*Quasi-Monte Carlo (QMC) methods* are a class of *equal weight* cubature rules

$$Q_{n,s} f = \frac{1}{n} \sum_{i=0}^{n-1} f(\boldsymbol{t}_i),$$

where $(\boldsymbol{t}_i)_{i=0}^{n-1}$ is an ensemble of *deterministic* nodes in $[0,1]^s$ (**not** a random sample of $\mathcal{U}([0,1]^s)$).

QMC methods exploit the smoothness and anisotropy of an integrand in order to achieve better-than-Monte Carlo cubature convergence rates.

# Rank-1 lattice rules

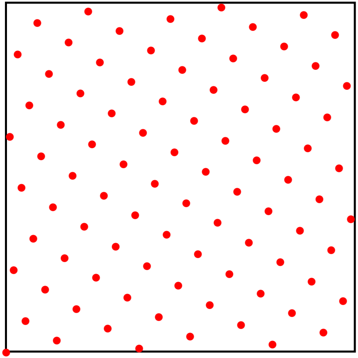
*Rank-1 lattice rules*

$$Q_{n,s}f = \frac{1}{n}\sum_{i=0}^{n-1} f(\boldsymbol{t}_i)$$

have the points

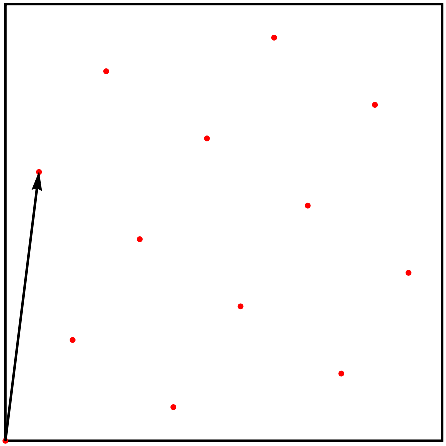
$$\boldsymbol{t}_i = \mathrm{mod}\left(\frac{i\boldsymbol{z}}{n}, 1\right), \quad i \in \{0, \ldots, n-1\},$$

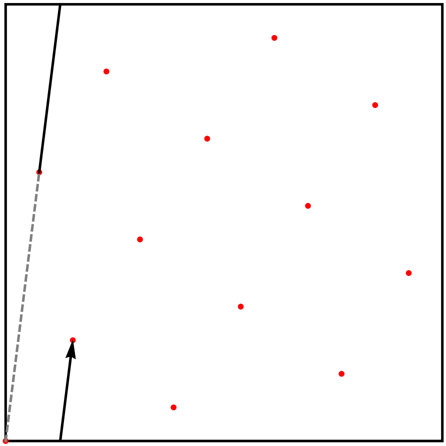
where the entire point set is determined by the *generating vector* $\boldsymbol{z} \in \mathbb{N}^s$, with all components *coprime* to $n$.

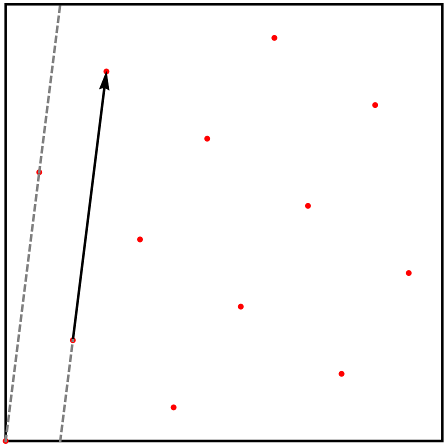


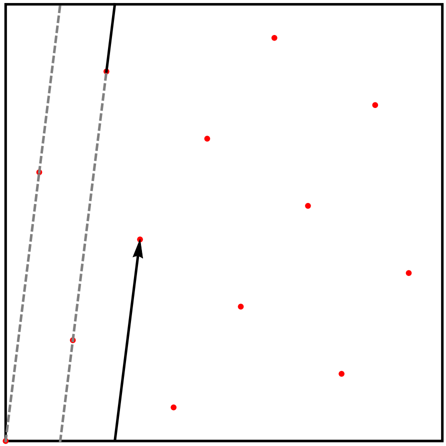
Lattice rule with $\boldsymbol{z} = (1, 55)$ and $n = 89$ nodes in $[0, 1]^2$

