On some aspects of high-dimensional integration and approximation

Mario Ullrich

Linz/Moscow

Seminar on the theory of functions

Moscow, Feb 2022

The Problem

For a function $f : D_d \to \mathbb{R}$, $D_d \subset \mathbb{R}^d$ (often $D_d = [0,1]^d$),
we want to approximate the integral

$$A_n(f) \approx I_d(f) = \int_{D_d} f(y)\,\mathrm{d}y$$

or the function itself

$$A_n(f) \approx f \qquad \text{(in } L_\infty\text{)}$$

using at most n function evaluations of f.

The Problem: errors

For classes of functions $F_d \subset \mathbb{R}^{D_d}$ and algorithm A_n, we want to bound

$$e^{int}(A_n, F_d) := \sup_{f \in F_d} |I_d(f) - A_n(f)|$$

in case of integration,

or

$$e^{app}(A_n, F_d) := \sup_{f \in F_d} ||f - A_n(f)||_\infty$$

for approximation.

(Clearly, one might consider many other problems.)

The Problem: optimal errors

In particular, we would like to know:

$$e_n^{\{int,app\}}(F_d) := \inf_{A_n} e^{\{int,app\}}(A_n, F_d),$$

i.e., the **minimal error** achievable with n function values.

Often, it is better stated by the **information complexity:**

$$n^{\{int,app\}}(\varepsilon, F_d) := \min\left\{n\colon e_n^{\{int,app\}}(F_d) \le \varepsilon\right\}.$$

i.e., the minimal number of function values needed by an optimal algorithm to approximate up to error $\varepsilon > 0$ for all $f \in F_d$.

Remark: linear and non-adaptive algorithms

By results of Bakhvalov and Smolyak, we may restrict ourselves to linear algorithms and non-adaptive sample points, i.e., to algorithms of the form

$$A_n(f) = \sum_{i=1}^{n} a_i\, f(x_i)$$

for some $a_i \in \{\mathbb{R}, L_\infty\}$ and some $x_i \in D_d$, whenever F_d is symmetric and convex. (This might not hold for other problems.)

Moreover, in this case,

$$e_n^{app}(F_d) \;=\; \inf_{x_i} \sup_{\substack{f\in F_d:\\ f(x_1)=\ldots=f(x_n)=0}} \|f\|_\infty .$$

The Problem: tractability and curse of dimension

The problem is **strongly polynomially tractable** iff

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

The problem is **polynomially tractable** iff

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

On the contrary: **curse of dimension** iff there are $c, \varepsilon_0, \gamma > 0$ with

$$n(\varepsilon, F_d) \ge c\,(1+\gamma)^d \quad \text{for all} \quad \varepsilon \le \varepsilon_0 \quad \text{and} \quad d \in \mathbb{N}.$$

"Simple" example: Lipschitz functions

Consider first the univariate Lipschitz class

$$F = \{f : [0,1] \to \mathbb{R} \mid |f(x) - f(y)| \le |x - y|\},$$

we want to compute the integral or f itself by using n function values.

The optimal x_i are equidistant, $x_i = \frac{2i-1}{2n}$.

An optimal algorithm for approximation is the piecewise linear spline, the worst case error is $\frac{1}{2n}$. So, $n^{app}(\varepsilon, F) = \lceil 1/(2\,\varepsilon) \rceil$.

Similarly, $\quad n^{int}(\varepsilon, F) = \lceil 1/(4\,\varepsilon) \rceil$.

"Simple" example: Lipschitz functions II

$$F_d = \left\{ f \in C([0,1]^d) \colon \; |f(x) - f(y)| \le \|x - y\|_\infty \right\}$$

Theorem (Maung Zho Newn and Sharygin 1971)

$$e_n^{int}(F_d) \;=\; \frac{d}{2d+2} \cdot n^{-1/d}$$

for $n = m^d$ with $m \in \mathbb{N}$.

Observe that, for $n = 2^d$, we have $e_{2^d}^{int}(F_d) = \frac{1}{2}\, e_0^{int}(F_d)$.

