Tractability of Multivariate Problems

Henryk Woźniakowski

Columbia University and University of Warsaw

Book

The talk is based on the book

Tractability of Multivariate Problems

- **Volume I: Linear Information**
- **Volume II: Function Values**

Erich Novak and Henryk Woźniakowski
EMS Tracts in Mathematics

A little tractability history

We are in early 1990.... **Finance CMO Problem**

$$I_d(f) = \int_{[0,1]^d} f(t)\,\mathrm{d}t$$

Then $d = 360$ by Spaskov. Today $d = 9125$ by Kuo and Waterhouse.

$$I_d(f) \approx \mathrm{QMC}_n(f) = \frac{1}{n}\sum_{j=1}^{n} f(t_j)$$

$t_1, t_2, \dots, t_n$ chosen as

Faure, Niederreiter, Sobol, Tezuka,... points

Empirical Results

Many reports that even for relatively small n, $n \approx 1000$,

$$I_d(f) - \mathrm{QMC}_n(f) \approx \frac{1}{n}$$

Usually similar results for QMC_n as for MC_{n^2}

Why?

Challenge to Theory

Roth [54,80], Frolov [80]: for $f \in H(K_d)$ with

$$K_d(x,y) = \prod_{j=1}^{d} (1 + \min(x_j, y_j))$$

we have

$$I_d(f) - \mathrm{QMC}_n(f) = \mathcal{O}\left(\frac{\ln^{(d-1)/2} n}{n} \, \|f\|_{H(K_d)} \right)$$

Ok for small d. But for $d = 360$

$$\frac{\ln^{(d-1)/2} n}{n} \quad \text{is increasing for} \quad n \leq e^{(d-1)/2}$$

$$d = 360 \quad \text{implies} \quad e^{(d-1)/2} \approx 9 \cdot 10^{77}$$

Challenge to Theory

$$I_d(f) \approx A_{n,d}(f) = \sum_{j=1}^{n} a_j f(t_j)$$

$$e(n,d) = \inf_{A_{n,d}} \sup_{\|f\|_{H(K_d)} \le 1} |I_d(f) - A_{n,d}(f)|$$

Information Complexity

$$n(\varepsilon,d) = \min\{\, n \;:\; e(n,d) \le \varepsilon \, \|I_d\| \,\}$$

Basic questions:

- **How does $n(\varepsilon,d)$ depend on ε^{-1} and d?**
- **Do we have the curse of dimensionality, i.e., an exponential dependence on d?**

Curse of Dimensionality

$$n(\varepsilon,d) \quad = \quad \Theta\left(\varepsilon^{-1}\,\ln^{(d-1)/2}\varepsilon^{-1}\right) \quad \textbf{as} \quad \varepsilon \to 0$$

$$n(\varepsilon,d) \quad \le \quad \left\lceil \frac{9-4\sqrt{2}}{3}\,\frac{1}{\varepsilon^2} \right\rceil = [1.1143\dots]^{\,d}\varepsilon^{-2}$$

$$n(\varepsilon,d) \quad \ge \quad [1.0202]^{\,d}\,(1-\varepsilon^2)$$

For QMC algorithms

$$n(\varepsilon,d) \quad \le \quad [1.125]^{\,d}\,\varepsilon^{-2}$$

$$n(\varepsilon,d) \quad \ge \quad [1.055]^{\,d}(1-\varepsilon^2) \qquad [1.055]^{360} \ge 2\cdot 10^8$$

1. Roth[54,80], Frolov[80], Chen[85] 2. Plaskota, Wasilkowski, Zhao[08]
3. Novak+W[01] 4. known 5. Sloan +W[98]

What's going on?

These results do not explain why QMC are so efficient

Analogy to Linear Systems: $Ax = b$ with an $n \times n$ matrix A

Large n often implies A sparse

For Multivariate Problems:

Large d often implies f has additional properties

Searching for Additional Properties

Standard spaces of functions are isotropic

That is,

all variables and groups of variables are equally important

In particular, let $g(t) = f(t_{j_1}, t_{j_2}, \dots, t_{j_d})$ – permutation of variables

$$f \in F_d \quad \text{implies that} \quad g \in F_d \quad \text{and} \quad \|g\|_{F_d} = \|f\|_{F_d}$$

Question:

It is really true for large d and practical f ?

