A Construction of Polynomial Lattice Rules with Small Gain Coefficients

Jan Baldeaux[1] and Josef Dick[2]

[1]School of Finance and Economics
University of Technology
Sydney, Australia

[2]School of Mathematics and Statistics
University of New South Wales
Sydney, Australia

MCQMC, 19th August 2010

QMC rules

- Quasi-Monte Carlo rules are equal weight integration formulas used to approximate high-dimensional integrals
- can roughly be divided into lattice rules and digital nets
- we focus on digital nets, in particular polynomial lattice rules, which are constructed using polynomials over finite fields

The need for randomization

- approximating integrals, we need information on integration errors
- in some cases, estimates are conservative or unknown
- randomization solves this problem, it allows to obtain statistical information on integration errors
- we study scrambling, a particular randomization method

Introducing the gain coefficients

- we are interested in the variance of estimators of the form

$$\widehat{I}(f) = \frac{1}{b^m} \sum_{h=0}^{b^m-1} f(\boldsymbol{y}_h) \approx \int_{[0,1]^s} f(\boldsymbol{x}) d\boldsymbol{x} = \mathbb{E}\left[\widehat{I}(f)\right]$$

  where $\{\boldsymbol{y}_h\}_{h=0}^{b^m-1}$ is obtained by applying the scrambling algorithm to a polynomial lattice rule
- we have

$$\mathrm{Var}(\widehat{I}(f)) = \frac{1}{N} \sum_{\boldsymbol{l} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} \Gamma_{\boldsymbol{l}} \sigma_{\boldsymbol{l}}^2(f),$$

  for any estimator obtained by scrambling a point set $\{\boldsymbol{x}_h\}_{h=0}^{b^m-1}$ such that $\boldsymbol{x}_h \in [0,1)^s$

A smoothness assumption

- we introduce a norm of the form

$$\|f\|_\alpha = \sup_{\boldsymbol{l} \in \mathbb{N}_0^s} b^{\alpha|\boldsymbol{l}|_1} \sigma_{\boldsymbol{l}}(f)$$

- hence

$$\mathrm{Var}(\widehat{I}(f)) \leq \|f\|_\alpha^2 \frac{1}{N} \sum_{\boldsymbol{l} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} \Gamma_{\boldsymbol{l}} b^{-2\alpha|\boldsymbol{l}|_1}$$

A quality criterion

- employ the quality criterion

$$\frac{1}{N} \sum_{\boldsymbol{l} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} \Gamma_{\boldsymbol{l}} b^{-2\alpha|\boldsymbol{l}|_1}$$

- for digital (t, m, s)-nets, a small t value yields small gain coefficients,

$$\Gamma_{\boldsymbol{l}} = 0 \text{ for } |\boldsymbol{l}|_1 \leq m - t$$

- we minimize the sum for all f for which $\|f\|_\alpha < \infty$ over the class of polynomial lattice rules

Weighted function spaces

- we introduce weights γ, in which case digital (t, m, s)-nets with small t-value do not necessarily yield the smallest possible gain coefficients
- component-by-component constructions have proven useful
- we show they achieve almost optimal convergence rates in the function space under consideration

Outline of this presentation

1. Preliminaries
   - Polynomial lattice rules
   - Scrambling
   - Weighted function spaces based on variance
2. Estimators based on scrambled polynomial lattice rules
3. Component-by-component construction
4. A lower bound

Outline

1. Preliminaries
   - Polynomial lattice rules
   - Scrambling
   - Weighted function spaces based on variance
2. Estimators based on scrambled polynomial lattice rules
3. Component-by-component construction
4. A lower bound

Introducing polynomial lattice rules I

- fix a prime b, denote by $\mathbb{Z}_b$ the finite field containing b elements and by $\mathbb{Z}_b((x^{-1}))$ the field of formal Laurent series

$$L = \sum_{l=w}^{\infty} t_l x^{-l},$$

where w is an arbitrary integer and all $t_l \in \mathbb{Z}_b$

- introduce a map from $\mathbb{Z}_b((x^{-1}))$ to $[0, 1)$

$$v_m\left(\sum_{l=w}^{\infty} t_l x^{-l}\right) = \sum_{l=\max(1,w)}^{m} t_l b^{-l}$$

