Monte Carlo methods for numerical integration

Mario Ullrich

Johannes Kepler University Linz

NPFSA Bedlewo, September 2017

The Problem

Let F_d be a class of d-variate functions $f : [0,1]^d \to \mathbb{R}$.

For $f \in F_d$ let

$$S(f) := \int_{[0,1]^d} f(y)\,\mathrm{d}y$$

$\rightsquigarrow S : F_d \to \mathbb{R}$ is called the *solution operator*.

The goal is to compute an approximation of $S(f)$ for $f \in F_d$ using only some *information* of the input f; here function evaluations.

Sobolev spaces

For $1 \leq p \leq \infty$ and $s \in \mathbb{N}$, we define the *Sobolev classes of (dominating) mixed smoothness*

$$\mathbf{W}_p^s = \left\{ f \in L_p([0,1]^d) \colon \|f\|_{\mathbf{W}_p^s} \leq 1 \right\},$$

where

$$\|f\|_{\mathbf{W}_p^s} := \left(\sum_{\alpha \in \mathbb{N}_0^d \colon |\alpha|_\infty \leq s} \|D^\alpha f\|_p^p \right)^{1/p}$$

For simplicity we sometimes consider $\mathring{\mathbf{W}}_p^s = \{ f \in \mathbf{W}_p^s \colon \operatorname{supp}(f) \subset (0,1)^d \}$.

Sobolev spaces

For practical applications other classes are often more suitable!

E.g. numerical simulations show an order of convergence n^{-2} for integration of $f(x) = |x - 1/2|$ $(d = 1)$. This can be explained by

$$f \in B^2_{1,\infty} \hookrightarrow E_2 := \{f : k^2 \hat{f}(k) < \infty\}.$$

Deterministic algorithms

The general form of a **deterministic** algorithm:

$$A_n(f) = \varphi_n\Big(f(x_1), \ldots, f(x_n)\Big)$$

with a (nonlinear) mapping $\varphi_n : \mathbb{R}^n \to \mathbb{R}$ and (adaptively chosen) sample points x_i.

We want to bound

$$e(A_n, F_d) \;=\; \sup_{f \in F_d} \big|S(f) - A_n(f)\big|$$

or even

$$e_n(F_d) \;=\; \inf_{A_n} \, e(A_n, F_d).$$

Remark (Bakhvalov and Smolyak)

(S is linear)

If F_d is symmetric and convex, then we may restrict ourselves to linear, non-adaptive algorithms (cubature rules) of the form

$$Q_n(f) = \sum_{x \in \mathcal{P}_n} a_x \, f(x),$$

where $(a_x)_{x \in \mathcal{P}_n} \subset \mathbb{R}$ and $\mathcal{P}_n \subset [0,1]^d$ with $\#\mathcal{P}_n = n$.

(If $a_x = 1/n$, we say Q_n is a quasi-Monte Carlo (QMC) algorithm.)

Randomized algorithms (a.k.a. Monte Carlo)

For a (possibly) random point set $\mathcal{U}_n$, let

$$M_n(f) := \sum_{x \in \mathcal{U}_n} c_x \, f(x)$$

with (possibly) random weights $c_x = c_x(\mathcal{U}_n)$.

Define

$$e^{\mathrm{ran}}(M_n, F_d) = \sup_{f \in F_d} \sqrt{\mathbb{E}\big|S(f) - M_n(f)\big|^2}$$

and analogously $e_n^{\mathrm{ran}}(F_d)$.

Randomized algorithms

The use of randomized algorithms in applications has many advantages (and a few disadvantages), e.g.,

$$e_n^{\mathrm{ran}}(F_d) \le \frac{2}{\sqrt{n}}$$

as long as $F_d \subset \{f \in L_2([0,1]^d) \colon \|f\|_2 \le 1\}$.
This is achieved by the classical Monte Carlo method.

Moreover,

$$e_n^{\mathrm{ran}}(F_d) \le e_n(F_d)$$

since every deterministic algorithm is also a random algorithm.

Randomized algorithms

Hence, we know that there exist randomized algorithms that

- provide dimension independent error bounds
- have higher order convergence for smooth functions

Randomized algorithms

Hence, we know that there exist randomized algorithms that

- provide dimension independent error bounds
- have higher order convergence for smooth functions

Is there an algorithm M_n that achieves both?

One could ask, e.g., for an M_n with

$$e_n^{\mathrm{ran}}(M_n, F_d) \leq 100 \cdot \min\left\{n^{-1/2}, e_n(F_d)\right\}.$$

Aim

Instead of considering this question, I have...