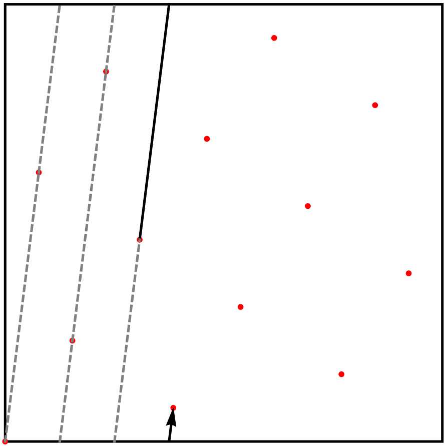
*"Lattice rules and periodic functions are a match made in heaven!" –VK*

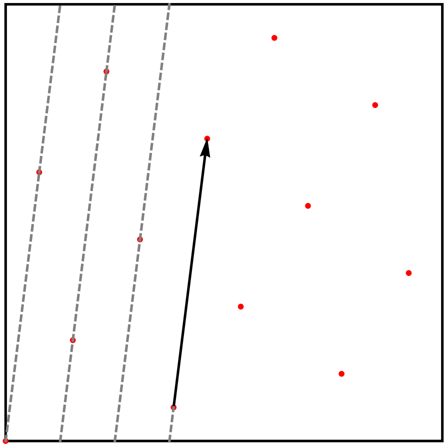


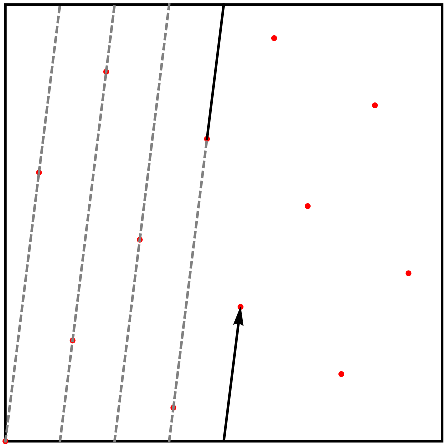


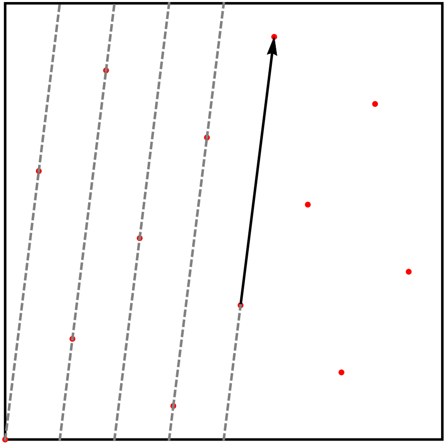


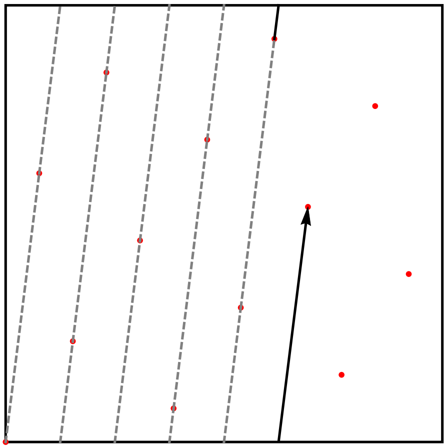


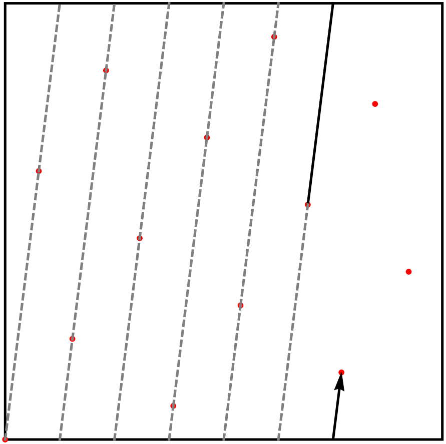


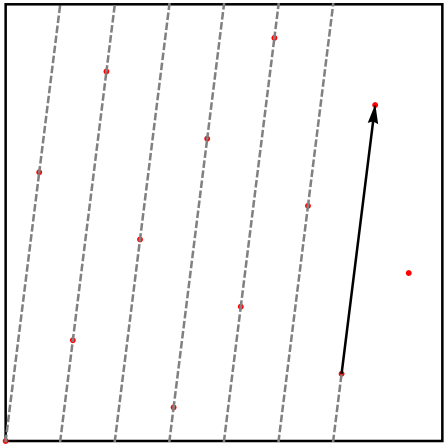


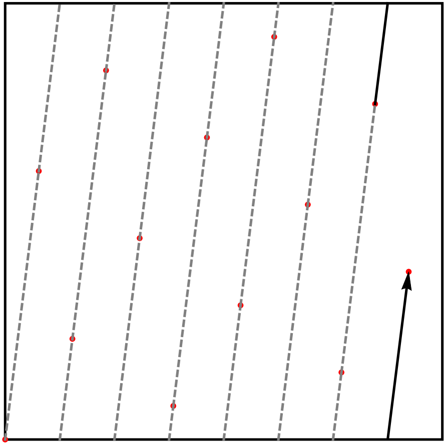












# Historical remarks on the development of lattice rules

- Number theorists (Korobov, Zaremba, Hua) in the 1950s and 60s.
- Lattice rules for multiple integration (Sloan and Kachoyan 1987; Sloan and Joe 1994).
- Weighted spaces (Sloan and Woźniakowski 1998; Hickernell 1996).
- Component-by-component (CBC) construction of lattice rules (Kuo, Joe, Sloan 2002).
- Fast CBC algorithm (Cools and Nuyens 2006; Kuo, Cools, and Nuyens 2006).
- Uncertainty quantification of PDEs using QMC methods (Kuo, Schwab, Sloan 2012).

and of course many, many others! (Dick, Giles, Goda, Graham, Kritzer, Niederreiter, Pillichshammer, Wasilkowski, . . .)

# Brief introduction to the classical theory of lattice rules

Let $f : [0,1]^s \to \mathbb{R}$ be an absolutely continuous and 1-periodic function, i.e.,