Introducing polynomial lattice rules II

Definition

Choose $p(x) \in \mathbb{Z}_b[x]$ with $\deg(p(x)) = m$, $q_1(x), \ldots, q_s(x) \in \mathbb{Z}_b[x]$. For $0 \le h < b^m$ let $h = h_0 + h_1 b + \ldots + h_{m-1} b^{m-1}$ and let

$$\overline{h}(x) = \sum_{r=0}^{m-1} h_r x^r \in \mathbb{Z}_b[x].$$

Then $S_{p,m}(\boldsymbol{q})$, where $\boldsymbol{q} = (q_1, \ldots, q_s)$, is the point set

$$\boldsymbol{x}_h = \left(v_m \left(\frac{\overline{h}(x) q_1(x)}{p(x)} \right), \ldots, v_m \left(\frac{\overline{h}(x) q_s(x)}{p(x)} \right) \right) \in [0,1)^s,$$

for $0 \le h < b^m$. A quasi-Monte Carlo rule using the point set $S_{p,m}(\boldsymbol{q})$ is called a polynomial lattice rule.

Dual lattice

- for a non-negative integer k with b-adic expansion $k = k_0 + k_1 b + \ldots,$

$$tr_m(k)(x) = k_0 + k_1 x + \ldots + k_{m-1} x^{m-1} \in \mathbb{Z}_b[x]$$

Definition

Let b be prime and $\boldsymbol{q}(x) = (q_1(x), \ldots, q_s(x)) \in \mathbb{Z}_b^s[x]$, then the dual polynomial lattice of $S_{p,m}(\boldsymbol{q})$ is given by

$$\mathcal{D} = \mathcal{D}_p(\boldsymbol{q}) = \{\boldsymbol{k} \in \mathbb{N}_0^s : tr_m(k_1)(x) q_1(x) + tr_m(k_2)(x) q_2(x) + \cdots + tr_m(k_s)(x) q_s(x) \equiv 0 \pmod{p(x)}\} .$$

Introducing scrambling I

- given $\boldsymbol{x} \in [0,1)^s$, where $\boldsymbol{x} = (x_1, \ldots, x_s)$ and

$$x_j = \frac{\xi_{j,1}}{b} + \frac{\xi_{j,2}}{b^2} + \ldots.$$

- then the scrambled point shall be denoted by $\boldsymbol{y} \in [0,1)^s$, where $\boldsymbol{y} = (y_1, \ldots, y_s)$ and

$$y_j = \frac{\eta_{j,1}}{b} + \frac{\eta_{j,2}}{b^2} + \ldots.$$

- the permutation applied to $\xi_{j,l}$, $j = 1, \ldots, s$ depends on $\xi_{j,k}$ for $1 \leq k < l$, we have

$$\eta_{j,1} = \pi_j(\xi_{j,1}), \; \eta_{j,2} = \pi_{j,\xi_{j,1}}(\xi_{j,2}), \; \eta_{j,3} = \pi_{j,\xi_{j,1},\xi_{j,2}}(\xi_{j,3})$$
$$\eta_{j,k} = \pi_{j,\xi_{j,1},\ldots,\xi_{j,k-1}}(\xi_{j,k}), \; k \geq 2.$$

Introducing scrambling II

- apply scrambling to a point $\boldsymbol{x}$ to obtain $\boldsymbol{y}$, $\boldsymbol{y}$ is uniformly distributed in $[0, 1)^s$
- scrambling preservers the (t, m, s)-net property with probability 1

An expression for the variance

- we study the estimator

$$\widehat{I}(f) = \frac{1}{N} \sum_{h=0}^{N-1} f(\boldsymbol{y}_h),$$

where $\{\boldsymbol{y}_h\}_{h=0}^{N-1}$ is obtained by applying the scrambling algorithm to $\{\boldsymbol{x}_h\}_{h=0}^{N-1}$, $\boldsymbol{x} \in [0,1)^s$,

$$\mathrm{Var}(\widehat{I}(f)) = \sum_{\boldsymbol{l} \in \mathbb{N}^s \setminus \{\boldsymbol{0}\}} \Gamma_{\boldsymbol{l}} \sigma_{\boldsymbol{l}}^2(f)$$

The gain coefficients

- the $\sigma_{\boldsymbol{l}}(f)$ only depend on f and can be expressed in terms of Walsh or Haar coefficients
- we define a weighted norm for functions $f \in L_2([0,1]^s)$ by

$$\|f\|_\alpha = \max_{\mathfrak{u} \subseteq [s]} \gamma_{\mathfrak{u}}^{-1/2} \sup_{\boldsymbol{l}_{\mathfrak{u}} \in \mathbb{N}^{|\mathfrak{u}|}} b^{\alpha |\boldsymbol{l}_{\mathfrak{u}}|_1} \sigma_{(\boldsymbol{l}_{\mathfrak{u}}, \boldsymbol{0})}(f).$$

