Transformations and Hardy-Krause Variation

Kinjal Basu

Joint work with Prof. Art Owen

Department of Statistics
Stanford University

Workshop in Discrepancy Theory
Varenna, Italy

16th June, 2016

Introduction

- Quasi-Monte Carlo (QMC) on $[0, 1]^m$.
- Triangles, disks, simplices, spheres, balls, etc are also of importance in applications.

Introduction

- Quasi-Monte Carlo (QMC) on $[0,1]^m$.
- Triangles, disks, simplices, spheres, balls, etc are also of importance in applications.
- In MC, find a transformation, $\tau : [0,1]^m \to \mathcal{X} \subset \mathbb{R}^d$ which is measure preserving.

Introduction

- Quasi-Monte Carlo (QMC) on $[0,1]^m$.
- Triangles, disks, simplices, spheres, balls, etc are also of importance in applications.
- In MC, find a transformation, $\tau : [0,1]^m \to \mathcal{X} \subset \mathbb{R}^d$ which is measure preserving.
- To estimate $\mu = \int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ we use

$$\hat{\mu} = \frac{\mathsf{vol}(\mathcal{X})}{n} \sum_{i=1}^{n} f(\tau(\boldsymbol{u}_i)) \qquad \text{for } \boldsymbol{u}_i \overset{\text{iid}}{\sim} \mathbf{U}[0,1]^m$$

Introduction

- Quasi-Monte Carlo (QMC) on $[0,1]^m$.
- Triangles, disks, simplices, spheres, balls, etc are also of importance in applications.
- In MC, find a transformation, $\tau : [0,1]^m \to \mathcal{X} \subset \mathbb{R}^d$ which is measure preserving.
- To estimate $\mu = \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$ we use

$$\hat{\mu} = \frac{\mathbf{vol}(\mathcal{X})}{n} \sum_{i=1}^{n} f(\tau(\boldsymbol{u}_i)) \qquad \text{for } \boldsymbol{u}_i \overset{\text{iid}}{\sim} \mathbf{U}[0,1]^m$$

- $O(1/\sqrt{n})$.

QMC Approach: Striving for a better rate

QMC Approach: Striving for a better rate

- Use the same transformation τ as in MC.
- Plug in QMC or randomized QMC (RQMC) instead of random $\boldsymbol{u}_i$.

QMC Approach: Striving for a better rate

- Use the same transformation τ as in MC.
- Plug in QMC or randomized QMC (RQMC) instead of random $\boldsymbol{u}_i$.
- Using Koskma-Hlawka inequality,

$$\begin{aligned}\left|\int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{1}{n}\sum_{i=1}^{n} f(\tau(\boldsymbol{u}_i))\right| &= \left|\int_{[0,1]^d} f(\tau(\boldsymbol{u}))\,\mathrm{d}\boldsymbol{u} - \frac{1}{n}\sum_{i=1}^{n} f(\tau(\boldsymbol{u}_i))\right| \\ &\le D_n^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)\mathrm{V}_{\mathrm{HK}}(f \circ \tau)\end{aligned}$$

QMC Approach: Striving for a better rate

- Use the same transformation τ as in MC.
- Plug in QMC or randomized QMC (RQMC) instead of random $\boldsymbol{u}_i$.
- Using Koskma-Hlawka inequality,

$$\left|\int_{\mathcal{X}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{1}{n}\sum_{i=1}^{n} f(\tau(\boldsymbol{u}_i))\right| = \left|\int_{[0,1]^d} f(\tau(\boldsymbol{u}))\,\mathrm{d}\boldsymbol{u} - \frac{1}{n}\sum_{i=1}^{n} f(\tau(\boldsymbol{u}_i))\right|$$
$$\leq D_n^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)\mathrm{V}_{\mathrm{HK}}(f \circ \tau)$$

- If $\mathrm{V}_{\mathrm{HK}}(f \circ \tau) < \infty$, we can attain $O(n^{-1+\epsilon})$.