$\Longrightarrow$ Curse of dimension

(Similar results: Babenko, Sukharev, Chernaya)

Another example: Smooth functions

We now consider the classical classes of functions

$$\mathcal{C}_d^k = \left\{ f \in C^k([0,1]^d) \colon \|D^\beta f\|_\infty \le 1 \ \text{ for all } \ \beta \in \mathbb{N}_0^d \ \text{ with } \ |\beta|_1 \le k \right\}$$

with the norm $\|f\|_{\mathcal{C}_d^k} := \max_{\beta\colon |\beta|_1 \le k} \|D^\beta f\|_\infty$.

Or the class of all (directional) bounded derivatives

$$\widetilde{\mathcal{C}}_d^k = \left\{ f \in C^k([0,1]^d) \colon \|D^\beta f\|_\infty \le 1 \ \text{ for all } \ \beta \in \mathbb{N}_0 \ \text{ with } \ \beta \le k \right\}$$

with the norm $\|f\|_{\widetilde{\mathcal{C}}_d^k} := \max_{\beta \le k} \|D^\beta f\|_\infty$.

Optimal order of convergence

Bakhvalov '59, Tikhomirov '60: For some $c_{k,d}, C_{k,d} > 0$,

$$\left(\frac{c_{k,d}}{\varepsilon}\right)^{d/k} \leq n^{\{int,app\}}(\varepsilon, \mathcal{C}_d^k) \leq \left(\frac{C_{k,d}}{\varepsilon}\right)^{d/k}$$

and the upper bounds are achieved by product rules, i.e., by function evaluations on (or close to) the full grid

$$G_m := \left\{\left(\frac{i_1}{m}, \dots, \frac{i_d}{m}\right) : \ i_\ell \in \{0, \dots, m-1\}\right\}, \qquad m = \lfloor n^{1/d} \rfloor.$$

But: The limited knowledge about $c_{k,d}$ and $C_{k,d}$ does not lead to any tractability statement.

Explicit-in-dimension bounds for integration

Theorem (HNUW '16 & '17)

For all $k \in \mathbb{N}$ there exist constants $c_k, C_k > 0$ such that for all $d \in \mathbb{N}$, we have

$$\min\left\{\tfrac{1}{2}, c_k\, d\, n^{-k/d}\right\} \;\le\; e_n^{int}(\mathcal{C}_d^k) \;\le\; \min\left\{1, C_k\, d\, n^{-k/d}\right\}.$$

or

$$\left(\frac{c_k d}{\varepsilon}\right)^{d/k} \;\le\; n^{int}(\varepsilon, \mathcal{C}_d^k) \;\le\; \left(\frac{C_k d}{\varepsilon}\right)^{d/k}.$$

$\Longrightarrow$ Curse of dimension for every $k \in \mathbb{N}$, even with super-exponential lower bound. Upper bound attained by product rule.

Proof of upper bound

Haber 1970: For the product rule

$$Q_m^d(f) = \sum_{i_1=0}^{m-1} \cdots \sum_{i_d=0}^{m-1} a_{i_1} \cdots a_{i_d} \cdot f(t_{i_1}, \ldots, t_{i_d}),$$

based on the one-dimensional quadrature rule

$$Q_m(f) = \sum_{i=0}^{m-1} a_i f(t_i)$$

we have

$$e(Q_m^d, \mathcal{C}_d^k) \leq \left(\sum_{\ell=0}^{d-1} A^\ell \right) \cdot e(Q_m, \mathcal{C}_1^k),$$

where $A = \sum_{i=1}^m |a_i|$. There are "good" rules with $A = 1$. □

Proof of lower bound

For every set $\mathcal{P} \subset [0,1]^d$ with $\#\mathcal{P} = n$ we will find a **fooling function**, i.e., a function with $f(x) = 0$ for $x \in \mathcal{P}$, and large integral.

For $r > 0$ we first define the initial function

$$h_r(x) = \begin{cases} 0, & \text{if } \min_{y \in \mathcal{P}_n} \|x-y\|_1 \le r, \\ 1, & \text{otherwise.} \end{cases}$$

Clearly,

$$\int_{[0,1]^d} h_r(x)\,\mathrm{d}x = 1 - \mathrm{vol}_d(\mathcal{P}_r) \ge 1 - n\,\mathrm{vol}_d(rB_1^d) = 1 - n\,\frac{(2r)^d}{d!}.$$

Smoothing the function

The initial function h_r is not in $\mathcal{C}_d^k$. We use convolution to make it differentiable.