$$f(y_1, y_2, \ldots, y_s) = f(y_1 + 1, y_2, \ldots, y_s) = f(y_1, y_2 + 1, \ldots, y_s) = \cdots,$$

with an *absolutely convergent Fourier series*

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{h} \in \mathbb{Z}^s} \widehat{f}(\boldsymbol{h}) \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{x}}, \quad \widehat{f}(\boldsymbol{h}) := \int_{[0,1]^s} f(\boldsymbol{x}) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x}.$$

Then the lattice rule error is precisely the sum of the integrand's Fourier coefficients over the so-called *dual lattice*.

**Theorem (Rank-1 lattice rule error)**

*Under the aforementioned conditions on $f : [0,1]^s \to \mathbb{R}$, there holds*

$$Q_{n,s}(f) - I_s(f) = \sum_{\boldsymbol{h} \in \Lambda^\perp \setminus \{\boldsymbol{0}\}} \widehat{f}(\boldsymbol{h}),$$

*where the dual lattice*

$$\Lambda^\perp := \{\boldsymbol{h} \in \mathbb{Z}^s \mid \boldsymbol{h} \cdot \boldsymbol{z} \equiv 0 \pmod{n}\}$$

*is determined entirely by the generating vector $\boldsymbol{z} \in \mathbb{N}^s$ and $n \in \mathbb{N}$.*

For future convenience, let us prove a couple of helpful auxiliary identities.