- for $0 < \alpha \leq 1$, $V_{\alpha,s,\gamma} \subseteq L_2([0,1]^s)$ consists of all functions f for which $\|f\|_\alpha < \infty$

What functions lie in this space? I

- for a subinterval $J = \prod_{j=1}^{s} [x_j, y_j)$ with $0 \leq x_j < y_j \leq 1$ and $f : [0,1)^s \to \mathbb{R}$, let $\Delta(j, J)$ denote the alternating sum of f at the vertices of J, adjacent vertices having opposite signs
- the generalized variation in the sense of Vitali of order $0 < \alpha \leq 1$ is

$$V_\alpha^{(s)}(f) = \sup_{\mathcal{P}} \left(\sum_{J \in \mathcal{P}} \mathrm{Vol}(J) \left| \frac{\Delta(f, j)}{\mathrm{Vol}(J)^\alpha} \right|^2 \right)^{1/2}$$

where the supremum is extended over all partitions $\mathcal{P}$ of $[0,1]^s$

What functions lie in this space? II

- for $\alpha = 1$, and f having continuous partial derivatives,

$$V_1^{(s)}(f) = \left(\int_{[0,1]^s} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s} \right|^2 \mathrm{d}\boldsymbol{x} \right)^{1/2} .$$

- taking into account projections onto lower-dimensional faces, we obtain the generalized Vitali variation with coefficient α

$$V_\alpha(f) = \left(\sum_{\mathfrak{u} \subseteq [s]} \left(V_\alpha^{(|\mathfrak{u}|)}(f_{\mathfrak{u}}; \mathfrak{u}) \right)^2 \right)^{1/2}$$

What functions lie in this space? III

Corollary

Let $b \geq 2$ be a natural number and let $f \in L_2([0,1]^s)$ have bounded variation $V_\alpha(f) < \infty$ of order $0 < \alpha \leq 1$. Then

$$\|f\|_\alpha \leq \max\left(\|f\|_{L_2} \gamma_\emptyset^{-1}, V_\alpha(f) \max_{\emptyset \neq \mathfrak{u} \subseteq [s]} \gamma_\mathfrak{u}^{-1/2} (b-1)^{(\alpha - 1/2)_+ |\mathfrak{u}|} \right).$$

Outline

1. Preliminaries
   - Polynomial lattice rules
   - Scrambling
   - Weighted function spaces based on variance
2. Estimators based on scrambled polynomial lattice rules
3. Component-by-component construction
4. A lower bound

The estimators

- we discuss the variance of the estimator

$$\widehat{I}(f) = \frac{1}{b^m} \sum_{h=0}^{b^m-1} f(\boldsymbol{y}_h),$$

where $\{\boldsymbol{y}_h\}_{h=0}^{b^m-1}$ are obtained by applying the scrambling algorithm to the polynomial lattice rule $\{\boldsymbol{x}_h\}$

Worst-case variance I

- we are interested in the worst-case variance of multivariate integration in $V_{\alpha,s,\gamma}$ using a scrambled quasi-Monte Carlo rule $Q_{b^m,s}$:

$$\mathrm{Var}\left(Q_{b^m,s}, V_{\alpha,s,\gamma}\right) = \sup_{f \in V_{\alpha,s,\gamma}, \|f\|_\alpha \leq 1} \mathrm{Var}\left[\widehat{I}(f, Q_{b^m,s})\right] ,$$

where $\widehat{I}(f, Q_{b^m,s})$ denotes the estimator obtained by scrambling $Q_{b^m,s}$

Worst-case variance II

- the quasi-Monte Carlo rule associated with $S_{p,m}(\boldsymbol{q})$ is denoted by $Q_{b^m,s}(\boldsymbol{q})$:

$$\mathrm{Var}\left(Q_{b^m,s}(\boldsymbol{q}), V_{\alpha,s,\gamma}\right) = \sup_{f \in V_{\alpha,s,\gamma}, \|f\|_\alpha \le 1} \mathrm{Var}\left[\widehat{I}(f, Q_{b^m,s})\right],$$