QMC Approach: Striving for a better rate

- Use the same transformation τ as in MC.
- Plug in QMC or randomized QMC (RQMC) instead of random $\boldsymbol{u}_i$.
- Using Koskma-Hlawka inequality,

$$\left| \int_{\mathcal{X}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=1}^{n} f(\tau(\boldsymbol{u}_i)) \right| = \left| \int_{[0,1]^d} f(\tau(\boldsymbol{u})) \, \mathrm{d}\boldsymbol{u} - \frac{1}{n} \sum_{i=1}^{n} f(\tau(\boldsymbol{u}_i)) \right| \le D_n^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \mathrm{V}_{\mathrm{HK}}(f \circ \tau)$$

- If $\mathrm{V}_{\mathrm{HK}}(f \circ \tau) < \infty$, we can attain $O(n^{-1+\epsilon})$.
- Under additional smoothness RQMC methods (scrambled nets) can yield $O(n^{-3/2+\epsilon})$.

The Problem

- Does this always work?

The Problem

- Does this always work?
- Answer : Sadly no!

The Problem

- Does this always work?
- Answer : Sadly no!
- Can we find conditions on which this technique works?

The Problem

- Does this always work?
- Answer : Sadly no!
- Can we find conditions on which this technique works?
- Answer : YES!

Overview

1. Smoothness Conditions
   - Function Composition
2. Necessary and Sufficient Conditions
3. Counter-Examples
   - Infinite Hardy-Krause Variation
   - Non L^2 Mapping
4. Non-Uniform Transformations

Overview

1. Smoothness Conditions
   - Function Composition
2. Necessary and Sufficient Conditions
3. Counter-Examples
   - Infinite Hardy-Krause Variation
   - Non L^2 Mapping
4. Non-Uniform Transformations

Smoothness Conditions

- QMC attains an error rate of $O(n^{-1}(\log n)^{m-1})$, if $f \in \mathrm{BVHK}$.

Smoothness Conditions

- QMC attains an error rate of $O(n^{-1}(\log n)^{m-1})$, if $f \in \mathrm{BVHK}$.
- If the mixed partial derivative $\partial^{1:m} f$ exists then

$$\mathrm{V}_{\mathrm{HK}}(f) \le \sum_{u \neq \emptyset} \int_{[0,1]^{|u|}} |\partial^u f(\boldsymbol{x}_u{:}\boldsymbol{1}_{-u})| \,\mathrm{d}\boldsymbol{x}_u.$$

[Owen (2005)]

Smoothness Conditions

- QMC attains an error rate of $O(n^{-1}(\log n)^{m-1})$, if $f \in \mathrm{BVHK}$.
- If the mixed partial derivative $\partial^{1:m} f$ exists then

$$\mathrm{V}_{\mathrm{HK}}(f) \le \sum_{u \neq \emptyset} \int_{[0,1]^{|u|}} |\partial^u f(\boldsymbol{x}_u{:}\mathbf{1}_{-u})| \,\mathrm{d}\boldsymbol{x}_u.$$

[Owen (2005)]

- For scrambled nets to attain $O(n^{-3/2}(\log n)^{(m-1)/2})$, f must be smooth in the following sense.

$$\|\partial^u f\|_2^2 \equiv \int (\partial^u f(\boldsymbol{x}))^2 \,\mathrm{d}\boldsymbol{x} < \infty, \quad \forall u \subseteq 1{:}m.$$

[Dick and Pillichshammer (2010)]

Function Composition

- We want condition under which $f \circ \tau : [0,1]^m \to \mathbb{R}^d \to \mathbb{R}$ is in BVHK.

Function Composition

- We want condition under which $f \circ \tau : [0, 1]^m \to \mathbb{R}^d \to \mathbb{R}$ is in BVHK.
- If $d = m = 1$ we reduce to the case of ordinary BV.

Function Composition

- We want condition under which $f \circ \tau : [0,1]^m \to \mathbb{R}^d \to \mathbb{R}$ is in BVHK.
- If $d = m = 1$ we reduce to the case of ordinary BV.
- If τ is of bounded variation and f is Lipschitz, then $f \circ \tau$ is of bounded variation. **[Josephy (1981)]**

Function Composition

- We want condition under which $f \circ \tau : [0,1]^m \to \mathbb{R}^d \to \mathbb{R}$ is in BVHK.
- If $d = m = 1$ we reduce to the case of ordinary BV.
- If τ is of bounded variation and f is Lipschitz, then $f \circ \tau$ is of bounded variation. **[Josephy (1981)]**
- Not the case for BVHK in higher dimensions.

A counter-example

A counter-example

- Let τ be the identity map on $[0,1]^2$ so that both τ_1 and τ_2 are in BVHK.