**Lemma**

*Let* $\boldsymbol{h} = (h_1, \ldots, h_s) \in \mathbb{Z}^s$ *and* $n \in \mathbb{N}$. *Then*

$$\int_{[0,1]^s} \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x} = \begin{cases} 1 & \text{if } \boldsymbol{h} = \boldsymbol{0} \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathrm{e}^{2\pi\mathrm{i}k\boldsymbol{h}\cdot\boldsymbol{z}/n} = \begin{cases} 1 & \text{if } \boldsymbol{h}\cdot\boldsymbol{z} \equiv 0 \pmod{n} \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* By Fubini's theorem

$$\int_{[0,1]^s} \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x} = \prod_{j=1}^{s}\int_0^1 \mathrm{e}^{2\pi\mathrm{i}h_j x_j}\,\mathrm{d}x_j, \tag{1}$$

where

$$\int_0^1 \mathrm{e}^{2\pi\mathrm{i}h_j x_j}\,\mathrm{d}x_j = \begin{cases} \int_0^1 \mathrm{d}x_j & \text{if } h_j = 0 \\ \left[\frac{\mathrm{e}^{2\pi\mathrm{i}h_j x_j}}{2\pi\mathrm{i}h_j}\right]_{x_j=0}^{x_j=1} & \text{if } h_j \neq 0 \end{cases} = \begin{cases} 1 & \text{if } h_j = 0 \\ 0 & \text{if } h_j \neq 0. \end{cases}$$

Thus the expression (1) is zero unless $h_1 = h_2 = \cdots = h_s = 0$.

To prove the second claim

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathrm{e}^{2\pi \mathrm{i} k \boldsymbol{h}\cdot\boldsymbol{z}/n} = \begin{cases} 1 & \text{if } \boldsymbol{h}\cdot\boldsymbol{z} \equiv 0 \pmod{n} \\ 0 & \text{otherwise} \end{cases}$$

consider two cases:

- If $\boldsymbol{h}\cdot\boldsymbol{z}$ is a multiple of $n$, i.e., $\boldsymbol{h}\cdot\boldsymbol{z} \equiv 0 \pmod{n}$, then clearly

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathrm{e}^{2\pi \mathrm{i} k \boldsymbol{h}\cdot\boldsymbol{z}/n} = \frac{1}{n}\sum_{k=0}^{n-1} \mathrm{e}^{0} = 1.$$

- If $\boldsymbol{h}\cdot\boldsymbol{z}$ is not a multiple of $n$, then by the geometric sum formula

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathrm{e}^{2\pi \mathrm{i} k \boldsymbol{h}\cdot\boldsymbol{z}/n} = \frac{1}{n}\sum_{k=0}^{n-1} \left(\mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{z}/n}\right)^{k} = \frac{1}{n}\frac{1-(\mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{z}/n})^{n}}{1-\mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{z}/n}} = 0.$$

This yields the assertion. □

*Proof (Rank-1 lattice rule error).* Using the Fourier series representation

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{h}\in\mathbb{Z}^s} \widehat{f}(\boldsymbol{h})\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}}, \quad \widehat{f}(\boldsymbol{h}) := \int_{[0,1]^s} f(\boldsymbol{x})\mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x},$$

and noting that $\mathrm{e}^{2\pi\mathrm{i}\left\{\frac{k\boldsymbol{z}}{n}\right\}\cdot\boldsymbol{h}} = \mathrm{e}^{2\pi\mathrm{i}k\boldsymbol{z}\cdot\boldsymbol{h}/n}$, we can change the order of the series (note that the Fourier series is absolutely convergent!) to obtain

$$\begin{aligned}
Q_{n,s}(f) - I_s(f) &= \frac{1}{n}\sum_{k=0}^{n-1} f\left(\left\{\frac{k\boldsymbol{z}}{n}\right\}\right) - \int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \frac{1}{n}\sum_{k=0}^{n-1}\sum_{\boldsymbol{h}\in\mathbb{Z}^s} \widehat{f}(\boldsymbol{h})\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}/n} - \widehat{f}(\boldsymbol{0}) \\
&= \sum_{\boldsymbol{h}\in\mathbb{Z}^s} \widehat{f}(\boldsymbol{h}) \underbrace{\frac{1}{n}\sum_{k=0}^{n-1} \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}/n}}_{\substack{=1 \text{ if } \boldsymbol{h}\cdot\boldsymbol{z}\equiv 0 \pmod{n} \\ =0 \text{ otherwise}}} - \widehat{f}(\boldsymbol{0}) \\
&= \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^s \\ \boldsymbol{h}\cdot\boldsymbol{z}\equiv 0 \pmod{n}}} \widehat{f}(\boldsymbol{h}) - \widehat{f}(\boldsymbol{0}) = \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^s\setminus\{\boldsymbol{0}\} \\ \boldsymbol{h}\cdot\boldsymbol{z}\equiv 0 \pmod{n}}} \widehat{f}(\boldsymbol{h}). \quad \square
\end{aligned}$$

