Optimal Importance Sampling for the Approximation of Integrals

Special Session on Monte Carlo Methods and Functional Analysis in Honor of Stefan Heinrich's 60th Birthday

Aicke Hinrichs

Friedrich-Schiller University Jena

MCQMC
Warsaw 2010

The Problem

- $D \subseteq \mathbb{R}^d$ Borel measurable
- ϱ probability density on D
- H Hilbert space of functions $f : D \to \mathbb{R}$
- **Integration Problem:**

$$I(f) = \int_D f(x)\varrho(x)\,\mathrm{d}x$$

- Problem is well defined iff $H \subset L_1(\varrho)$ iff

$$C^{\mathrm{init}} = \left(\int_D \int_D K(x,y)\varrho(x)\varrho(y)\,\mathrm{d}x\mathrm{d}y\right)^{1/2} < \infty$$

- **Algorithms:** Randomized Algorithm using n function values, in particular importance sampling
- to have function values well defined we assume that H is a **reproducing kernel Hilbert space** with kernel $K : D \times D \to \mathbb{R}$.

The Problem

- $D \subseteq \mathbb{R}^d$ Borel measurable
- ϱ probability density on D
- H Hilbert space of functions $f : D \to \mathbb{R}$
- **Integration Problem:**

$$I(f) = \int_D f(x)\varrho(x)\,\mathrm{d}x$$

- Problem is well defined iff $H \subset L_1(\varrho)$ iff

$$C^{\mathrm{init}} = \left(\int_D \int_D K(x,y)\varrho(x)\varrho(y)\,\mathrm{d}x\mathrm{d}y\right)^{1/2} < \infty$$

- **Algorithms:** Randomized Algorithm using n function values, in particular importance sampling
- to have function values well defined we assume that H is a **reproducing kernel Hilbert space** with kernel $K : D \times D \to \mathbb{R}$.

The Problem

- $D \subseteq \mathbb{R}^d$ Borel measurable
- ϱ probability density on D
- H Hilbert space of functions $f : D \to \mathbb{R}$
- **Integration Problem:**

$$I(f) = \int_D f(x)\varrho(x)\,\mathrm{d}x$$

- Problem is well defined iff $H \subset L_1(\varrho)$ iff

$$C^{\mathrm{init}} = \left(\int_D \int_D K(x,y)\varrho(x)\varrho(y)\,\mathrm{d}x\mathrm{d}y\right)^{1/2} < \infty$$

- **Algorithms:** Randomized Algorithm using n function values, in particular importance sampling
- to have function values well defined we assume that H is a **reproducing kernel Hilbert space** with kernel $K : D \times D \to \mathbb{R}$.

The Problem

- $D \subseteq \mathbb{R}^d$ Borel measurable
- ϱ probability density on D
- H Hilbert space of functions $f : D \to \mathbb{R}$
- **Integration Problem:**

$$I(f) = \int_D f(x)\varrho(x)\,\mathrm{d}x$$

- Problem is well defined iff $H \subset L_1(\varrho)$ iff

$$C^{\mathrm{init}} = \left(\int_D \int_D K(x,y)\varrho(x)\varrho(y)\,\mathrm{d}x\mathrm{d}y\right)^{1/2} < \infty$$

- **Algorithms:** Randomized Algorithm using n function values, in particular importance sampling
- to have function values well defined we assume that H is a **reproducing kernel Hilbert space** with kernel $K : D \times D \to \mathbb{R}$.

The Problem

- $D \subseteq \mathbb{R}^d$ Borel measurable
- ϱ probability density on D
- H Hilbert space of functions $f : D \to \mathbb{R}$
- **Integration Problem:**

$$I(f) = \int_D f(x)\varrho(x)\,\mathrm{d}x$$

- Problem is well defined iff $H \subset L_1(\varrho)$ iff

$$C^{\text{init}} = \left(\int_D \int_D K(x,y)\varrho(x)\varrho(y)\,\mathrm{d}x\mathrm{d}y\right)^{1/2} < \infty$$

- **Algorithms:** Randomized Algorithm using n function values, in particular importance sampling
- to have function values well defined we assume that H is a **reproducing kernel Hilbert space** with kernel $K : D \times D \to \mathbb{R}$.

The Problem

- $D \subseteq \mathbb{R}^d$ Borel measurable
- ϱ probability density on D
- H Hilbert space of functions $f : D \to \mathbb{R}$
- **Integration Problem:**

$$I(f) = \int_D f(x)\varrho(x)\,\mathrm{d}x$$

- Problem is well defined iff $H \subset L_1(\varrho)$ iff

$$C^{\mathrm{init}} = \left(\int_D \int_D K(x,y)\varrho(x)\varrho(y)\,\mathrm{d}x\mathrm{d}y\right)^{1/2} < \infty$$

- **Algorithms:** Randomized Algorithm using n function values, in particular importance sampling
- to have function values well defined we assume that H is a **reproducing kernel Hilbert space** with kernel $K : D \times D \to \mathbb{R}$.