1. proven the optimal order for random methods for $\mathbf{W}_p^s$
   (with an algorithm that is useless in high dimension).
2. considered a promising algorithm for **very** high dimensions
   (that is certainly not optimal).

Deterministic algorithms: General lattice rules

Let T be an invertible $d \times d$-matrix with $\det(T) = 1$ and let

$$\mathbb{X} := T(\mathbb{Z}^d) \qquad \text{and} \qquad \mathcal{L}_n := c_n^{1/d}\mathbb{X} \cap [0,1]^d$$

with $n \in \mathbb{N}$ and $c_n > 0$ such that $\#\mathcal{L}_n = n$. Clearly, $c_n \asymp 1/n$.

The cubature rule is then defined by

$$Q_n(f) = c_n \sum_{x \in \mathcal{L}_n} f(x).$$

Frolov's construction from 1976

It remains to construct a "good" generator T for our cubature rule. For this define the (irreducible) polynomial

$$p_d(t) = \prod_{j=1}^{d} (t - 2j + 1) - 1, \qquad t \in \mathbb{R},$$

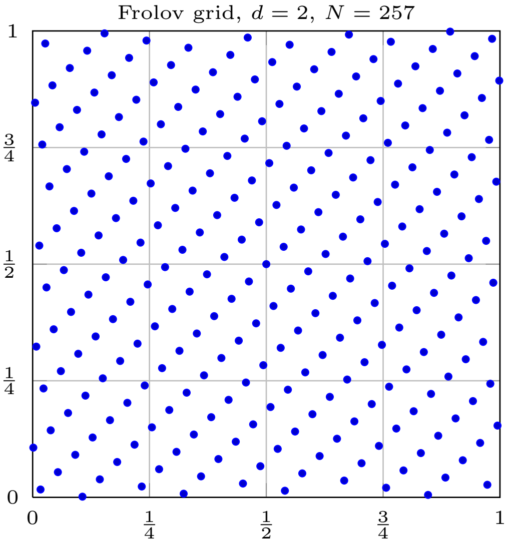
and let $\xi_1, \ldots, \xi_d \in \mathbb{R}$ be its roots. Now, define the invertible matrix

$$T' = \begin{pmatrix} 1 & \xi_1 & \xi_1^2 & \cdots & \xi_1^{d-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \xi_d & \xi_d^2 & \cdots & \xi_d^{d-1} \end{pmatrix}$$

and let $T := T'/\det(T')$. Infact, every T with

$$\inf_{m \in \mathbb{Z}^d \setminus \{0\}} \prod_{j=1}^{d} (Tm)_j > 0 \qquad \text{would work.}$$

Frolov's construction from 1976



Results (optimal deterministic order)

Theorem [Frolov '76 / Skriganov '94 / Bykovski '85 / Temlyakov '90]

Let Q_n be the Frolov cubature rule as defined above. Then, for each $1 < p < \infty$ and $s > \max\{1/2, 1/p\}$, we have

$$e(Q_n, \mathring{\mathbf{W}}_p^s) \asymp e_n(\mathbf{W}_p^s) \asymp n^{-s} (\log n)^{(d-1)/2}.$$

There are numerous more known results in different settings of, e.g., Bakhvalov, Dubinin, Dũng, Frolov, Hinrichs, Korobov, Markhasin, Skriganov, Temlyakov, Triebel, T. Ullrich, myself...

$\rightsquigarrow$ Frolov's construction is **universal** for $\left\{\mathring{\mathbf{W}}_p^s \colon\ s \geq 1,\ p \in (1, \infty)\right\}$.

(Large d? Implementation?)

Monte Carlo

We consider the randomized algorithm of Krieg & Novak

$$M_n(f) = c(n, U, V) \sum_{x \in \mathcal{P}_{n,U,V}} f(x)$$

$$\mathcal{P}_{n,U,V} \ := \ c^{1/d}\, U\, T(\mathbb{Z}^d + V) \cap [0,1]^d$$

with $c = c(n, U, V)$ such that $\#\mathcal{P}_{n,U,V} = n$.
Here, $V \sim \mathcal{U}[0,1]^d$ and $U = \mathrm{diag}(u_1, \ldots, u_d)$ with $u_i \sim \mathcal{U}[1,2]$.

Monte Carlo

Theorem (MU '17)

Let M_n be the randomized Frolov cubature rule as defined above.