Weighted Spaces

For $f \in H(K_d)$

$$\|f\|^2_{H(K_d)} = \sum_{\mathfrak{u}\subseteq[d]} \int_{[0,1]^{|\mathfrak{u}|}} \left(\frac{\partial^{|\mathfrak{u}|}}{\partial x_{\mathfrak{u}}} f(x_{\mathfrak{u}},1)\right)^2 \mathrm{d}x_{\mathfrak{u}} \qquad [d] := \{1,2,\dots,d\}$$

Let $\gamma = \{\gamma_{d,\mathfrak{u}}\}$ **with** $\gamma_{d,\mathfrak{u}} \ge 0$

$$K_{d,\gamma}(x,y) = \sum_{\mathfrak{u}\subseteq[d]} \gamma_{d,\mathfrak{u}} \prod_{j\in\mathfrak{u}} \min(x_j,y_j)$$

For any γ,

$$\|f\|^2_{H(K_{d,\gamma})} = \sum_{\mathfrak{u}\subseteq[d]} \frac{1}{\gamma_{d,\mathfrak{u}}} \int_{[0,1]^{|\mathfrak{u}|}} \left(\frac{\partial^{|\mathfrak{u}|}}{\partial x_{\mathfrak{u}}} f(x_{\mathfrak{u}},1)\right)^2 \mathrm{d}x_{\mathfrak{u}} \qquad \frac{0}{0} = 0$$

For $\gamma_{d,\mathfrak{u}} = 1$ $K_{d,\gamma} = K_d$, as before.

Weighted Spaces

$$f = \sum_{\mathfrak{u} \subseteq [d]} f_{\mathfrak{u}}$$

$f_{\mathfrak{u}}$ orthogonal, $f_{\mathfrak{u}} \in H(K_{\mathfrak{u}})$ with $K_{\mathfrak{u}}(x,y) = \prod_{j \in \mathfrak{u}} \min(x_j, y_j)$

Anova-type decomposition, Kuo, Sloan, Wasilkowski+W [08]

$$\|f\|^2_{H(K_{d,\gamma})} = \sum_{\mathfrak{u} \subseteq [d]} \frac{1}{\gamma_{d,\mathfrak{u}}} \|f_{\mathfrak{u}}\|^2_{H(K_{\mathfrak{u}})} \qquad \frac{0}{0} = 0$$

$f_{\mathfrak{u}}$ depends only on variables in $\mathfrak{u}$

We can model various properties of f for various weights

Various Weights

- **Product weights:** Sloan+W [98], $\gamma_{d,\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_{d,j}$. Then

$$H(K_{d,\gamma}) = H(K_{1,\gamma_{d,1}}) \otimes \cdots \otimes H(K_{1,\gamma_{d,d}})$$

  and $\gamma_{d,j}$ moderates the influence of x_j

- **Finite-order weights:** Dick, Sloan, Wang + W [2003],

$$\gamma_{d,\mathfrak{u}} = 0 \text{ for all } |\mathfrak{u}| > \omega. \text{ Then}$$

$$f = \sum_{\mathfrak{u}\subseteq[d],|\mathfrak{u}|\le\omega} f_u$$

  is a sum of functions depending on at most ω variables.

Tractability

Information Complexity

$$n(\varepsilon, d) \;=\; \min\{\, n \;:\; e(n,d) \,\le\, \varepsilon\, \|I_d\|\,\}$$

$I = \{I_d\}$ is **polynomially tractable** iff

$$n(\varepsilon,d) \,\le\, C\, d^{\,q}\, \varepsilon^{-p} \quad \text{for all} \quad \varepsilon \in (0,1) \quad d = 1, 2, \dots$$

If $q = 0$ then $I = \{I_d\}$ is **strongly polynomially tractable**

$I = \{I_d\}$ is **weakly tractable** iff

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln\, n(\varepsilon,d)}{\varepsilon^{-1} + d} \;=\; 0$$

Other notions of tractability, T-tractability, Gnewuch+W [07,...]