The Problem

- $D \subseteq \mathbb{R}^d$ Borel measurable
- ϱ probability density on D
- H Hilbert space of functions $f : D \to \mathbb{R}$
- **Integration Problem:**

$$I(f) = \int_D f(x)\varrho(x)\,\mathrm{d}x$$

- Problem is well defined iff $H \subset L_1(\varrho)$ iff

$$C^{\mathrm{init}} = \left(\int_D \int_D K(x,y)\varrho(x)\varrho(y)\,\mathrm{d}x\mathrm{d}y\right)^{1/2} < \infty$$

- **Algorithms:** Randomized Algorithm using n function values, in particular importance sampling
- to have function values well defined we assume that H is a **reproducing kernel Hilbert space** with kernel $K : D \times D \to \mathbb{R}$.

Importance Sampling

- another density function ω on D
- **Alternative Integration Problem:**

$$I(f) = \int_D \frac{f(x)\varrho(x)}{\omega(x)}\,\omega(x)\,\mathrm{d}x$$

- **Monte-Carlo:** $x_1, \ldots, x_n$ iid according to probability density ω

$$Q_n(f) = \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)\varrho(x_i)}{\omega(x_i)}$$

- **Error:**

$$\begin{aligned} e_n^2 &= \sup_{\|f\|_H \le 1} \mathbb{E}\,|I(f) - Q_n(f)|^2 \\ &= \frac{1}{n} \sup_{\|f\|_H \le 1} \left(\int_D \frac{f(x)^2 \varrho(x)^2}{\omega(x)}\,\mathrm{d}x - I(f)^2 \right) \end{aligned}$$

Importance Sampling

- another density function ω on D
- **Alternative Integration Problem:**

$$I(f) = \int_D \frac{f(x)\varrho(x)}{\omega(x)}\,\omega(x)\,\mathrm{d}x$$

- **Monte-Carlo:** $x_1, \ldots, x_n$ iid according to probability density ω

$$Q_n(f) = \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)\varrho(x_i)}{\omega(x_i)}$$

- **Error:**

$$\begin{aligned} e_n^2 &= \sup_{\|f\|_H \le 1} \mathbb{E}\,|I(f) - Q_n(f)|^2 \\ &= \frac{1}{n} \sup_{\|f\|_H \le 1} \left(\int_D \frac{f(x)^2 \varrho(x)^2}{\omega(x)}\,\mathrm{d}x - I(f)^2 \right) \end{aligned}$$

Importance Sampling

- another density function ω on D
- **Alternative Integration Problem:**

$$I(f) = \int_D \frac{f(x)\varrho(x)}{\omega(x)}\,\omega(x)\,\mathrm{d}x$$

- **Monte-Carlo:** $x_1, \ldots, x_n$ iid according to probability density ω

$$Q_n(f) = \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)\varrho(x_i)}{\omega(x_i)}$$

- **Error:**

$$\begin{aligned} e_n^2 &= \sup_{\|f\|_H \le 1} \mathbb{E}\,|I(f) - Q_n(f)|^2 \\ &= \frac{1}{n} \sup_{\|f\|_H \le 1} \left(\int_D \frac{f(x)^2 \varrho(x)^2}{\omega(x)}\,\mathrm{d}x - I(f)^2 \right) \end{aligned}$$

Importance Sampling

- another density function ω on D
- **Alternative Integration Problem:**

$$I(f) = \int_D \frac{f(x)\varrho(x)}{\omega(x)}\,\omega(x)\,\mathrm{d}x$$

- **Monte-Carlo:** $x_1, \ldots, x_n$ iid according to probability density ω

$$Q_n(f) = \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)\varrho(x_i)}{\omega(x_i)}$$

- **Error:**

$$\begin{aligned} e_n^2 &= \sup_{\|f\|_H \le 1} \mathbb{E}\,|I(f) - Q_n(f)|^2 \\ &= \frac{1}{n}\sup_{\|f\|_H \le 1} \left(\int_D \frac{f(x)^2 \varrho(x)^2}{\omega(x)}\,\mathrm{d}x - I(f)^2 \right) \end{aligned}$$

Analysis of the error

- Independent of the concrete integral $I(f)$:

$$e_n \leq n^{-1/2}\, C(\omega)$$

where $C(\omega)$ is given by

$$C(\omega) = \left(\sup_{\|f\|_H \leq 1} \int_D \frac{f(x)^2 \varrho(x)^2}{\omega(x)} \,\mathrm{d}x \right)^{1/2}$$

- **Consequence:** Importance sampling has worst case error of order $n^{-1/2}$ if

$$C^{\mathrm{imps}} = \inf_{\omega} C(\omega) < \infty$$

Analysis of the error

- Independent of the concrete integral $I(f)$:

$$e_n \leq n^{-1/2} \, C(\omega)$$

where $C(\omega)$ is given by

$$C(\omega) = \left(\sup_{\|f\|_H \leq 1} \int_D \frac{f(x)^2 \varrho(x)^2}{\omega(x)} \, \mathrm{d}x \right)^{1/2}$$