Let

$$g_{r,k}(x) \;=\; \frac{1}{\mathrm{vol}_d(\varrho_r B_1^d)} \begin{cases} 1, & \text{if } \|x\|_1 \le \frac{r}{k+1}, \\ 0, & \text{otherwise}, \end{cases}$$

and define the fooling function

$$f_r \;:=\; h_r * \underbrace{g_{r,k} * \cdots * g_{r,k}}_{(k+1)\text{-fold}} .$$

We have $f_r(x) = 0$ for $x \in \mathcal{P}$ and $\int f_r \,\mathrm{d}x = \int h_r \,\mathrm{d}x$.

Bounding the derivatives

For a continuous function f it is easy to prove that

$$\|D^{e_i}[f * g_{r,k}]\|_\infty \leq \frac{d(k+1)}{r}\,\|f\|_\infty$$

Inductively, we obtain

$$\|f_r\|_{\mathcal{C}_d^k} \leq \max\left\{1,\, r^{-k}\big(d(k+1)\big)^{k-1}\right\}$$

Finishing the proof

We define

$$f_r^* = \frac{f_r}{\|f_r\|_{\mathcal{C}_d^k}} \in \mathcal{C}_d^k.$$

Using

$$\int_{[0,1]^d} f_r(x)\,\mathrm{d}x \geq 1 - n\,\frac{(2r)^d}{d!} > 1 - n\left(\frac{4er}{d}\right)^d$$

we obtain that $\int_{[0,1]^d} f_r^*(x)\,\mathrm{d}x \leq \varepsilon$ implies that

$$n \geq \left(1 - \varepsilon \cdot \|f_r\|_{\mathcal{C}_d^k}\right) \left(\frac{d}{4er}\right)^d.$$

Finishing the proof

$$n \geq \left(1 - \varepsilon \cdot \|f_r\|_{\mathcal{C}_d^k}\right) \left(\frac{d}{4er}\right)^d .$$

Now we choose

$$r = (2\varepsilon)^{1/r} \big(d(k+1)\big)^{1-1/k}$$

to obtain

$$\|f_r\|_{\mathcal{C}_d^k} \leq \frac{1}{2\varepsilon}$$

and

$$n(\varepsilon, \mathcal{C}_d^k) \geq \frac{1}{2} \left(\frac{c_k\, d}{\varepsilon}\right)^{d/r}$$

with $c_k = 1/((4e)^k\,(k+1)^{k-1})$.

Curse of dimension for integration for $\widetilde{\mathcal{C}}_d^k$

Partial derivatives: $$n(\varepsilon, \mathcal{C}_d^k) \approx \left(\frac{d}{\varepsilon}\right)^{d/k}.$$

Theorem (HNUW '16 & '17)

For all $k \in \mathbb{N}$ there exist constants $c_k, C_k > 0$ such that for all $d \in \mathbb{N}$, we have

$$\min(1/2,\, c_k d^{1/2} n^{-k/d}) \leq e_n(\widetilde{\mathcal{C}}_d^k) \leq \min(1,\, \tilde{c}_k d n^{-k/d}).$$

or

$$\left(\frac{c_k \sqrt{d}}{\varepsilon}\right)^{d/k} \leq n^{int}(\varepsilon, \widetilde{\mathcal{C}}_d^k) \leq \left(\frac{C_k d}{\varepsilon}\right)^{d/k}.$$

Open Problems

Open Problem 1: directional derivatives

What is the dependence of $n^{int}(\varepsilon_0, \widetilde{\mathcal{C}}_d^k)$ on d?

Open Problem 2: arbitrary domains

The lower bounds hold for arbitrary $D_d \subset \mathbb{R}^d$ with $\mathrm{vol}(D_d) \approx 1$.