A counter-example

- Let τ be the identity map on $[0,1]^2$ so that both τ_1 and τ_2 are in BVHK.
- Then we construct a Lipschitz function $f : [0,1]^2 \to \mathbb{R}$ with $f \circ \tau = f \notin \mathrm{BVHK}$.

Sierpenkski function

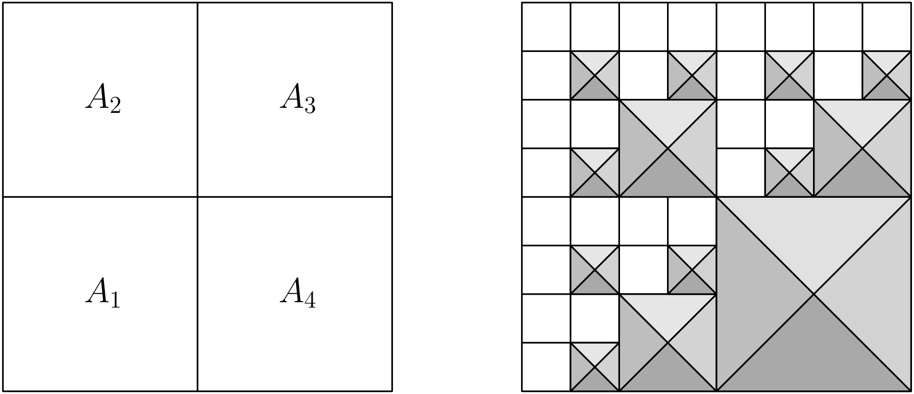

Figure: The plot on the left shows the square partition $\mathcal{P}$ which is repeated in a recursive manner. The right figure shows the function as a 2-dimensional projection for $k = 3$. Each such pyramidal structure has a height of half the length of its base square.

Results

Lemma 1

The function f is Lipschitz on $[0,1]^2$ with respect to the Euclidean norm.

Lemma 2

The function $f \notin \mathrm{BVHK}$. If we define a d-dimensional function $f_d(x_1,\dots,x_d) := f(x_1,x_2)$, then f_d is Lipschitz on $[0,1]^d$ but $f_d \notin \mathrm{BVHK}$.

Faa di Bruno formula

Faa di Bruno formula

- Remember that,

$$\mathrm{V}_{\mathrm{HK}}(f) \le \sum_{u \neq \emptyset} \int_{[0,1]^{|u|}} |\partial^u f(\boldsymbol{x}_u{:}\boldsymbol{1}_{-u})| \,\mathrm{d}\boldsymbol{x}_u.$$

Faa di Bruno formula

- Remember that,

$$\mathrm{V}_{\mathrm{HK}}(f) \le \sum_{u \ne \emptyset} \int_{[0,1]^{|u|}} |\partial^u f(\boldsymbol{x}_u{:}\boldsymbol{1}_{-u})| \, \mathrm{d}\boldsymbol{x}_u.$$

- Study $f \circ \tau$ using the Multivariate Faa di Bruno formula due to Constantine and Savits (1996).

Faa di Bruno formula

- Remember that,

$$\mathrm{V}_{\mathrm{HK}}(f) \leq \sum_{u \neq \emptyset} \int_{[0,1]^{|u|}} |\partial^u f(\boldsymbol{x}_u{:}\boldsymbol{1}_{-u})| \, \mathrm{d}\boldsymbol{x}_u.$$

- Study $f \circ \tau$ using the Multivariate Faa di Bruno formula due to Constantine and Savits (1996).
- $\tau : [0,1]^m \to \mathcal{X} \subset \mathbb{R}^d$ and $f : \mathcal{X} \to \mathbb{R}$.

Faa di Bruno formula

- Remember that,

$$\mathrm{V}_{\mathrm{HK}}(f) \le \sum_{u \neq \emptyset} \int_{[0,1]^{|u|}} |\partial^u f(\boldsymbol{x}_u{:}\boldsymbol{1}_{-u})| \, \mathrm{d}\boldsymbol{x}_u.$$

- Study $f \circ \tau$ using the Multivariate Faa di Bruno formula due to Constantine and Savits (1996).
- $\tau : [0,1]^m \to \mathcal{X} \subset \mathbb{R}^d$ and $f : \mathcal{X} \to \mathbb{R}$.
- Let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d) \in \mathbb{N}_0^d$. Then $f_{\boldsymbol{\lambda}}$ is the derivative of f taken λ_i times with respect to x_i.