- for $k = \kappa_0 + \kappa_1 b + \ldots + \kappa_{a-1} b^{a-1} \in \mathbb{N}_0$ let

$$r_{\alpha,\gamma}(k) = \begin{cases} 1 & \text{if } k = 0, \\ \gamma \frac{b}{(b-1)b^{\alpha a}} & \text{if } k > 0, \end{cases}$$

and for $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$ let $r_{\alpha,\boldsymbol{\gamma}}(\boldsymbol{k}) = \prod_{j=1}^s r_{\alpha,\gamma_j}(k_j)$

A bound on the worst-case variance

Corollary

Let $0 < \alpha \leq 1$, $\boldsymbol{q} \in \mathbb{Z}_b[x]^s$ *be a generating vector for a classical polynomial lattice point set with modulus p, and* $\mathrm{Var}(Q_{b^m,s}(\boldsymbol{q}), V_{\alpha,s,\gamma})$ *the worst-case variance associated with* $Q_{b^m}(\boldsymbol{q})$. *Then*

$$\mathrm{Var}(Q_{b^m,s}(\boldsymbol{q}), V_{\alpha,s,\gamma}) \leq \sum_{\boldsymbol{k} \in \mathcal{D}_p'(\boldsymbol{q})} r_{2\alpha+1,\gamma}(\boldsymbol{k}),$$

where $\mathcal{D}_p'(\boldsymbol{q}) = \mathcal{D}_p(\boldsymbol{q}) \setminus \{\mathbf{0}\}$ *and* $\mathcal{D}_p(\boldsymbol{q})$ *is the dual polynomial lattice.*

A quality criterion

- the bound is denoted by

$$B(\boldsymbol{q}, \alpha, \gamma) := \sum_{\boldsymbol{k} \in \mathcal{D}_p'(\boldsymbol{q})} r_{2\alpha+1,\gamma}(\boldsymbol{k}) \qquad (1)$$

Theorem

The following equality holds:

$$B(\boldsymbol{q}, \alpha, \gamma) = \frac{1}{b^m} \sum_{h=0}^{b^m-1} \prod_{j=1}^{s} \left(1 + \frac{b}{b-1} \gamma_j \phi(\boldsymbol{x}_{h,j}, \alpha)\right) - 1 \,,$$

$$\phi(x, \alpha) = \frac{b - 1 - b^{2\alpha \lfloor \log_b x \rfloor}(b^{2\alpha+1} - 1)}{b(b^{2\alpha} - 1)} \,.$$

Self-adjustment property

- assume $S_{p,m}(\boldsymbol{q})$ is a polynomial lattice rule so that

$$B(\boldsymbol{q}, \alpha, \gamma) \leq C_{s,\alpha,\gamma} N^{-(1+2\alpha)+\delta},$$

- then from Jensen's inequality

$$B(\boldsymbol{q}, \alpha, \gamma)^{\frac{1+2\alpha'}{1+2\alpha}} \geq B(\boldsymbol{q}, \alpha', \gamma^{\frac{1+2\alpha'}{1+2\alpha}}),$$

for $\alpha \leq \alpha' \leq 1$.

- Hence

$$B(\boldsymbol{q}, \alpha', \gamma^{\frac{1+2\alpha'}{1+2\ \text{alpha}}}) \leq C_{s,\alpha,\gamma}^{\frac{1+2\alpha'}{1+2\ \text{alpha}}} N^{-(1+2\alpha')+\delta\frac{1+2\alpha'}{1+2\ \text{alpha}}},$$

so the polynomial lattice rule constructed to achieve optimal convergence rates for functions of smoothness α, adjusts itself to the optimal rate.

Outline

1. Preliminaries
   - Polynomial lattice rules
   - Scrambling
   - Weighted function spaces based on variance
2. Estimators based on scrambled polynomial lattice rules
3. Component-by-component construction
4. A lower bound

The CBC algorithm

$$R_{b,m} := \{q \in \mathbb{Z}_b[x] : deg(q) < m \text{ and } q \neq 0\} \ .$$

Algorithm 1 CBC algorithm

Require: b a prime, $s, m \in \mathbb{N}$ and weights $\gamma = (\gamma_j)_{j \geq 1}$.