Verify that $n^{int}(\varepsilon_0, \mathcal{C}^k(D_d)) \lesssim \left(\frac{d}{\varepsilon}\right)^{d/k}$ for arbitrary domains D_d?

Open Problem 3: more arbitrary domains

One can even ask if the asymptotic constant

$$\limsup_{n\to\infty} e_n^{int}(C^k(D_d)) \cdot n^{k/d}$$

depend on D_d (except volume)? Is it always a lim? (True for $k = 1$.)

Infinite smoothness

Surprisingly, we don't know much about infinite smoothness:

$$\mathcal{C}_d^\infty = \left\{ f \in [0,1]^d \to \mathbb{R} \colon \|D^\beta f\|_\infty \le 1 \text{ for all } \beta \in \mathbb{N}_0^d \right\}$$

Note that it is often stated that integration in $\mathcal{C}_d^\infty$ is "easy" since

$$e_n^{int}(\mathcal{C}_d^\infty) \lesssim_{r,d} n^{-r} \qquad \text{for all } r > 0.$$

But what about high dimensions?

The best lower bound (unpublished) is linear, i.e., $n^{int}(\varepsilon, \mathcal{C}_d^\infty) \ge c \cdot d$ for small ε while all known upper bounds are exponential in d.

Infinite smoothness II

We only know: $d \lesssim n(\varepsilon, \mathcal{C}_d^\infty) \lesssim c^d.$

(Important) Open Problem 4: infinte smoothness

Is there a constant $\gamma > 0$ such that for some $\varepsilon_0 \in (0,1)$ we have

$$n^{int}(\varepsilon_0, \mathcal{C}_d^\infty) \geq (1+\gamma)^d?$$

For $\widetilde{C}_d^\infty$ we know (HNUW '14) that the problem is *weakly tractable*:

$$\lim_{\mathrm{e}^{-1}+d\to\infty} \frac{\ln\, n(\varepsilon, \widetilde{C}_d^\infty)}{\mathrm{e}^{-1}+d} = 0\,.$$

Curse of dimension for approximation

Let us turn approximation.

NW '09: For every $\varepsilon \in (0,1)$ we have

$$n^{app}(\varepsilon, \mathcal{C}_d^\infty) \geq 2^{\lfloor d/2 \rfloor}.$$

HNUW '17: There exists $c_k > 0$ such that for all $d \in \mathbb{N}$ and $\varepsilon \in (0,1)$ we have

$$n^{app}(\varepsilon, \mathcal{C}_d^k) \geq n^{int}(\varepsilon, \mathcal{C}_d^k) \geq \left(\frac{c_k d}{\varepsilon}\right)^{d/k}.$$

When is approximation more difficult than integration?

Curse of dimension for approximation II

Theorem (Krieg '19)

For all $k \in \mathbb{N}$ there exist constants $c_k, C_k > 0$ such that for all $d \in \mathbb{N}$ and $\varepsilon \in (0, 1/2)$, we have

$$\left(\frac{c_k d^{k/2}}{\varepsilon}\right)^{d/k} \leq n^{app}(\varepsilon, \mathcal{C}_d^k) \leq \left(\frac{C_k d^{k/2}}{\varepsilon}\right)^{d/k}, \quad \text{if } k \text{ even,}$$

and

$$\left(\frac{c_k d^{k/2}}{\varepsilon}\right)^{d/k} \leq n^{app}(\varepsilon, \mathcal{C}_d^k) \leq \left(\frac{C_k d^{(k+1)/2}}{\varepsilon}\right)^{d/k}, \quad \text{if } k \text{ odd.}$$

Approximation is essentially harder than integration iff $k \geq 3$.

Curse of dimension for approximation III

Again, we know a bit more for directional derivatives:

Theorem (Krieg '19)

For all $k \in \mathbb{N}$ there exist constants $c_k, C_k > 0$ such that for all $d \in \mathbb{N}$ and $\varepsilon \in (0, 1/2)$, we have

$$\left(\frac{c_k d^{k/2}}{\varepsilon}\right)^{d/k} \le n^{app}(\varepsilon, \widetilde{\mathcal{C}}_d^k) \le \left(\frac{C_k d^{k/2}}{\varepsilon}\right)^{d/k}.$$

Approximation is essentially harder than integration:

yes, for $k \ge 3$; no, for $k = 1$; unclear, for $k = 2$.