Multivariate Faa di Bruno formula

Multivariate Faa di Bruno formula

- For any $v \subseteq 1:m$,

$$\partial^v (f \circ \tau) = \sum_{\substack{\boldsymbol{\lambda} \in \mathbb{N}_0^d \\ 1 \le |\boldsymbol{\lambda}| \le |v|}} f_{\boldsymbol{\lambda}} \sum_{s=1}^{|v|} \sum_{(\ell_r, k_r) \in \widetilde{\mathrm{KL}}(s, v, \boldsymbol{\lambda})} \prod_{r=1}^{s} \partial^{\ell_r} \tau_{k_r}$$

where $\widetilde{\mathrm{KL}}(s, v, \boldsymbol{\lambda})$ equals

$$\Big\{ (\ell_r, k_r),\ r = 1, \dots, s, \ \Big|\ \ell_r \subseteq 1{:}m,\ k_r \in 1{:}d,\ \cup_{r=1}^{s} \ell_r = v, \\ \ell_r \cap \ell_{r'} = \emptyset \text{ for } r \neq r' \text{ and } |\{ j \in 1{:}s \mid k_j = i \}| = \lambda_i \Big\}.$$

Overview

Main Result for QMC point set

Theorem 1. B and Owen (2016)

Let $\tau(\boldsymbol{u})$ be as described. Assume that

$$\int_{[0,1]^{|v|}} \prod_{r=1}^{s} \left|\partial^{\ell_r} \tau_{k_r}(\boldsymbol{u}_v{:}\mathbf{1}_{-v})\right| \mathrm{d}\boldsymbol{u}_v < \infty$$

holds under appropriate set-up. Then $f \circ \tau \in \mathrm{BVHK}$ for all $f \in C^m(\mathcal{X})$.

Sufficient Condition

Corollary 1. B and Owen (2016)

If $\partial^v \tau_j(\boldsymbol{u}_v{:}\mathbf{1}_{-v}) \in L^{p_j}\big([0,1]^{|v|}\big)$ for all j and $v \subseteq 1{:}m$, where $p_j \in [1,\infty]$ and $\sum_{j=1}^{d} 1/p_j \le 1$ then $f \circ \tau \in \mathrm{BVHK}$ for all $f \in C^m(\mathcal{X})$.

Sufficient Condition

Corollary 1. B and Owen (2016)

If $\partial^v \tau_j(\boldsymbol{u}_v{:}\mathbf{1}_{-v}) \in L^{p_j}\big([0,1]^{|v|}\big)$ for all j and $v \subseteq 1{:}m$, where $p_j \in [1,\infty]$ and $\sum_{j=1}^{d} 1/p_j \le 1$ then $f \circ \tau \in \mathrm{BVHK}$ for all $f \in C^m(\mathcal{X})$.

- Proof: Generalized Holder inequality and L^{p_j} conditions establish,

$$\int_{[0,1]^{|v|}} \prod_{r=1}^{s} \big|\partial^{\ell_r} \tau_{k_r}(\boldsymbol{u}_v{:}\mathbf{1}_{-v})\big| \,\mathrm{d}\boldsymbol{u}_v < \infty$$

Main Result for RQMC (Scrambled Net)

Theorem 2. B and Owen (2016)

Let $\tau(\boldsymbol{u})$ be as described. Assume that

$$\int_{[0,1]^d} \prod_{r=1}^{s} \left|\partial^{\ell_r} \tau_{k_r}(\boldsymbol{u})\right|^2 \mathrm{d}\boldsymbol{u} < \infty$$

holds under appropriate set-up. Then $f \circ \tau$ is smooth enough to benefit from randomization.

Sufficient Condition

Corollary 2. B and Owen (2016)

If $\partial^v \tau_j \in L^{p_j}([0,1]^m)$ for all j and $v \subseteq 1{:}m$, where $p_j \in [1,\infty]$. and $\sum_{j=1}^d 1/p_j \le 1/2$, then $f \circ \tau$ is smooth enough to benefit from randomization.

Necessary Conditions

Necessary Conditions

- Much more subtle.

Necessary Conditions

- Much more subtle.
- f can repair any problem by being constant.