- **Consequence:** Importance sampling has worst case error of order $n^{-1/2}$ if

$$C^{\mathrm{imps}} = \inf_{\omega} C(\omega) < \infty$$

Known Result (Plaskota, Wasilkowski, Zhao 2009)

- use $|f(t)| \le \sqrt{K(t,t)}$ for $\|f\| \le 1$ to obtain

$$C(\omega) \le \left(\int_D \frac{K(x,x)\varrho^2(x)}{\omega(x)} \, \mathrm{d}x \right)^{1/2}$$

- **Standard Monte-Carlo:** $\omega = \varrho$

$$C^{\mathrm{std}} := \left(\int_D K(x,x)\varrho(x) \, \mathrm{d}x \right)^{1/2} < \infty$$

  sufficient for standard MC to have error of order $n^{-1/2}$

- **Optimization for** ω**:** with the condition

$$C^{\mathrm{sqrt}} = \int_D \sqrt{K(x,x)}\varrho(x) \, \mathrm{d}x < \infty$$

  the density

$$\omega^*(x) = \frac{\sqrt{K(x,x)}\varrho(x)}{C^{\mathrm{sqrt}}}$$

  gives error of order $n^{-1/2}$

Known Result (Plaskota, Wasilkowski, Zhao 2009)

- use $|f(t)| \leq \sqrt{K(t,t)}$ for $\|f\| \leq 1$ to obtain

$$C(\omega) \leq \left(\int_D \frac{K(x,x)\varrho^2(x)}{\omega(x)} \, \mathrm{d}x \right)^{1/2}$$

- **Standard Monte-Carlo:** $\omega = \varrho$

$$C^{\mathrm{std}} := \left(\int_D K(x,x)\varrho(x) \, \mathrm{d}x \right)^{1/2} < \infty$$

sufficient for standard MC to have error of order $n^{-1/2}$

- **Optimization for ω:** with the condition

$$C^{\mathrm{sqrt}} = \int_D \sqrt{K(x,x)}\varrho(x) \, \mathrm{d}x < \infty$$

the density

$$\omega^*(x) = \frac{\sqrt{K(x,x)}\varrho(x)}{C^{\mathrm{sqrt}}}$$

gives error of order $n^{-1/2}$

Known Result (Plaskota, Wasilkowski, Zhao 2009)

- use $|f(t)| \leq \sqrt{K(t,t)}$ for $\|f\| \leq 1$ to obtain

$$C(\omega) \leq \left(\int_D \frac{K(x,x)\varrho^2(x)}{\omega(x)} \,\mathrm{d}x \right)^{1/2}$$

- **Standard Monte-Carlo:** $\omega = \varrho$

$$C^{\mathrm{std}} := \left(\int_D K(x,x)\varrho(x) \,\mathrm{d}x \right)^{1/2} < \infty$$

sufficient for standard MC to have error of order $n^{-1/2}$

- **Optimization for ω:** with the condition

$$C^{\mathrm{sqrt}} = \int_D \sqrt{K(x,x)}\varrho(x) \,\mathrm{d}x < \infty$$

the density

$$\omega^*(x) = \frac{\sqrt{K(x,x)}\varrho(x)}{C^{\mathrm{sqrt}}}$$

gives error of order $n^{-1/2}$

The Result

- $C^{\text{init}} \le C^{\text{imps}} \le C^{\text{sqrt}} \le C^{\text{std}}$
- $C^{\text{std}} < \infty \implies$ Standard MC has error of order $n^{-1/2}$
- $C^{\text{sqrt}} < \infty \implies$ PWZ-importance sampling has error of order $n^{-1/2}$
- $C^{\text{init}} < \infty \iff H \subset L_1(\varrho) \iff$ Problem is well defined $\iff$ $J_H : H \to L_1(\varrho)$ is a bounded operator

The Result

- $C^{\text{init}} \leq C^{\text{imps}} \leq C^{\text{sqrt}} \leq C^{\text{std}}$
- $C^{\text{std}} < \infty \implies$ Standard MC has error of order $n^{-1/2}$
- $C^{\text{sqrt}} < \infty \implies$ PWZ-importance sampling has error of order $n^{-1/2}$
- $C^{\text{init}} < \infty \iff H \subset L_1(\varrho) \iff$ Problem is well defined $\iff$ $J_H : H \to L_1(\varrho)$ is a bounded operator

The Result

- $C^{\mathrm{init}} \leq C^{\mathrm{imps}} \leq C^{\mathrm{sqrt}} \leq C^{\mathrm{std}}$
- $C^{\mathrm{std}} < \infty \Longrightarrow$ Standard MC has error of order $n^{-1/2}$
- $C^{\mathrm{sqrt}} < \infty \Longrightarrow$ PWZ-importance sampling has error of order $n^{-1/2}$
- $C^{\mathrm{init}} < \infty \Longleftrightarrow H \subset L_1(\varrho) \Longleftrightarrow$ Problem is well defined $\Longleftrightarrow$ $J_H : H \to L_1(\varrho)$ is a bounded operator

