Construction-free median lattice rules

Takashi Goda

School of Engineering, University of Tokyo

Joint work with Pierre L'Ecuyer (Montréal)

MCQMC 2022, Linz, Austria, July 2022

Problem setting

Approximate the integral

$$I(f) = \int_{[0,1)^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \quad \approx \quad I(f;P_N) = \frac{1}{N}\sum_{n=0}^{N-1} f(\boldsymbol{x}_n)$$

with a point set $P_N = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\} \subset [0,1)^s$.

The quantity on the right is called **quasi-Monte Carlo integration** over P_N.

Lattice rules

Definition (rank-1 lattice rules)

Given $N \geq 2$ and $\boldsymbol{z} \in \{1, \ldots, N-1\}^s$, define $P_N := P_{N,\boldsymbol{z}}$ by

$$\boldsymbol{x}_n = \left\{ \frac{n\boldsymbol{z}}{N} \right\}, \quad n = 0, 1, \ldots, N-1,$$

where $\{\cdot\}$ takes the fractional part of each component. QMC integration over $P_{N,\boldsymbol{z}}$ is called **rank-1 lattice rule** with generating vector $\boldsymbol{z}$.

Lattice rules

Definition (rank-1 lattice rules)

Given $N \geq 2$ and $\boldsymbol{z} \in \{1, \ldots, N-1\}^s$, define $P_N := P_{N,\boldsymbol{z}}$ by

$$\boldsymbol{x}_n = \left\{ \frac{n\boldsymbol{z}}{N} \right\}, \quad n = 0, 1, \ldots, N-1,$$

where $\{\cdot\}$ takes the fractional part of each component. QMC integration over $P_{N,\boldsymbol{z}}$ is called **rank-1 lattice rule** with generating vector $\boldsymbol{z}$.

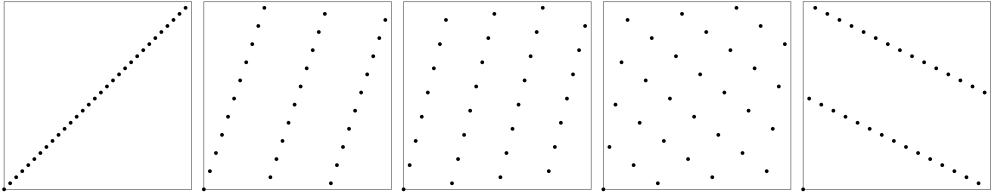


Figure 1: 2-d lattices with $N = 31$, $z_1 = 1$ and $z_2 = 1, 3, 4, 7, 15$.

Periodic functions and dual lattice

Let $f : [0,1)^s \to \mathbb{R}$ be periodic with an absolutely convergent Fourier series

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^s} \hat{f}(\boldsymbol{k}) e^{2\pi i \boldsymbol{k}\cdot\boldsymbol{x}},$$

with $\hat{f}(\boldsymbol{k}) = \int_{[0,1)^s} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{k}\cdot\boldsymbol{x}} \,\mathrm{d}\boldsymbol{x}$ being the $\boldsymbol{k}$-th Fourier coefficient.

Periodic functions and dual lattice

Let $f : [0,1)^s \to \mathbb{R}$ be periodic with an absolutely convergent Fourier series

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^s} \hat{f}(\boldsymbol{k}) e^{2\pi i \boldsymbol{k}\cdot\boldsymbol{x}},$$

with $\hat{f}(\boldsymbol{k}) = \int_{[0,1)^s} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{k}\cdot\boldsymbol{x}} \,\mathrm{d}\boldsymbol{x}$ being the $\boldsymbol{k}$-th Fourier coefficient. The error is

$$\begin{aligned} I(f; P_{N,\boldsymbol{z}}) - I(f) &= \frac{1}{N}\sum_{n=0}^{N-1}\sum_{\boldsymbol{k}\in\mathbb{Z}^s} \hat{f}(\boldsymbol{k}) \exp(2\pi i \boldsymbol{k}\cdot\boldsymbol{x}_n) - \hat{f}(\boldsymbol{0}) \\ &= \sum_{\boldsymbol{k}\in\mathbb{Z}^s\setminus\{\boldsymbol{0}\}} \hat{f}(\boldsymbol{k}) \frac{1}{N}\sum_{n=0}^{N-1} \exp\left(2\pi i \frac{n}{N}\boldsymbol{k}\cdot\boldsymbol{z}\right) \\ &= \sum_{\boldsymbol{k}\in P_{N,\boldsymbol{z}}^{\perp}\setminus\{\boldsymbol{0}\}} \hat{f}(\boldsymbol{k}), \end{aligned}$$