Necessary Conditions

- Much more subtle.
- f can repair any problem by being constant.
- τ is unsuitable for QMC when one or more of the components τ_j has $\partial^v \tau_j(\cdot : \mathbf{1}_{-v}) \notin L^1$ for some $v \subset 1{:}m$.
- Thus $\tau_j \notin \mathrm{BVHK}$.

Necessary Conditions

- Much more subtle.
- f can repair any problem by being constant.
- τ is unsuitable for QMC when one or more of the components τ_j has $\partial^v \tau_j(\cdot : \mathbf{1}_{-v}) \notin L^1$ for some $v \subset 1{:}m$.
- Thus $\tau_j \notin \mathrm{BVHK}$.
- Similarly, if $\partial^v \tau_j \notin L^2$ for any j and v, then τ is not a good candidate for RQMC (scrambled nets).

Overview

Map from $[0, 1]^3$ to Equilateral Triangle in 3-dimensions

Map from $[0, 1]^3$ to Equilateral Triangle in 3-dimensions

- Let $T^3 = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1\}$ be an equilateral triangle in 3- dimensions. Consider the map $\tau : [0, 1]^3 \to T^3$ defined by

$$\tau_j(u_1, u_2, u_3) = \frac{\log u_j}{\sum_{i=1}^{3} \log u_i} \qquad j = 1, 2, 3.$$

Map from $[0,1]^3$ to Equilateral Triangle in 3-dimensions

- Let $T^3 = \{(x,y,z) \in \mathbb{R}^3 : x+y+z=1\}$ be an equilateral triangle in 3- dimensions. Consider the map $\tau : [0,1]^3 \to T^3$ defined by

$$\tau_j(u_1,u_2,u_3) = \frac{\log u_j}{\sum_{i=1}^3 \log u_i} \qquad j = 1,2,3.$$

- $$\int_0^1 \int_0^1 \left| \frac{\partial^2 \tau_1}{\partial u_1 \partial u_2} \right|_{u_3=1} du_1 du_2 = \infty$$

Map from $[0,1]^3$ to Equilateral Triangle in 3-dimensions

- Let $T^3 = \{(x,y,z) \in \mathbb{R}^3 : x+y+z=1\}$ be an equilateral triangle in 3- dimensions. Consider the map $\tau : [0,1]^3 \to T^3$ defined by

$$\tau_j(u_1,u_2,u_3) = \frac{\log u_j}{\sum_{i=1}^3 \log u_i} \qquad j = 1,2,3.$$

- $$\int_0^1 \int_0^1 \left| \frac{\partial^2 \tau_1}{\partial u_1 \partial u_2} \right|_{u_3=1} du_1 du_2 = \infty$$

- Thus $\tau \notin \mathrm{BVHK}$.

Map from $[0,1]^d$ to Sphere in d-dimensions via Inverse Gaussian CDF

- The mapping from $[0,1]^d$ to $\mathcal{X} = \mathbb{S}^{d-1}$ is

$$\tau_j(\boldsymbol{u}) = \frac{\Phi^{-1}(u_j)}{\sqrt{\sum_{i=1}^{d} \Phi^{-1}(u_i)^2}}.$$

Map from $[0,1]^d$ to Sphere in d-dimensions via Inverse Gaussian CDF

- The mapping from $[0,1]^d$ to $\mathcal{X} = \mathbb{S}^{d-1}$ is

$$\tau_j(\boldsymbol{u}) = \frac{\Phi^{-1}(u_j)}{\sqrt{\sum_{i=1}^{d} \Phi^{-1}(u_i)^2}}.$$

-

$$\int \left| \partial^{1:d} \tau_j \right| du = \infty$$

Map from $[0,1]^d$ to Sphere in d-dimensions via Inverse Gaussian CDF

- The mapping from $[0,1]^d$ to $\mathcal{X} = \mathbb{S}^{d-1}$ is

$$\tau_j(\boldsymbol{u}) = \frac{\Phi^{-1}(u_j)}{\sqrt{\sum_{i=1}^{d} \Phi^{-1}(u_i)^2}}.$$

-

$$\int \left| \partial^{1:d} \tau_j \right| du = \infty$$

- Thus $\tau \notin \mathrm{BVHK}$.

Non L^2 Mapping

- "Necessary" Condition : $\partial^v \tau_j \in L^2$ to benefit from RQMC.