The Result

- $C^{\text{init}} \leq C^{\text{imps}} \leq C^{\text{sqrt}} \leq C^{\text{std}}$
- $C^{\text{std}} < \infty \implies$ Standard MC has error of order $n^{-1/2}$
- $C^{\text{sqrt}} < \infty \implies$ PWZ-importance sampling has error of order $n^{-1/2}$
- $C^{\text{init}} < \infty \iff H \subset L_1(\varrho) \iff$ Problem is well defined $\iff$ $J_H : H \to L_1(\varrho)$ is a bounded operator

The Result

- $C^{\text{init}} \leq C^{\text{imps}} \leq C^{\text{sqrt}} \leq C^{\text{std}}$
- $C^{\text{std}} < \infty \Longrightarrow$ Standard MC has error of order $n^{-1/2}$
- $C^{\text{sqrt}} < \infty \Longrightarrow$ PWZ-importance sampling has error of order $n^{-1/2}$
- $C^{\text{init}} < \infty \Longleftrightarrow H \subset L_1(\varrho) \Longleftrightarrow$ Problem is well defined $\Longleftrightarrow$ $J_H : H \to L_1(\varrho)$ is a bounded operator

Theorem

If the initial error is finite or, equivalently, if the embedding $J_H : H \to L_1(\varrho)$ is a bounded operator then importance sampling has error of order $n^{-1/2}$. More precisely,

$$C^{\text{imps}} \leq \sqrt{\frac{\pi}{2}}\, \|J_H : H \to L_1(\varrho)\|.$$

In particular, if the kernel K is nonnegative then

$$C^{\text{imps}} \leq \sqrt{\frac{\pi}{2}}\, C^{\text{init}}.$$

p-Summing Operators

- $1 \leq p < \infty$
- X, Y Banach spaces, $T : X \to Y$ bounded linear operator
- T is called p-summing if T maps weakly p-summable sequences in X to strongly p-summable sequences in Y.
- $$\sum_{i=1}^{n} \|Tx_i\|^p \leq c^p \sup_{\|a\|_{X'} \leq 1} \sum_{i=1}^{n} |a(x_i)|^p$$
- $\pi_p(T) = \inf c$
- **Pietsch Domination Theorem:** $T : X \to Y$ is p-summing if and only if there exists a constant $c \geq 0$ and a regular Borel probability measure ν on the weak-$*$-compact closed unit ball $B_{X'}$ of X' such that for all $x \in X$

$$\|Tx\|^p \leq c^p \int_{B_{X'}} |a(x)|^p \, \mathrm{d}\nu(a)$$

p-Summing Operators

- $1 \leq p < \infty$
- X, Y Banach spaces, $T : X \to Y$ bounded linear operator
- T is called p-summing if T maps weakly p-summable sequences in X to strongly p-summable sequences in Y.
- $$\sum_{i=1}^{n} \|Tx_i\|^p \leq c^p \sup_{\|a\|_{X'} \leq 1} \sum_{i=1}^{n} |a(x_i)|^p$$
- $\pi_p(T) = \inf c$
- **Pietsch Domination Theorem:** $T : X \to Y$ is p-summing if and only if there exists a constant $c \geq 0$ and a regular Borel probability measure ν on the weak-$*$-compact closed unit ball $B_{X'}$ of X' such that for all $x \in X$

$$\|Tx\|^p \leq c^p \int_{B_{X'}} |a(x)|^p \, \mathrm{d}\nu(a)$$

p-Summing Operators

- $1 \leq p < \infty$
- X, Y Banach spaces, $T : X \to Y$ bounded linear operator
- T is called p-summing if T maps weakly p-summable sequences in X to strongly p-summable sequences in Y.
- $$\sum_{i=1}^{n} \|Tx_i\|^p \leq c^p \sup_{\|a\|_{X'} \leq 1} \sum_{i=1}^{n} |a(x_i)|^p$$
- $\pi_p(T) = \inf c$
- **Pietsch Domination Theorem:** $T : X \to Y$ is p-summing if and only if there exists a constant $c \geq 0$ and a regular Borel probability measure ν on the weak-$*$-compact closed unit ball $B_{X'}$ of X' such that for all $x \in X$

$$\|Tx\|^p \leq c^p \int_{B_{X'}} |a(x)|^p \, \mathrm{d}\nu(a)$$

p-Summing Operators

- $1 \leq p < \infty$
- X, Y Banach spaces, $T : X \to Y$ bounded linear operator
- T is called p-summing if T maps weakly p-summable sequences in X to strongly p-summable sequences in Y.
- $$\sum_{i=1}^{n} \|Tx_i\|^p \leq c^p \sup_{\|a\|_{X'} \leq 1} \sum_{i=1}^{n} |a(x_i)|^p$$
- $\pi_p(T) = \inf c$
- **Pietsch Domination Theorem:** $T : X \to Y$ is p-summing if and only if there exists a constant $c \geq 0$ and a regular Borel probability measure ν on the weak-$*$-compact closed unit ball $B_{X'}$ of X' such that for all $x \in X$