Then, for each $s \in \mathbb{N}$, $1 \le p \le \infty$, we have

$$e^{\mathrm{ran}}(M_n, \mathring{\mathbf{W}}_p^s) \asymp e_n^{\mathrm{ran}}(\mathbf{W}_p^s) \asymp n^{-s-1/2+\sigma_p}$$

with $\sigma_p := \min\{0, 1/2 - 1/p\}$.

- universal for $\left\{\mathring{\mathbf{W}}_p^s \colon\ s \in \mathbb{N},\ p \in [1,\infty]\right\}$
- The constant in front is $\sim d^d$ (Geometry of numbers)

(Implementation?)

Online archive for point sets

In collaboration with C. Kacwin, J. Oettershagen and T. Ullrich we're designing the homepage

http://wissrech.ins.uni-bonn.de/research/software/frolov/

where one can download the point sets (presently for $d \leq 10$, $n \leq 2^{20}$) together with programs to use them.

Optimally weighted MC

WORK IN PROGRESS

As discussed above, the random method based on Frolov's cubature is hard (or impossible) to implement in very high dimension.

Instead, we use **i.i.d. random points** and assign "optimal" weights.

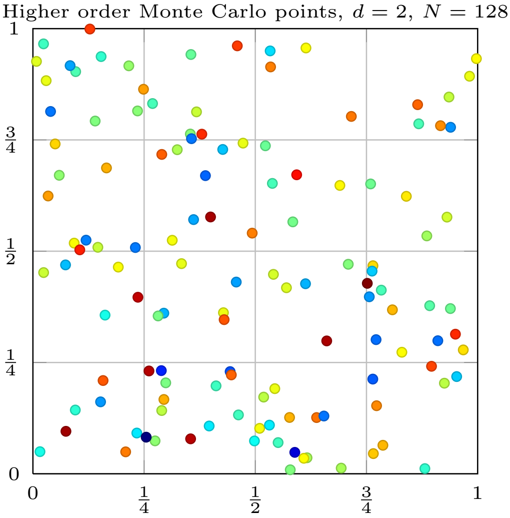
For such a random set $\mathcal{U}_n$, we arrange the weights $c_x = c_x(\mathcal{U}_n)$ in

$$M_n(f) := \sum_{x \in \mathcal{U}_n} c_x \, f(x)$$

such that certain (basis) functions are **integrated exactly**.

Optimally weighted MC

WORK IN PROGRESS



Optimally weighted MC

WORK IN PROGRESS

Together with A. Hinrichs and J. Oettershagen, we proved

Theorem

Let M_n be defined as above. Then, for each $s \in \mathbb{N}$, we have

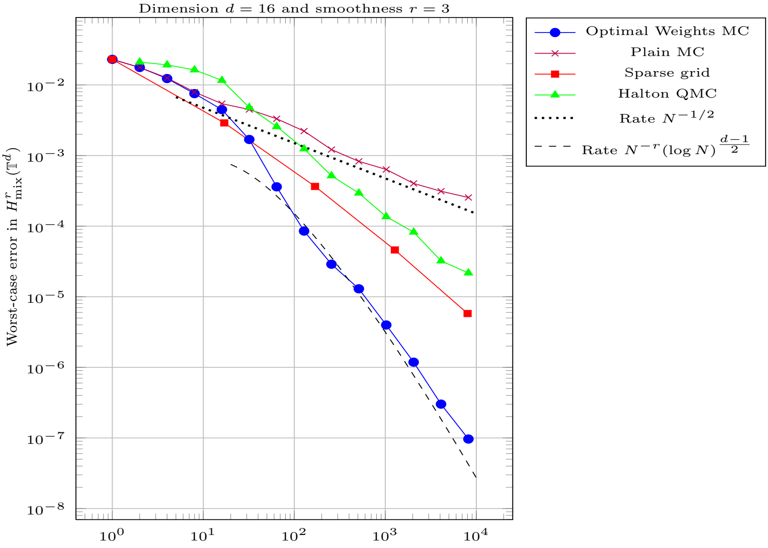
$$e(M_n, \mathring{\mathbf{W}}_2^s) \lesssim n^{-s}(\log n)^{c(s,d)}$$

with probability $\geq 1 - n^{-K}$ for all $K > 0$.

(based on ideas of Cohen/Davenport/Leviatan '13)

Optimally weighted MC

WORK IN PROGRESS



Online archive for point sets

In collaboration with C. Kacwin, J. Oettershagen and T. Ullrich we're designing the homepage

`http://wissrech.ins.uni-bonn.de/research/software/frolov/`

where one can download the point sets (presently for $d \leq 10$, $n \leq 2^{20}$) together with programs to use them.

Thank you!