Conditions on Weights

For product weights: $\gamma_{d,\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_{d,j}$

- Strong Pol. Tract. iff $\limsup_d \sum_{j=1}^{d} \gamma_{d,j} < \infty$
- Pol. Tract. iff $\limsup_d \frac{\sum_{j=1}^{d} \gamma_{d,j}}{\ln d} < \infty$
- Weak Tract. iff $\lim_d \frac{\sum_{j=1}^{d} \gamma_{d,j}}{d} = 0$

For finite-order weights: $\gamma_{d,\mathfrak{u}} = 0$ **for all** $|\mathfrak{u}| > \omega$

$$n(\varepsilon, d) \le \frac{\sum_{\mathfrak{u}} \gamma_{d,\mathfrak{u}} 2^{-|\mathfrak{u}|}}{\sum_{\mathfrak{u}} \gamma_{d,\mathfrak{u}} 3^{-|\mathfrak{u}|}} \frac{1}{\varepsilon^2} \le \left(\frac{3}{2}\right)^{\omega} \frac{1}{\varepsilon^2}$$

For all such weights polynomial tractability holds. Proven by Sloan+W [98], Gnewuch +W [07], Sloan, Wang+W [04].

Semi-Constructive Proofs for FoW

Shifted Lattice Rules: $\mathrm{QMC}_{n,d}(f) = \frac{1}{n}\sum_{k=1}^{n} f\left(\left\{\frac{k-1}{n}\,\mathbf{z} + \mathbf{\Delta}\right\}\right)$

$\mathbf{z} \in [n-1]^d$ by the CBC algorithm at cost $\mathcal{O}(n\,d\,\ln\,n)$, Cools and Nuyens[06], and $\mathbf{\Delta} \in [0,1)^d$

Sloan, Wang +W [04]: For some $\mathbf{\Delta}$, the error is ε with

$$n \le C_a\,\varepsilon^{-2/a}\,d^{\,q^*(1-1/a)} \quad \text{for all } a \in [1,2)$$

- $a = 1$ best dependence on ε^{-1} + pol. tract.,
- $a = 2$ strong pol. tract.
- but $\mathbf{z}$ and $\mathbf{\Delta}$ depends on weights

Constructive Proofs for FoW

Low Discrepancy Sequences: $\mathrm{QMC}_{n,d}(f) = \frac{1}{n}\sum_{k=1}^{n} f(t_k)$

Sloan, Wang+ W. [04]: The Niederreiter sequence in base b solves the problem with

$$n \le C_\delta\, \varepsilon^{-(1+\delta)} \left(d^{\,q^*} \log(d+b)\right)^{1+\delta} \quad \forall\, \delta > 0$$

- best dependence on ε^{-1} + pol. tractability
- Niederreiter sequence does **not** depend on weights
- similar results for Halton and Sobol

General Case

$$S_d \,:\, F_d \;\to\; G_d$$

F_d a class of d-variate functions

$$S_d(f) \;\approx\; A_{n,d}(f) = \phi(L_1(f), \dots, L_n(f))$$

L_j arb. linear functionals or function values

Errors in Different Settings

WORST CASE:

$$e(n,d) = \inf_{A_n} \sup_{f \in F_d} \|S_d(f) - A_{n,d}(f)\|$$

AVERAGE CASE:

$$e(n,d) = \inf_{A_n} \int_{F_d} \|S_d(f) - A_{n,d}(f)\| \, \mu(\mathrm{d}f)$$

RANDOMIZED:

$$e(n,d) = \inf_{A_n} \sup_{f \in F_d} \int_{\Omega} \|S_d(f) - A_{n,d,\omega}(f)\| \, \rho(\mathrm{d}\omega)$$

etc. . . .