$$\|Tx\|^p \leq c^p \int_{B_{X'}} |a(x)|^p \, \mathrm{d}\nu(a)$$

p-Summing Operators

- $1 \leq p < \infty$
- X, Y Banach spaces, $T : X \to Y$ bounded linear operator
- T is called p-summing if T maps weakly p-summable sequences in X to strongly p-summable sequences in Y.
- $$\sum_{i=1}^{n} \|Tx_i\|^p \leq c^p \sup_{\|a\|_{X'} \leq 1} \sum_{i=1}^{n} |a(x_i)|^p$$
- $\pi_p(T) = \inf c$
- **Pietsch Domination Theorem:** $T : X \to Y$ is p-summing if and only if there exists a constant $c \geq 0$ and a regular Borel probability measure ν on the weak-$*$-compact closed unit ball $B_{X'}$ of X' such that for all $x \in X$

$$\|Tx\|^p \leq c^p \int_{B_{X'}} |a(x)|^p \, \mathrm{d}\nu(a)$$

p-Summing Operators

- $1 \leq p < \infty$
- X, Y Banach spaces, $T : X \to Y$ bounded linear operator
- T is called p-summing if T maps weakly p-summable sequences in X to strongly p-summable sequences in Y.
-
$$\sum_{i=1}^{n} \|Tx_i\|^p \leq c^p \sup_{\|a\|_{X'} \leq 1} \sum_{i=1}^{n} |a(x_i)|^p$$
- $\pi_p(T) = \inf c$
- **Pietsch Domination Theorem:** $T : X \to Y$ is p-summing if and only if there exists a constant $c \geq 0$ and a regular Borel probability measure ν on the weak-$*$-compact closed unit ball $B_{X'}$ of X' such that for all $x \in X$

$$\|Tx\|^p \leq c^p \int_{B_{X'}} |a(x)|^p \, \mathrm{d}\nu(a)$$

Change of Density

Theorem (Rosenthal, Johnson/Schechtman)

$X \subset L_1(\Omega, \mu)$, μ is a probability measure, $J : X \to L_1(\mu)$ bounded embedding. If the dual operator $J' : L_\infty(\Omega, \mu) \to X'$ is q-summing for some $1 \le q < \infty$ then there exists a measurable function $g > 0$ on Ω such that $\int_\Omega g \, \mathrm{d}\mu = 1$ and such that the isometry

$$M : L_1(\Omega, \mu) \to L_1(\Omega, g \, \mathrm{d}\mu) \quad \textit{given by } Mf = fg^{-1}$$

maps X to a space $\widetilde{X} = M(X)$ which is contained in $L_p(\Omega, g \, \mathrm{d}\mu)$, where p is the dual index of q defined as $1/p + 1/q = 1$.
Moreover, if we equip $\widetilde{X}$ with the norm from X, i.e. if we set

$$\|Mf | \widetilde{X}\| = \|f | X\| \quad \textit{for } f \in X,$$

then the embedding $\widetilde{J} : \widetilde{X} \to L_p(\Omega, g \, \mathrm{d}\mu)$ has norm

$$\|\widetilde{J} : \widetilde{X} \to L_p(\Omega, g \, \mathrm{d}\mu)\| \le \pi_q(J' : L_\infty(\Omega, \mu) \to X').$$

The Method

- $C^{\text{init}} < \infty$ means $H \subset L_1(\varrho)$
- we want to change the density so that $H \subset L_2$
- recall: $J_H : H \to L_1(\varrho)$ is a linear bounded operator
- the **Little Grothendieck Theorem** tells you that the dual operator is 2-summing
- the **Change of Density Theorem** tells you that then the measure ϱdx can be changed with a density so that H then actually becomes a subspace of L_2
- that is exactly what we need
- the density ω for the importance sampling algorithm can be obtained from the **Pietsch measure** in the **Pietsch Domination Theorem** associated with the 2-summing operator J_H'

The Method

- $C^{\text{init}} < \infty$ means $H \subset L_1(\varrho)$
- we want to change the density so that $H \subset L_2$
- recall: $J_H : H \to L_1(\varrho)$ is a linear bounded operator
- the **Little Grothendieck Theorem** tells you that the dual operator is 2-summing
- the **Change of Density Theorem** tells you that then the measure $\varrho \mathrm{d}x$ can be changed with a density so that H then actually becomes a subspace of L_2
- that is exactly what we need
- the density ω for the importance sampling algorithm can be obtained from the **Pietsch measure** in the **Pietsch Domination Theorem** associated with the 2-summing operator J'_H

The Method

- $C^{\text{init}} < \infty$ means $H \subset L_1(\varrho)$
- we want to change the density so that $H \subset L_2$
- recall: $J_H : H \to L_1(\varrho)$ is a linear bounded operator
- the **Little Grothendieck Theorem** tells you that the dual operator is 2-summing
- the **Change of Density Theorem** tells you that then the measure $\varrho \mathrm{d}x$ can be changed with a density so that H then actually becomes a subspace of L_2
- that is exactly what we need
- the density ω for the importance sampling algorithm can be obtained from the **Pietsch measure** in the **Pietsch Domination Theorem** associated with the 2-summing operator J'_H