where $P_{N,\boldsymbol{z}}^{\perp} := \{\boldsymbol{k}\in\mathbb{Z}^s \mid \boldsymbol{k}\cdot\boldsymbol{z} \equiv 0 \pmod{N}\}$ denotes the **dual lattice** of $P_{N,\boldsymbol{z}}$.

Korobov spaces

- $\alpha > 1/2$: decay of Fourier coefficients (smoothness).
- $\boldsymbol{\gamma} = \{\gamma_u\}_{u \subseteq \{1,\ldots,s\}}$: non-negative weights [Sloan/Woźniakowski 1998].
- For $\boldsymbol{k} \in \mathbb{Z}^s$, write $\mathrm{supp}(\boldsymbol{k}) = \{1 \le j \le s \mid k_j \ne 0\}$ and define

$$r_{\alpha,\boldsymbol{\gamma}}(\boldsymbol{k}) := \gamma_{\mathrm{supp}(\boldsymbol{k})} \prod_{j \in \mathrm{supp}(\boldsymbol{k})} \frac{1}{|k_j|^\alpha}.$$

Definition

The weighted Korobov space $H^{\mathrm{kor}}_{s,\alpha,\boldsymbol{\gamma}}$ is a RKHS with norm

$$\|f\|^{\mathrm{kor}}_{s,\alpha,\boldsymbol{\gamma}} = \left(\sum_{\boldsymbol{k} \in \mathbb{Z}^s} \frac{|\hat{f}(\boldsymbol{k})|^2}{(r_{\alpha,\boldsymbol{\gamma}}(\boldsymbol{k}))^2} \right)^{1/2}.$$

Koksma-Hlawka type inequality

$$
\begin{aligned}
|I(f;P_{N,\boldsymbol{z}})-I(f)| &= \left|\sum_{\boldsymbol{k}\in P_{N,\boldsymbol{z}}^{\perp}\setminus\{\boldsymbol{0}\}} \hat{f}(\boldsymbol{k})\right| = \left|\sum_{\boldsymbol{k}\in P_{N,\boldsymbol{z}}^{\perp}\setminus\{\boldsymbol{0}\}} \frac{\hat{f}(\boldsymbol{k})}{r_{\alpha,\gamma}(\boldsymbol{k})}\cdot r_{\alpha,\gamma}(\boldsymbol{k})\right| \\
&\le \left(\sum_{\boldsymbol{k}\in P_{N,\boldsymbol{z}}^{\perp}\setminus\{\boldsymbol{0}\}} \frac{|\hat{f}(\boldsymbol{k})|^2}{(r_{\alpha,\gamma}(\boldsymbol{k}))^2}\right)^{1/2} \left(\sum_{\boldsymbol{k}\in P_{N,\boldsymbol{z}}^{\perp}\setminus\{\boldsymbol{0}\}} (r_{\alpha,\gamma}(\boldsymbol{k}))^2\right)^{1/2} \\
&\le \|f\|_{\boldsymbol{s},\alpha,\gamma}^{\mathrm{kor}} \left(\sum_{\boldsymbol{k}\in P_{N,\boldsymbol{z}}^{\perp}\setminus\{\boldsymbol{0}\}} (r_{\alpha,\gamma}(\boldsymbol{k}))^2\right)^{1/2}.
\end{aligned}
$$

Thus it is natural to study how to find good $\boldsymbol{z}$ such that the second factor, denoted by $\mathcal{S}_{\alpha,\gamma}(\boldsymbol{z})$, becomes small.