Non L^2 Mapping

- "Necessary" Condition : $\partial^v \tau_j \in L^2$ to benefit from RQMC.
- Fang and Wang (1994) gave mappings to the following domains

$$A_d = \{(x_1, \ldots, x_d) : 0 \le x_1 \le \ldots \le x_d \le 1\}$$
$$B_d = \{(x_1, \ldots, x_d) : x_1^2 + \ldots + x_d^2 \le 1\}$$
$$U_d = \{(x_1, \ldots, x_d) : x_1^2 + \ldots + x_d^2 = 1\}$$
$$V_d = \{(x_1, \ldots, x_d) \in \mathbb{R}_+^d : x_1 + \ldots + x_d \le 1\}$$
$$T_d = \{(x_1, \ldots, x_d) \in \mathbb{R}_+^d : x_1 + \ldots + x_d = 1\}$$

Non L^2 Mapping

- "Necessary" Condition : $\partial^v \tau_j \in L^2$ to benefit from RQMC.
- Fang and Wang (1994) gave mappings to the following domains

$$
\begin{aligned}
A_d &= \{(x_1, \ldots, x_d) : 0 \le x_1 \le \ldots \le x_d \le 1\} \\
B_d &= \{(x_1, \ldots, x_d) : x_1^2 + \ldots + x_d^2 \le 1\} \\
U_d &= \{(x_1, \ldots, x_d) : x_1^2 + \ldots + x_d^2 = 1\} \\
V_d &= \{(x_1, \ldots, x_d) \in \mathbb{R}_+^d : x_1 + \ldots + x_d \le 1\} \\
T_d &= \{(x_1, \ldots, x_d) \in \mathbb{R}_+^d : x_1 + \ldots + x_d = 1\}
\end{aligned}
$$

- Each $\tau \in BVHK$.

Non L^2 Mapping

- "Necessary" Condition : $\partial^v \tau_j \in L^2$ to benefit from RQMC.
- Fang and Wang (1994) gave mappings to the following domains

$$
\begin{aligned}
A_d &= \{(x_1, \ldots, x_d) : 0 \leq x_1 \leq \ldots \leq x_d \leq 1\} \\
B_d &= \{(x_1, \ldots, x_d) : x_1^2 + \ldots + x_d^2 \leq 1\} \\
U_d &= \{(x_1, \ldots, x_d) : x_1^2 + \ldots + x_d^2 = 1\} \\
V_d &= \{(x_1, \ldots, x_d) \in \mathbb{R}_+^d : x_1 + \ldots + x_d \leq 1\} \\
T_d &= \{(x_1, \ldots, x_d) \in \mathbb{R}_+^d : x_1 + \ldots + x_d = 1\}
\end{aligned}
$$

- Each $\tau \in BVHK$.
- None of them satisfy $\partial^v \tau_j \in L^2$.

Overview

Importance Sampling

- Aim : Estimate $\mu = \int f(x)dP$.

Importance Sampling

- Aim : Estimate $\mu = \int f(x)dP$.
- Assume the measure P has a density p.

Importance Sampling

- Aim : Estimate $\mu = \int f(x) dP$.
- Assume the measure P has a density p.
- Use τ on $[0,1]^m$ which yields $\boldsymbol{x} = \tau(\boldsymbol{u}) \sim q$ on $\mathcal{X}$ when $\boldsymbol{u} \sim \mathbf{U}[0,1]^m$.

Importance Sampling

- Aim : Estimate $\mu = \int f(x)dP$.
- Assume the measure P has a density p.
- Use τ on $[0,1]^m$ which yields $\boldsymbol{x} = \tau(\boldsymbol{u}) \sim q$ on $\mathcal{X}$ when $\boldsymbol{u} \sim \mathbf{U}[0,1]^m$.
- We estimate μ by

$$\hat{\mu}_q^n = \frac{1}{n}\sum_{i=1}^{n} \frac{f(\tau(\boldsymbol{u}_i))p(\tau(\boldsymbol{u}_i))}{q(\tau(\boldsymbol{u}_i))} = \frac{1}{n}\sum_{i=1}^{n} \left(\frac{fp}{q} \circ \tau\right)(\boldsymbol{u}_i).$$