The Method

- $C^{\text{init}} < \infty$ means $H \subset L_1(\varrho)$
- we want to change the density so that $H \subset L_2$
- recall: $J_H : H \to L_1(\varrho)$ is a linear bounded operator
- the **Little Grothendieck Theorem** tells you that the dual operator is 2-summing
- the **Change of Density Theorem** tells you that then the measure $\varrho \mathrm{d}x$ can be changed with a density so that H then actually becomes a subspace of L_2
- that is exactly what we need
- the density ω for the importance sampling algorithm can be obtained from the **Pietsch measure** in the **Pietsch Domination Theorem** associated with the 2-summing operator J'_H

The Method

- $C^{\text{init}} < \infty$ means $H \subset L_1(\varrho)$
- we want to change the density so that $H \subset L_2$
- recall: $J_H : H \to L_1(\varrho)$ is a linear bounded operator
- the **Little Grothendieck Theorem** tells you that the dual operator is 2-summing
- the **Change of Density Theorem** tells you that then the measure $\varrho \mathrm{d}x$ can be changed with a density so that H then actually becomes a subspace of L_2
- that is exactly what we need
- the density ω for the importance sampling algorithm can be obtained from the **Pietsch measure** in the **Pietsch Domination Theorem** associated with the 2-summing operator J'_H

The Method

- $C^{\text{init}} < \infty$ means $H \subset L_1(\varrho)$
- we want to change the density so that $H \subset L_2$
- recall: $J_H : H \to L_1(\varrho)$ is a linear bounded operator
- the **Little Grothendieck Theorem** tells you that the dual operator is 2-summing
- the **Change of Density Theorem** tells you that then the measure $\varrho \mathrm{d}x$ can be changed with a density so that H then actually becomes a subspace of L_2
- that is exactly what we need
- the density ω for the importance sampling algorithm can be obtained from the **Pietsch measure** in the **Pietsch Domination Theorem** associated with the 2-summing operator J'_H

The Method

- $C^{\text{init}} < \infty$ means $H \subset L_1(\varrho)$
- we want to change the density so that $H \subset L_2$
- recall: $J_H : H \to L_1(\varrho)$ is a linear bounded operator
- the **Little Grothendieck Theorem** tells you that the dual operator is 2-summing
- the **Change of Density Theorem** tells you that then the measure $\varrho \mathrm{d}x$ can be changed with a density so that H then actually becomes a subspace of L_2
- that is exactly what we need
- the density ω for the importance sampling algorithm can be obtained from the **Pietsch measure** in the **Pietsch Domination Theorem** associated with the 2-summing operator J_H'

Tractability of Multivariate Integration

Theorem

Assume that we have a sequence I_d of d-dimensional integration problems with normalized initial error. If the embeddings $J_{H_d} : H_d \to L_1(\varrho_d)$ are uniformly bounded, then the multivariate weighted integration problem is strongly polynomially tractable in the randomized setting with exponent 2. This is in particular the case if all the kernels K_d are nonnegative.

- **Novak, Wozniakowski:** The exponent 2 is sharp for tensor product Hilbert spaces whose univariate reproducing kernel is nonnegative and decomposable and univariate integration is not trivial for the two spaces corresponding to decomposable parts.

Tractability of Multivariate Integration

Theorem

Assume that we have a sequence I_d of d-dimensional integration problems with normalized initial error. If the embeddings $J_{H_d} : H_d \to L_1(\varrho_d)$ are uniformly bounded, then the multivariate weighted integration problem is strongly polynomially tractable in the randomized setting with exponent 2. This is in particular the case if all the kernels K_d are nonnegative.

- **Novak, Wozniakowski:** The exponent 2 is sharp for tensor product Hilbert spaces whose univariate reproducing kernel is nonnegative and decomposable and univariate integration is not trivial for the two spaces corresponding to decomposable parts.

Toy Example

- $D = [0, \infty)$
- $K(x,y) = \sum_{j=1}^{\infty} a_j^2 \, \mathbf{1}_j(x) \, \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \, \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\mathrm{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\mathrm{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\mathrm{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\mathrm{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\mathrm{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \, \mathbf{1}_j(x)$$

leads to $C^{\mathrm{imps}} = C^{\mathrm{init}}$

Toy Example

- $D = [0, \infty)$
- $K(x,y) = \sum_{j=1}^{\infty} a_j^2 \, \mathbf{1}_j(x) \, \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \, \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\mathrm{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\mathrm{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\mathrm{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\mathrm{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\mathrm{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \, \mathbf{1}_j(x)$$