How to construct good $\boldsymbol{z}$? (quick review)

- Korobov (1959): introduced greedy **component-by-component (CBC)** algorithm; for $1 \leq j \leq s$, one component $z_j \in \{1, \ldots, N-1\}$ is optimized while the earlier components $z_1, \ldots, z_{j-1}$ are fixed.
- Sloan/Reztsov (2002): reintroduced and popularized CBC.
- Kuo (2003): gave an error bound of $\mathcal{O}(N^{-\alpha+\varepsilon})$ for CBC.
- Nuyens/Cools (2006): introduced fast CBC using FFT.

$\vdots$

- Dick/TG (2021): studied stability of CBC (i.e., an error bound for $H^{\mathrm{kor}}_{s,\alpha',\gamma'}$ while $\boldsymbol{z}$ is constructed for $H^{\mathrm{kor}}_{s,\alpha,\gamma}$ with $\alpha \neq \alpha'$ and $\boldsymbol{\gamma} \neq \boldsymbol{\gamma}'$).
- Ebert/Kritzer/Nuyens/Osisiogu (2021): introduced α-free CBC.

Limitation of CBC

To make fast CBC work well, **the weights γ are modeled by a small number of parameters.**

- Product weights: $\gamma_u = \prod_{j \in u} \gamma_j$
- Order-dependent weights: $\gamma_u = \Gamma_{|u|}$
- Product and order-dependent (POD) weights: $\gamma_u = \Gamma_{|u|} \prod_{j \in u} \gamma_j$
- Smoothness-driven POD (SPOD) weights:

$$\gamma_u = \sum_{\boldsymbol{\tau} \in \{1,\ldots,\alpha\}^{|u|}} \Gamma_{\|\boldsymbol{\tau}\|_1} \prod_{j \in u} \gamma_{j,\tau_j}$$

The latter two choices are motivated by PDEs with random coefficients [Kuo/Schwab/Sloan 2012; Dick/Kuo/Le Gia/Nuyens/Schwab, 2014].

Most of the generating vectors are good

For a given application, we may not know what is an appropriate choice for α and γ. Even more, an efficient CBC does not exist for general weights.

However, good news is that **most of the possible $\boldsymbol{z}$'s are good** for any α and $\boldsymbol{\gamma}$ [Dick/Sloan/Wang/Woźniakowski 2006; Kritzer/Kuo/Nuyens/Ullrich 2019]. This fact can be shown by averaging argument and Markov inequality.

Recall that our error criterion is

$$\mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}) = \left(\sum_{\boldsymbol{k} \in P_{N,\boldsymbol{z}}^{\perp} \setminus \{\boldsymbol{0}\}} (r_{\alpha,\gamma}(\boldsymbol{k}))^2 \right)^{1/2}.$$

Let α and $\boldsymbol{\gamma}$ be given and N be a prime. For $1/(2\alpha) < \lambda \leq 1$, we have

$$\frac{1}{(N-1)^s} \sum_{z_1,\ldots,z_s=1}^{N-1} (\mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}))^{2\lambda} \leq \frac{1}{N-1} \sum_{\emptyset \neq u \subseteq \{1,\ldots,s\}} \gamma_u^{2\lambda} (2\zeta(2\alpha\lambda))^{|u|}.$$

This implies that the average of $(\mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}))^{2\lambda}$ over all admissible $\boldsymbol{z}$ is of $\mathcal{O}(N^{-1})$ for any $1/(2\alpha) < \lambda \leq 1$.

Markov inequality states that a proportion of the admissible $\boldsymbol{z}$ which satisfy

$$\mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}) \leq \inf_{1/(2\alpha)<\lambda\leq 1} \left(\frac{1}{\eta(N-1)} \sum_{\emptyset \neq u \subseteq \{1,\ldots,s\}} \gamma_u^{2\lambda} (2\zeta(2\alpha\lambda))^{|u|} \right)^{1/(2\lambda)} =: (*)$$

is greater than or equal to $1-\eta$, for any $0 < \eta < 1$.

Why not consider avoiding bad generating vectors, instead of searching for good ones?

Median helps

Construction-free median lattice rules

Let N be a prime and r be odd.

1. Draw $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r \in \{1, \ldots, N-1\}^s$ uniformly and independently.
2. Approximate $I(f)$ by $\mathrm{median}(I(f; P_{N,\boldsymbol{z}_1}), \ldots, I(f; P_{N,\boldsymbol{z}_r}))$.