Open Problems

Open Problem 5: partial derivatives

What is the precise dependence of $n^{app}(\varepsilon, \mathcal{C}_d^k)$ on d for odd k?

Open Problem 6: arbitrary domains

The lower bounds hold for arbitrary $D_d \subset \mathbb{R}^d$.

What about arbitrary domains D_d?

Open Problem 7: infinite smoothness

$n^{app}(\varepsilon, \mathcal{C}_d^k) \asymp_{k,\varepsilon} d^{d/2}$, but 'only' $n^{app}(\varepsilon, \mathcal{C}_d^\infty) \geq 2^{\lfloor d/2 \rfloor}$.

Is there a constant $\gamma > 0$ such that for some $\varepsilon_0 \in (0,1)$ we have

$$n^{app}(\varepsilon_0, \mathcal{C}_d^\infty) \geq (1+\gamma)^{\Omega(d \log d)}?$$

L_2-Approximation

Let us shortly discuss L_2-approximation.

We are interested in the L_2-**sampling numbers**

$$g_n(F) := \inf_{\substack{x_1,\dots,x_n\in D\\ \varphi_1,\dots,\varphi_n\in L_2}} \sup_{f\in F} \left\| f - \sum_{i=1}^n f(x_i)\,\varphi_i \right\|_{L_2},$$

i.e., the minimal error that can be achieved with n function values.

We compare that with the **Gelfand width**

$$c_n(F) := \inf_{\substack{\varphi\colon \mathbb{R}^n\to L_2\\ N\in(F')^n}} \sup_{f\in F} \left\| f - \varphi\circ N(f) \right\|_{L_2},$$

A comparison

We proved this general result on the **power of function values**.

Theorem [Krieg/U 2019; U 2020; Krieg/U 2021]

Let $F \hookrightarrow L_2$ be a separable metric space of functions on D, such that point evaluation is continuous on F.

Then, for every $0 < p < 2$, there is a constant $c_p > 0$, depending only on p, such that, for all $n \geq 2$, we have

$$g_N(F) \leq \sqrt{\log n} \left(\frac{1}{n} \sum_{k \geq n} \left(\sqrt{k} \cdot c_k(F) \right)^p \right)^{1/p}$$

for $N \geq c_p \cdot n$.

Information complexities

We define

$$n^{L_2}(\varepsilon, F) \;:=\; \min\left\{n\colon\, g_n(F) \le \varepsilon\right\}$$

and

$$n^{L_2,all}(\varepsilon, F) \;:=\; \min\left\{n\colon\, c_n(F) \le \varepsilon\right\}.$$

L_2-approximation: polynomial convergence

Theorem (Krieg/U/Wozniakowski '22)

Assume that

$$n^{L_2,all}(\varepsilon, F_d) \leq C\, d^q\, \varepsilon^{-p}$$

for some $q > 0$, $p < 2$ and all $\varepsilon \in (0,1)$. Then

$$n^{L_2}(\varepsilon, F_d) \leq D\, d^q\, \varepsilon^{-p} \log(d/\varepsilon)^s$$

for all $\varepsilon \in (0,1)$ and $d \in \mathbb{N}$, and some $D, s > 0$ that depends only on C, p and q.

This was known (for Hilbert spaces?) up to optimal exponents, see books of Novak and Wozniakowski.

L_2-approximation: exponential convergence

Theorem (Krieg/U/Wozniakowski '22)

Assume that

$$n^{L_2,all}(\varepsilon, F_d) \leq C\, d^q\, (1 + \ln \varepsilon^{-1})^p$$

for some $p, q > 0$ and all $\varepsilon \in (0,1)$. Then

$$n^{L_2}(\varepsilon, F_d) \leq D\, d^q\, (\ln d)^p\, (1 + \ln \varepsilon^{-1})^p$$

for all $\varepsilon \in (0,1)$ and $d \in \mathbb{N}$, and some $D > 0$ that depends only on C, p and q.

Thank you!