leads to $C^{\mathrm{imps}} = C^{\mathrm{init}}$

Toy Example

- $D = [0, \infty)$
- $K(x, y) = \sum_{j=1}^{\infty} a_j^2 \, \mathbf{1}_j(x) \, \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \, \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\mathrm{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\mathrm{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\mathrm{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\mathrm{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\mathrm{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \, \mathbf{1}_j(x)$$

leads to $C^{\mathrm{imps}} = C^{\mathrm{init}}$

Toy Example

- $D = [0, \infty)$
- $K(x,y) = \sum_{j=1}^{\infty} a_j^2 \, \mathbf{1}_j(x) \, \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \, \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\text{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\text{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\text{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\text{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\text{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \, \mathbf{1}_j(x)$$

leads to $C^{\text{imps}} = C^{\text{init}}$

Toy Example

- $D = [0, \infty)$
- $K(x, y) = \sum_{j=1}^{\infty} a_j^2 \, \mathbf{1}_j(x) \, \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \, \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\mathrm{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\mathrm{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\mathrm{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\mathrm{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\mathrm{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \, \mathbf{1}_j(x)$$

leads to $C^{\mathrm{imps}} = C^{\mathrm{init}}$

Toy Example

- $D = [0, \infty)$
- $K(x, y) = \sum_{j=1}^{\infty} a_j^2 \mathbf{1}_j(x) \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\mathrm{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\mathrm{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\mathrm{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\mathrm{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\mathrm{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \mathbf{1}_j(x)$$

leads to $C^{\mathrm{imps}} = C^{\mathrm{init}}$

Toy Example

- $D = [0, \infty)$
- $K(x, y) = \sum_{j=1}^{\infty} a_j^2 \, \mathbf{1}_j(x) \, \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \, \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\mathrm{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\mathrm{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\mathrm{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\mathrm{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\mathrm{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \, \mathbf{1}_j(x)$$

leads to $C^{\mathrm{imps}} = C^{\mathrm{init}}$

Toy Example

- $D = [0, \infty)$
- $K(x,y) = \sum_{j=1}^{\infty} a_j^2 \, \mathbf{1}_j(x) \, \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \, \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\mathrm{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\mathrm{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\mathrm{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\mathrm{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\mathrm{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \, \mathbf{1}_j(x)$$

leads to $C^{\mathrm{imps}} = C^{\mathrm{init}}$

Toy Example

- $D = [0, \infty)$
- $K(x, y) = \sum_{j=1}^{\infty} a_j^2 \, \mathbf{1}_j(x) \, \mathbf{1}_j(y)$
- $\varrho(x) = \sum_{j=1}^{\infty} r_j \, \mathbf{1}_j(x)$ for some $r_j \geq 0$ summing to 1
- The functions $a_j \mathbf{1}_j$ are an orthonormal basis of H.
- $C^{\mathrm{init}} < \infty \iff (a_j r_j) \in \ell_2$
- $C^{\mathrm{sqrt}} < \infty \iff (a_j r_j) \in \ell_1$
- $C^{\mathrm{std}} < \infty \iff (a_j^2 r_j) \in \ell_1$
- $C^{\mathrm{imps}} < \infty \iff (a_j r_j) \in \ell_2$
- direct construction of the density

$$\omega(x) = \left(C^{\mathrm{init}}\right)^{-2} \sum_{j=1}^{\infty} a_j^2 r_j^2 \, \mathbf{1}_j(x)$$

leads to $C^{\mathrm{imps}} = C^{\mathrm{init}}$

Finally - A Real Example

- tractability of uniform integration on weighted Sobolev spaces
- $K_d(x,t) = \prod_{j=1}^{d} K^{\gamma_j}(x_j, t_j)$ for $x, t \in [0,1]^d$
- $K^\gamma(x,t) = 1 + \gamma \min(x,t)$ (non-periodic case) or $K^\gamma(x,t) = 1 + \gamma(\min(x,t) - xt)$ (periodic case) for $x, t \in [0,1]$
- $H_d = \otimes_{j=1}^{d} H^{\gamma_j}$ with H^γ consisting of absolutely continuous functions on $[0,1]$ whith first derivatives in $L_2[0,1]$ with norm

$$\|f\|_{H^\gamma}^2 = |f(0)|^2 + \frac{1}{\gamma} \int_0^1 |f'(x)|^2 \, \mathrm{d}x$$

- consider the normalized error criterion

Finally - A Real Example

- tractability of uniform integration on weighted Sobolev spaces
- $K_d(x,t) = \prod_{j=1}^d K^{\gamma_j}(x_j,t_j)$ for $x,t \in [0,1]^d$
- $K^{\gamma}(x,t) = 1 + \gamma \min(x,t)$ (non-periodic case) or $K^{\gamma}(x,t) = 1 + \gamma(\min(x,t) - xt)$ (periodic case) for $x,t \in [0,1]$
- $H_d = \otimes_{j=1}^d H^{\gamma_j}$ with H^{γ} consisting of absolutely continuous functions on $[0,1]$ whith first derivatives in $L_2[0,1]$ with norm

$$\|f\|_{H^\gamma}^2 = |f(0)|^2 + \frac{1}{\gamma}\int_0^1 |f'(x)|^2 \,\mathrm{d}x$$