Taking the median plays a role in **filtering bad $\boldsymbol{z}$ out adaptively (with very high prob.) without specifying α and γ!**

This median approach for scrambled nets is studied by Pan/Owen (2021).

A probabilistic error guarantee

Given $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r$, the error is

$$\begin{aligned}
&|\mathrm{median}(I(f; P_{N,\boldsymbol{z}_1}), \ldots, I(f; P_{N,\boldsymbol{z}_r})) - I(f)| \\
&\quad \leq \mathrm{median}(|I(f; P_{N,\boldsymbol{z}_1}) - I(f)|, \ldots, |I(f; P_{N,\boldsymbol{z}_r}) - I(f)|) \quad \text{(Jensen)} \\
&\quad \leq \|f\|_{s,\alpha,\gamma}^{\mathrm{kor}} \, \mathrm{median}(\mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}_1), \ldots, \mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}_r)).
\end{aligned}$$

In order to have $\mathrm{median}(\mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}_1), \ldots, \mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}_r)) > (*)$, **at least half of** $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r$ must satisfy $\mathcal{S}_{\alpha,\gamma}(\boldsymbol{z}_j) > (*)$. Thus, under a random choice of $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r$ (with replacement), **such a "failure event" probability is very small!**

A probabilistic error bound [TG/L'Ecuyer, 2022]

Let $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r$ be chosen independently and randomly. **For any α, γ and $f \in H^{\mathrm{kor}}_{s,\alpha,\gamma}$,** we have

$$\left|\mathrm{median}(I(f; P_{N,\boldsymbol{z}_1}), \ldots, I(f; P_{N,\boldsymbol{z}_r})) - I(f)\right|$$

$$\leq \|f\|^{\mathrm{kor}}_{s,\alpha,\gamma} \times \inf_{1/(2\alpha)<\lambda\leq 1} \left(\frac{1}{\eta(N-1)} \sum_{\emptyset \neq u \subseteq \{1,\ldots,s\}} \gamma_u^{2\lambda} (2\zeta(2\alpha\lambda))^{|u|} \right)^{1/(2\lambda)}$$

with a probability of at least

$$1 - (4\eta)^{(r+1)/2}/4$$

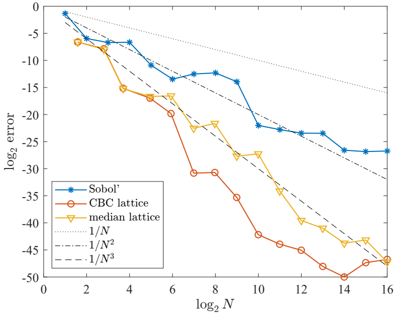
for any $0 < \eta < 1/4$.

Numerical experiments

Smooth, periodic test function

$$f_{\beta,\omega}(\boldsymbol{x}) = \prod_{j=1}^{s} \left[1 + \omega_j \left((2\beta+1) \binom{2\beta}{\beta} x_j^{\beta} (1-x_j)^{\beta} - 1 \right) \right]$$

with $s = 50$ and $\beta = 5$. We set $r = 11$ for our median rule.



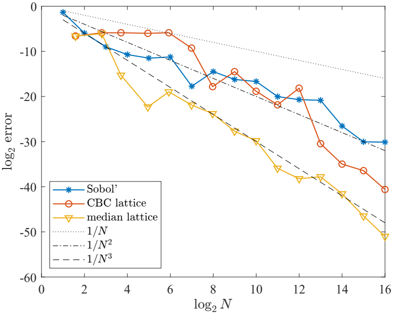


Figure 2: Absolute error for $\omega_j = j^{-6}$ (left) and $\omega_j = (s-j+1)^{-6}$ (right).

Non-periodic functions?

Higher order digital nets constructed by the digit interlacing [Dick 2008; TG/Dick 2015; etc] **require the interlacing factor** (essentially corresponding to smoothness of integrands) **as an input.**

High-order polynomial lattice rules [Niederreiter 1992; Dick/Pillichshammer 2005] do not. The precision n is required instead, which can be as large as machine precision. Therefore they are a good candidate to have an analogous result for non-periodic functions (**a median rule with a prob. error bound without specifying smoothness and weights**).