Ultimately, we are interested in applying lattice rules for *non-periodic*, smooth functions. We will need to put in a bit more effort to make this method work in the non-periodic setting...

# Worst-case error and reproducing kernel Hilbert space (RKHS)

# Worst-case error

In the classical study of quadrature and cubature rules, we usually consider the so-called *worst-case error*. Suppose that $f \in H$, where $H$ is a Hilbert space continuously embedded in $C([0,1]^s)$. Let $I_s\colon H \to \mathbb{R}$ be an integral operator

$$I_s f := \int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

and let $Q_{n,s}\colon H \to \mathbb{R}$ be a QMC rule

$$Q_{n,s} f := \frac{1}{n}\sum_{i=0}^{n-1} f(\boldsymbol{t}_i),$$

where $P := \{\boldsymbol{t}_i \in [0,1]^s \mid 0 \leq i \leq n-1\}$ is a collection of cubature nodes. The worst-case error of cubature rule $Q_{n,s}$ in $H$ is defined by

$$e_{n,s}(P;H) := \sup_{\substack{f\in H\\ \|f\|_H\leq 1}} |I_s f - Q_{n,s} f|.$$

*Note that this is precisely the operator norm of* $\|I_s - Q_{n,s}\|_{H\to\mathbb{R}}$.

Since the worst-case error is just the operator norm of $I_s - Q_{n,s}$, we can express the cubature error as

$$|I_s f - Q_{n,s} f| \leq e_{n,s}(P; H)\|f\|_H.$$

Worst-case errors are in general hard to compute – except for the special case, when $H$ is a *reproducing kernel Hilbert space* (RKHS).

Our strategy will be to *choose* the Hilbert space $H$ (where our integrand $f$ lives) to be such that it is possible to write down the expression for $e_{n,s}(P; H)$ *explicitly* given a family of QMC rules. This allows us to analyze the dependence of the cubature error w.r.t. $n$ and $s$.

We will end up taking $H$ as an *unanchored, weighted Sobolev space* since this choice turns out to be "compatible" with the family of (randomly shifted) lattice rules!

# Reproducing kernel Hilbert space (RKHS)

Let $H$ be a Hilbert space of functions on $D \subseteq \mathbb{R}^s$, with the property that *every point evaluation is a bounded linear functional*. That is, for any $\boldsymbol{y} \in D$, let

$$T_{\boldsymbol{y}}(f) := f(\boldsymbol{y}) \quad \text{for all } f \in H.$$

Then, since $T_{\boldsymbol{y}}$ is a bounded linear functional, by Riesz representation theorem there exists a unique representer $a_{\boldsymbol{y}} := K(\cdot, \boldsymbol{y}) \in H$ such that

$$T_{\boldsymbol{y}}(f) = \langle f, a_{\boldsymbol{y}} \rangle = \langle f, K(\cdot, \boldsymbol{y}) \rangle \quad \text{for all } f \in H.$$

The function $K(\boldsymbol{x}, \boldsymbol{y})$ is known as the *reproducing kernel* of $H$.

## Definition (Reproducing kernel)

A *reproducing kernel* of a Hilbert space $H$ of functions on $D \subseteq \mathbb{R}^s$ is a function $K \colon D \times D \to \mathbb{R}$ which satisfies

$$K(\cdot, \boldsymbol{y}) \in H \quad \text{for all } \boldsymbol{y} \in D$$
$$\text{and } f(\boldsymbol{y}) = \langle f, K(\cdot, \boldsymbol{y}) \rangle \quad \text{for all } f \in H \text{ and } \boldsymbol{y} \in D.$$

The latter property is known as the *reproducing property*.