- consider the normalized error criterion

Finally - A Real Example

- tractability of uniform integration on weighted Sobolev spaces
- $K_d(x,t) = \prod_{j=1}^{d} K^{\gamma_j}(x_j, t_j)$ for $x, t \in [0,1]^d$
- $K^{\gamma}(x,t) = 1 + \gamma \min(x,t)$ (non-periodic case) or $K^{\gamma}(x,t) = 1 + \gamma(\min(x,t) - xt)$ (periodic case) for $x, t \in [0,1]$
- $H_d = \otimes_{j=1}^{d} H^{\gamma_j}$ with H^{γ} consisting of absolutely continuous functions on $[0,1]$ whith first derivatives in $L_2[0,1]$ with norm

$$\|f\|_{H^{\gamma}}^2 = |f(0)|^2 + \frac{1}{\gamma}\int_0^1 |f'(x)|^2 \, \mathrm{d}x$$

- consider the normalized error criterion

Finally - A Real Example

- tractability of uniform integration on weighted Sobolev spaces
- $K_d(x,t) = \prod_{j=1}^d K^{\gamma_j}(x_j,t_j)$ for $x,t \in [0,1]^d$
- $K^\gamma(x,t) = 1 + \gamma \min(x,t)$ (non-periodic case) or $K^\gamma(x,t) = 1 + \gamma(\min(x,t) - xt)$ (periodic case) for $x,t \in [0,1]$
- $H_d = \otimes_{j=1}^d H^{\gamma_j}$ with H^γ consisting of absolutely continuous functions on $[0,1]$ whith first derivatives in $L_2[0,1]$ with norm

$$\|f\|_{H^\gamma}^2 = |f(0)|^2 + \frac{1}{\gamma}\int_0^1 |f'(x)|^2 \, \mathrm{d}x$$

- consider the normalized error criterion

Finally - A Real Example

- tractability of uniform integration on weighted Sobolev spaces
- $K_d(x,t) = \prod_{j=1}^d K^{\gamma_j}(x_j,t_j)$ for $x,t \in [0,1]^d$
- $K^\gamma(x,t) = 1 + \gamma \min(x,t)$ (non-periodic case) or $K^\gamma(x,t) = 1 + \gamma(\min(x,t) - xt)$ (periodic case) for $x,t \in [0,1]$
- $H_d = \otimes_{j=1}^d H^{\gamma_j}$ with H^γ consisting of absolutely continuous functions on $[0,1]$ whith first derivatives in $L_2[0,1]$ with norm

$$\|f\|_{H^\gamma}^2 = |f(0)|^2 + \frac{1}{\gamma}\int_0^1 |f'(x)|^2 \,\mathrm{d}x$$

- consider the normalized error criterion

Finally - A Real Example

- Sloan/Wozniakowski 2004: strongly polynomially tractable in the deterministic setting iff $\sum \gamma_j < \infty$
- Sloan/Wozniakowski 2004: standard MC is strongly polynomial iff $\sum \gamma_j^2 < \infty$
- Plaskota, Wasilkowski, Zhao 2009: importance sampling is strongly polynomial for the periodic case if $\sum \gamma_j^3 < \infty$
- Now: importance sampling is strongly polynomial without any condition in both the periodic and the non-periodic case

Finally - A Real Example

- Sloan/Wozniakowski 2004: strongly polynomially tractable in the deterministic setting iff $\sum \gamma_j < \infty$
- Sloan/Wozniakowski 2004: standard MC is strongly polynomial iff $\sum \gamma_j^2 < \infty$
- Plaskota, Wasilkowski, Zhao 2009: importance sampling is strongly polynomial for the periodic case if $\sum \gamma_j^3 < \infty$
- Now: importance sampling is strongly polynomial without any condition in both the periodic and the non-periodic case

Finally - A Real Example

- Sloan/Wozniakowski 2004: strongly polynomially tractable in the deterministic setting iff $\sum \gamma_j < \infty$
- Sloan/Wozniakowski 2004: standard MC is strongly polynomial iff $\sum \gamma_j^2 < \infty$
- Plaskota, Wasilkowski, Zhao 2009: importance sampling is strongly polynomial for the periodic case if $\sum \gamma_j^3 < \infty$
- Now: importance sampling is strongly polynomial without any condition in both the periodic and the non-periodic case

Finally - A Real Example

- Sloan/Wozniakowski 2004: strongly polynomially tractable in the deterministic setting iff $\sum \gamma_j < \infty$
- Sloan/Wozniakowski 2004: standard MC is strongly polynomial iff $\sum \gamma_j^2 < \infty$
- Plaskota, Wasilkowski, Zhao 2009: importance sampling is strongly polynomial for the periodic case if $\sum \gamma_j^3 < \infty$
- Now: importance sampling is strongly polynomial without any condition in both the periodic and the non-periodic case

The End

HERZLICHEN GLÜCKWUNSCH!