In fact, for a class of weighted non-periodic function spaces, we could establish the corresponding result, see [TG/L'Ecuyer 2022].

High-order polynomial lattice rules

Definition (High-order polynomial lattice rules)

Let $m, n \geq 2$ with $n \geq m$ be given. For $p \in \mathbb{F}_b[x]$ with $\deg(p) = n$ and $\boldsymbol{q} \in (\mathbb{F}_b[x])^s$ with $\deg(q_j) < n$, define $P_{b^m} := P_{b^m,p,\boldsymbol{q}}$ by

$$\boldsymbol{x}_h = \nu_n\left(\frac{h(x)\boldsymbol{q}(x)}{p(x)}\right), \quad h = 0, 1, \ldots, b^m - 1,$$

where $\nu_n : \mathbb{F}_b((x^{-1})) \to [0, 1)$ is given by

$$\nu_n\left(\sum_{i=w}^{\infty} \frac{a_i}{x^i}\right) = \sum_{i=\max(1,w)}^{n} \frac{a_i}{b^i}$$

and is applied componentwise to a vector.

Examples

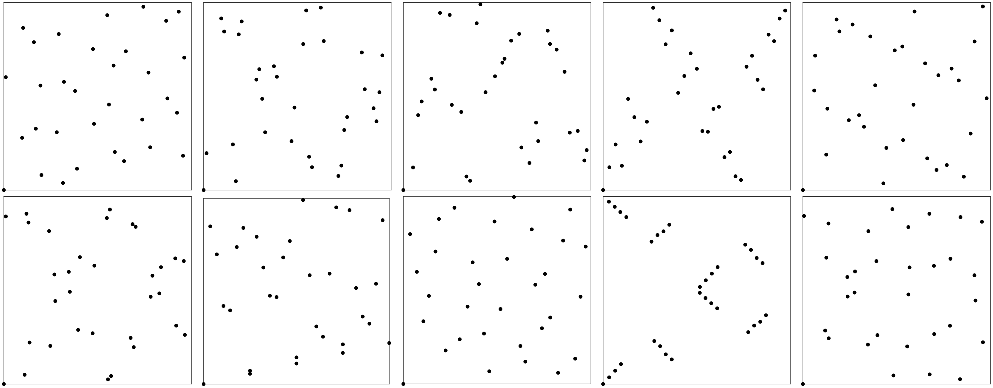


Figure 3: 2-d high-order polynomial lattices with $N = 2^5$ ($b = 2$ and $m = 5$), $n = 52$, $p(x) = x^{52} + x^3 + 1$ (irreducible) and $q_1(x), q_2(x)$ being both random.

Numerical experiments

1-d smooth, non-periodic test function

$$f_1(x) = x^3(1/4 + \log x) \qquad \text{and} \qquad f_2(x) = xe^{x/4}.$$

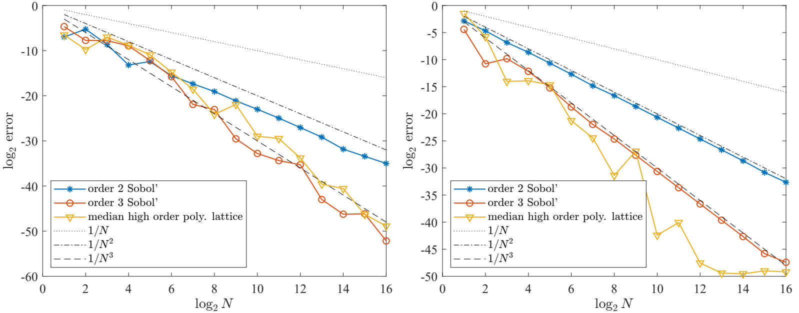


Figure 4: Absolute error for f_1 (left) and f_2 (right).

Future works?

1. What about median of randomly shifted lattice rules?
2. Looking only at the median seems wasteful.

Thank you for your attention!

For more details, please refer to

- TG & P. L'Ecuyer, Construction-free median quasi-Monte Carlo rules for function spaces with unspecified smoothness and general weights, *SIAM Journal on Scientific Computing* (in press). arXiv:2201.09413