# Remarks

- A *reproducing kernel Hilbert space* (RKHS) is a Hilbert space equipped with a reproducing kernel, or equivalently, it is a Hilbert space in which *every point evaluation is a bounded linear functional*.
- For any other bounded linear functional $A\colon H \to \mathbb{R}$, its representer $a \in H$ satisfying $A(f) = \langle f, a \rangle$ for all $f \in H$ is given by

$$a(\boldsymbol{y}) = \langle a, K(\cdot, \boldsymbol{y}) \rangle = \langle K(\cdot, \boldsymbol{y}), a \rangle = A(K(\cdot, \boldsymbol{y})) \quad \text{for all } \boldsymbol{y} \in D.$$

- Any reproducing kernel $K(\boldsymbol{x}, \boldsymbol{y})$ is symmetric in its arguments:

$$K(\boldsymbol{x}, \boldsymbol{y}) = K(\boldsymbol{y}, \boldsymbol{x}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in D.$$

*Proof.* For fixed $\boldsymbol{y} \in D$, apply the reproducing property to the function $f = K(\cdot, \boldsymbol{y})$ to get

$$\begin{aligned} K(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{x}) = \langle f, K(\cdot, \boldsymbol{x}) \rangle &= \langle K(\cdot, \boldsymbol{y}), \langle K(\cdot, \boldsymbol{x})\rangle \\ &= \langle K(\cdot, \boldsymbol{x}), K(\cdot, \boldsymbol{y}) \rangle = K(\boldsymbol{y}, \boldsymbol{x}). \quad \square \end{aligned}$$

## Example

Suppose that we have a Hilbert space containing continuous functions on $[0, 1]$ with square-integrable first order derivatives, equipped with the inner product

$$\langle f, g \rangle = \left( \int_0^1 f(x)\,\mathrm{d}x \right)\left( \int_0^1 g(x)\,\mathrm{d}x \right) + \int_0^1 f'(x)g'(x)\,\mathrm{d}x.$$

Then this space has the *reproducing kernel*

$$K(x, y) = 1 + \eta(x, y),\ \eta(x, y) = \tfrac{1}{2}B_2(|x - y|) + (x - \tfrac{1}{2})(y - \tfrac{1}{2}),$$

where $B_2(x) := x^2 - x + \frac{1}{6}$ denotes the *Bernoulli polynomial of degree 2*.

That is, we claim that

$$\langle f, K(\cdot, y) \rangle = f(y) \quad \text{for all } y \in [0, 1].$$

Example (continued)

By observing that

$$\int_0^1 K(x,y)\,\mathrm{d}x = 1 \quad \text{and} \quad \frac{\partial}{\partial x}K(x,y) = x - \frac{1}{2} - \frac{1}{2}\operatorname{sign}(x-y),$$

there holds

$$\langle f, K(\cdot,y)\rangle = \left(\int_0^1 f(x)\,\mathrm{d}x\right)\underbrace{\left(\int_0^1 K(x,y)\,\mathrm{d}x\right)}_{=1} + \int_0^1 f'(x)\left(x - \frac{1}{2} - \frac{1}{2}\operatorname{sign}(x-y)\right)\mathrm{d}x$$

$$= \int_0^1 f(x)\,\mathrm{d}x + \int_0^1 f'(x)x\,\mathrm{d}x - \frac{1}{2}\int_0^1 f'(x)\,\mathrm{d}x + \frac{1}{2}\int_0^y f'(x)\,\mathrm{d}x - \frac{1}{2}\int_y^1 f'(x)\,\mathrm{d}x$$

$$= \cancel{\int_0^1 f(x)\,\mathrm{d}x} + \cancel{f(1)} - \cancel{\int_0^1 f(x)\,\mathrm{d}x} - \cancel{\frac{1}{2}f(1)} + \cancel{\frac{1}{2}f(0)} + \frac{1}{2}f(y) - \cancel{\frac{1}{2}f(0)} - \cancel{\frac{1}{2}f(1)} + \frac{1}{2}f(y)$$

$$= f(y)$$

for all $y \in [0,1]$, as desired.