1: Choose an irreducible polynomial $p \in \mathbb{Z}_b[x]$ with $\deg(p) = m$.
2: Set $q_1 = 1$.
3: **for** $d = 2$ to s **do**
4: find $q_d \in R_{b,m}$ by minimizing $B((q_1, \ldots, q_d), \alpha, \gamma)$ as a function of q_d.
5: **end for**
6: **return** $\boldsymbol{q} = (q_1, \ldots, q_s)$.

The convergence rate of the CBC algorithm

Theorem

Let b be prime and suppose that $\boldsymbol{q}^*$ is constructed using the CBC algorithm. Then

$$B(\boldsymbol{q}_s^*, \alpha, \gamma) \leq c_{s,\alpha,\gamma,\delta} N^{-(2\alpha+1)+\delta}, \; 0 < \delta \leq 2\alpha .$$

If $\sum_{j=1}^{\infty} \gamma_j^{\frac{1}{2\alpha+1-\delta}} < \infty$, then

$$B(\boldsymbol{q}_s^*, \alpha, \gamma) \leq c_{\infty,\alpha,\gamma,\delta} N^{-(2\alpha+1)+\delta}, \; 0 < \delta \leq 2\alpha .$$

Outline

1. Preliminaries
   - Polynomial lattice rules
   - Scrambling
   - Weighted function spaces based on variance
2. Estimators based on scrambled polynomial lattice rules
3. Component-by-component construction
4. A lower bound

Introducing a class of approximations I

- establish a lower bound for a large class of algorithms, following [Novak '88]
- consider approximating

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

a mapping $I : V_{\alpha,s,\gamma} \to \mathbb{R}$ using $\widetilde{I} : V_{\alpha,s,\gamma} \to \mathbb{R}$

Introducing a class of approximations II

- we consider approximations of the form

$$\widetilde{I} = \varphi \circ L$$

where

- $L : V_{\alpha,s,\gamma} \to \mathbb{R}^N$ represents information
- $\varphi : \mathbb{R}^N \to \mathbb{R}$ is the algorithm showing how to use the information

Defining a class of approximations

- recall that our approximations have the form $\tilde{I} = \varphi \circ L$
- define the following information operator

$$I^{ad} = \{L : V_{\alpha,s,\gamma} \to \mathbb{R}^N | L(f) = (f(\boldsymbol{a}_1), \ldots, f(\boldsymbol{a}_N(f(\boldsymbol{a}_1), \ldots, f(\boldsymbol{a}_{N-1})))), \text{ where } \boldsymbol{a}_1 \in [0,1]^s \text{ and } \boldsymbol{a}_i : \mathbb{R}^{i-1} \to [0,1]^s \text{ for } i = 2, \ldots, s\}$$

- introduce the class of approximations

$$A_N^{ad} = \left\{\tilde{I} : V_{\alpha,s,\gamma} \to \mathbb{R} | \tilde{I} = \varphi \circ L \text{ with } \varphi : \mathbb{R}^N \to \mathbb{R} \text{ and } L \in I_N^{ad}\right\}$$

Randomized algorithms

- recall the class of approximations

$$A_N^{ad} = \left\{ \tilde{S} : V_{\alpha,s,\gamma} \to \mathbb{R} | \tilde{S} = \phi \circ M \text{ with } \phi : \mathbb{R}^N \to \mathbb{R} \text{ and } M \in I_N^{ad} \right\}$$

- $Q = (Q(\omega))_{\omega \in \Omega}$ is a randomized algorithm in A_N^{ad} if (Ω, B, μ) is a probability space and $Q(\omega) \in A_N^{ad}$ for all $\omega \in \Omega$
- the set of all randomized algorithms is denoted by $C(A_N^{ad})$

A lower bound

Theorem

We have the following lower bound

$$\inf_{Q \in C(A_N^{ad})} \sup_{\substack{f \in V_{\alpha,s,\gamma} \\ \|f\|_\alpha \le 1}} \mathrm{Var}(Q(f)) \geq \tilde{C} N^{-(2\alpha+1)},$$

for some constant $\tilde{C}$ independent of N, where

$$\mathrm{Var}(Q(f)) = \int_\Omega \left[Q(\omega)(f) - \int_\Omega Q(\omega')(f) d\mu(\omega') \right]^2 d\mu(\omega).$$

Related works

- same convergence rates established for functions in $V_{\alpha,s,\gamma}$, $0 \leq \alpha \leq 1$, using scrambled digital nets, see the book by Dick and Pillichshammer
- optimal root mean square error convergence rates for $\alpha \geq 2$, for functions of higher order variation, recently obtained by Dick