Tractability

Information Complexity

$$n(\varepsilon,d) \;=\; \min\{\, n : \;\; e(n,d) \,\le\, \varepsilon\, \mathrm{CRI}_d \,\}$$

for the absolute error criterion $\mathrm{CRI}_d = 1$

for the normalized error criterion $\mathrm{CRI}_d = e(0,d) \;=\; \|S_d\|$

Polynomial Tractability

$n(\varepsilon,d) \,\le\, C\, d^{\,q}\, \varepsilon^{-p}$ for all $\varepsilon \in (0,1)$, $d \in \mathbb{N}$

Weak Tractability

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln\, n(\varepsilon,d)}{\varepsilon^{-1}+d} \;=\; 0$$

A Sample of the Results

Linear Unweighted Tensor Product Problems

$$S_d = S_1 \otimes \cdots \otimes S_1 \quad \text{and} \quad S_1 \quad \text{linear}$$

$$F_d = F_1 \otimes \cdots \otimes F_1, \quad G_d = G_1 \otimes \cdots \otimes G_1$$

$\{\lambda_j\}$ ordered eigenvalues of $S_1^* S_1$, $\lambda_2 > 0$

For Worst Case, Normalized Error Criterion, Arbitrary Linear Functionals

$$n(\varepsilon, d) = \left| \left\{ j = [j_1, j_2, \dots, j_d] \, : \, \frac{\lambda_{j_1}}{\lambda_1} \cdots \frac{\lambda_{j_d}}{\lambda_1} > \varepsilon^2 \right\} \right|$$

- no polynomial tractability
- weak tractability iff $\lambda_2 < \lambda_1$ and $\lambda_n = o((\ln n)^{-2})$

Linear Weighted Tensor Product Problems

Let λ_1 be of multiplicity $p \geq 2$. Then

$$n(\varepsilon, d) \ \geq p^d \quad \text{curse of dimensionality}$$

For weighted spaces $F_d = H(K_d)$ and $\lambda_n = o((\ln\, n)^{-2})$

$$K_d(x,y) \ = \sum_{\mathfrak{u}} \gamma_{d,\mathfrak{u}} \prod_{j\in\mathfrak{u}} K_1(x_j,y_j) \qquad 1 \not\in \ H(K_1)$$

$$m_p(\varepsilon,d) \ = \sum_{\mathfrak{u}:\ \gamma_{d,\mathfrak{u}}>\varepsilon^2} (p-1)^{|\mathfrak{u}|}$$

$$\text{weak tractability iff} \lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln m_p(\varepsilon,d)}{\varepsilon^{-1}+d} \ = \ 0$$

Typical Results

Let $\sum_{j=1}^{\infty} \lambda_j^{\tau} < \infty$ for some $\tau > 0$.

- Unweighted problems $\Longrightarrow$ no polynomial tractability or even curse of dimensionality
- For product weights $\gamma_{d,\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_{d,j}$

$$\limsup_{d\to\infty} \sum_{j=1}^{d} \frac{\gamma_{d,j}^{\tau}}{\ln\, d} < \infty \qquad \Longrightarrow \qquad \text{polynomial tractability}$$

$$\lim_{d\to\infty} \sum_{j=1}^{d} \frac{\gamma_{d,j}^{\tau}}{d} = 0 \qquad \Longrightarrow \qquad \text{weak tractability}$$

- Finite-order weights $\gamma_{d,\mathfrak{u}} = 0$ for $|\mathfrak{u}| > \omega \Longrightarrow$ polynomial tractability.

Major Tractability Questions

For which spaces F_d and for which
linear or non-linear multivariate problems
do we have tractability?

Tractability is a popular research subject from 1994.
Many papers and results obtained so far by:

Dick, Fang, Gnewuch, Griebel, Heinrich, Hickernell, Hinrichs, Huang, Joe, Kritzer, Kuo, Larcher, Leobacher, Li, Niederreiter, Novak, Papageorgiou, Pillichshammer, Plaskota, Scheicher, Schmid, Sloan, Wang, Wasilkowski, Werschulz, Wojtaszczyk, Waterhouse, W, Yue, Zhao, Zhang, . . .

Still many open questions...

One of Many Open Problems

Finance integrands do not belong to $H(K_d)$!!!. Usually we have

$$I_d(f) = \int_{[0,1]^d} |f(t)| \, \mathrm{d}t \qquad f \in H(K_d)$$

Conjecture:

Some QMC are as good for f as for $|f|$

Tractability Number

- If you solve k open problems from our book and publish m "good" papers on tractability then your number is

$$2k + m$$

- Let $k = m = 0$. Let p be the largest tractability number among people with whom you publish a "good" paper. Then your number is

$$\frac{p}{1+p}$$

HA: Compute Your Tractability Number

Warning: HA may be intractable !!!