## Theorem

*Let $H := H_s(K)$ be an RKHS and let $K\colon [0,1]^s \times [0,1]^s \to \mathbb{R}$ be a reproducing kernel that satisfies*

$$\int_{[0,1]^s}\int_{[0,1]^s} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} < \infty.$$

*Then*

$$e_{n,s}^2(P; H_s(K)) = \int_{[0,1]^s}\int_{[0,1]^s} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} - \frac{2}{n}\sum_{i=0}^{n-1}\int_{[0,1]^s} K(\boldsymbol{t}_i,\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} + \frac{1}{n^2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} K(\boldsymbol{t}_i,\boldsymbol{t}_j).$$

*Proof.* For $f \in H$, we apply the reproducing property $f(\boldsymbol{t}_k) = \langle f, K(\cdot, \boldsymbol{t}_k)\rangle_H$ and average the results to obtain

$$Q_{n,s} f = \frac{1}{n}\sum_{k=0}^{n-1} f(\boldsymbol{t}_k) = \frac{1}{n}\sum_{k=0}^{n-1}\langle f, K(\cdot, \boldsymbol{t}_k)\rangle_H = \left\langle f, \frac{1}{n}\sum_{k=0}^{n-1} K(\cdot, \boldsymbol{t}_k)\right\rangle_H. \quad (2)$$

Similarly, we find that

$$I_s f = \int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{[0,1]^s}\langle f, K(\cdot, \boldsymbol{x})\rangle_H\,\mathrm{d}\boldsymbol{x} = \left\langle f, \int_{[0,1]^s} K(\cdot, \boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right\rangle_H, \quad (3)$$

which holds provided that $\int_{[0,1]^s} K(\cdot, \boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \in H$. However, this is guaranteed by our assumption since

$$\begin{aligned}\left\| \int_{[0,1]^s} K(\cdot, \boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right\|_H^2 &= \int_{[0,1]^s}\int_{[0,1]^s}\langle K(\cdot, \boldsymbol{x}), K(\cdot, \boldsymbol{y})\rangle_H\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} \\ &= \int_{[0,1]^s}\int_{[0,1]^s} K(\boldsymbol{x}, \boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} < \infty,\end{aligned}$$

which will hold for all the kernels we shall consider.

Taking the difference of (2) and (3) yields

$$I_s f - Q_{n,s} f = \left\langle f, \int_{[0,1]^s} K(\cdot, \boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=0}^{n-1} K(\cdot, \boldsymbol{t}_i) \right\rangle_H = \langle f, \xi \rangle_H,$$

where

$$\xi(\boldsymbol{y}) := \int_{[0,1]^s} K(\boldsymbol{x}, \boldsymbol{y}) \,\mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=0}^{n-1} K(\boldsymbol{y}, \boldsymbol{t}_i), \quad \boldsymbol{y} \in [0,1]^s$$

is called the *representer* of the integration error since

$$e_{n,s}(P; H) = \sup_{\|f\| \le 1} |\langle f, \xi \rangle_H| = \|\xi\|_H.$$

Especially, the supremum is attained by $f = \xi / \|\xi\| \in H$ and we obtain

$$e_{n,s}^2(P; H) = \left\| \int_{[0,1]^s} K(\cdot, \boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=0}^{n-1} K(\boldsymbol{x}, \boldsymbol{t}_i) \right\|^2$$

$$= \int_{[0,1]^s} \int_{[0,1]^s} K(\boldsymbol{x}, \boldsymbol{y}) \,\mathrm{d}\boldsymbol{x} \,\mathrm{d}\boldsymbol{y} - \frac{2}{n} \sum_{i=0}^{n-1} \int_{[0,1]^s} K(\boldsymbol{x}, \boldsymbol{t}_i) \,\mathrm{d}\boldsymbol{x} + \frac{1}{n^2} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} K(\boldsymbol{t}_i, \boldsymbol{t}_j),$$

as desired. □