Importance Sampling

- Aim : Estimate $\mu = \int f(x) dP$.
- Assume the measure P has a density p.
- Use τ on $[0,1]^m$ which yields $\boldsymbol{x} = \tau(\boldsymbol{u}) \sim q$ on $\mathcal{X}$ when $\boldsymbol{u} \sim \mathbf{U}[0,1]^m$.
- We estimate μ by

$$\hat{\mu}_q^n = \frac{1}{n} \sum_{i=1}^{n} \frac{f(\tau(\boldsymbol{u}_i)) p(\tau(\boldsymbol{u}_i))}{q(\tau(\boldsymbol{u}_i))} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{fp}{q} \circ \tau \right) (\boldsymbol{u}_i).$$

- If $q(\boldsymbol{x}) > 0$ whenever $f(\boldsymbol{x})p(\boldsymbol{x}) \neq 0$ (and if μ exists) then $\mathbb{E}(\hat{\mu}_q^n) = \mu$.

Importance Sampling

- Aim : Estimate $\mu = \int f(x)dP$.
- Assume the measure P has a density p.
- Use τ on $[0,1]^m$ which yields $\boldsymbol{x} = \tau(\boldsymbol{u}) \sim q$ on $\mathcal{X}$ when $\boldsymbol{u} \sim \mathbf{U}[0,1]^m$.
- We estimate μ by

$$\hat{\mu}_q^n = \frac{1}{n}\sum_{i=1}^{n} \frac{f(\tau(\boldsymbol{u}_i))p(\tau(\boldsymbol{u}_i))}{q(\tau(\boldsymbol{u}_i))} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{fp}{q}\circ\tau\right)(\boldsymbol{u}_i).$$

- If $q(\boldsymbol{x}) > 0$ whenever $f(\boldsymbol{x})p(\boldsymbol{x}) \neq 0$ (and if μ exists) then $\mathbb{E}(\hat{\mu}_q^n) = \mu$.
- To apply the Koksma-Hlawka inequality we only need $(fp/q)\circ\tau \in \mathrm{BVHK}$.

Sufficient Condition for Importance Sampling

Corollary 3. B and Owen (2016)

Under the above setup, assume τ satisfies the conditions of Theorem 1 and that $fp/q \in C^m(\mathcal{X})$. Then, for a low-discrepancy point set $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ in $[0, 1]^m$,

$$\left| \int_{\mathcal{X}} f(\boldsymbol{x}) p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=1}^{n} \left(\frac{fp}{q} \circ \tau \right) (\boldsymbol{u}_i) \right| = O\left(\frac{(\log n)^{m-1}}{n} \right).$$

Sufficient Condition for Importance Sampling

Corollary 3. B and Owen (2016)

Under the above setup, assume τ satisfies the conditions of Theorem 1 and that $fp/q \in C^m(\mathcal{X})$. Then, for a low-discrepancy point set $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ in $[0,1]^m$,

$$\left| \int_{\mathcal{X}} f(\boldsymbol{x}) p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=1}^{n} \left(\frac{fp}{q} \circ \tau \right) (\boldsymbol{u}_i) \right| = O\left(\frac{(\log n)^{m-1}}{n} \right).$$

Proof.

Follows from Theorem 1 and the Koksma-Hlawka inequality. □

Sufficient Condition for Importance Sampling

- The result works when $\mathcal{X}$ is bounded. Especially for spiky integrands on compact sets $\mathcal{X}$.

Sufficient Condition for Importance Sampling

- The result works when $\mathcal{X}$ is bounded. Especially for spiky integrands on compact sets $\mathcal{X}$.
- Note that if $f \in C^m$, then $fp/q \in C^m$ as long as $p/q \in C^m$.

Sufficient Condition for Importance Sampling

- The result works when $\mathcal{X}$ is bounded. Especially for spiky integrands on compact sets $\mathcal{X}$.
- Note that if $f \in C^m$, then $fp/q \in C^m$ as long as $p/q \in C^m$.
- Take $q(\boldsymbol{x}) \propto p(\boldsymbol{x}) \exp(\theta^T \boldsymbol{x})$ for a parameter $\theta \in \mathbb{R}^d$. Then $p/q \in C^m(\mathcal{X})$ when $\mathcal{X}$ is bounded. **[Asmussen and Glynn (2007)]**

Conclusion

- We give sufficient conditions for $V_{HK}(f \circ \tau) < \infty$ as well as well the transformation can benefit from RQMC.
- For most of the common known transformations there is no guarantee of QMC rate. Need constructive proof in almost all spaces and regions.
- For general measures, it might be possible to get QMC rate.

Thank you!

- For this amazing graduation gift!