How difficult is it to compute the L_q norm of a function

Stefan Heinrich

University of Kaiserslautern

Let $d \in \mathbb{N}$, $Q = [0,1]^d$, $r \in \mathbb{N}_0$, $1 \le p, q \le \infty$. We assume that $r/d > 1/p$, so $J : W_p^r(Q) \to L_q(Q)$ is continuous and function values are well-defined.

We want to approximate the (nonlinear) operator $S_q : W_p^r(Q) \to \mathbb{R}$, given by

$$S_q(f) \;=\; \|f\|_{L_q(Q)} = \begin{cases} \left(\int_Q |f(x)|^q dx \right)^{1/q} & \text{if} \quad 1 \le q < \infty \\ \operatorname{ess\,sup}_{x \in Q} |f(x)| & \text{if} \quad q = \infty, \end{cases}$$

that is, we seek to approximate/compute the L_q norm of a function. We consider standard information, that is, values of f.

Complexity: Determine the smallest error among all deterministic/ randomized algorithms that use not more than n function values

Let $n \in \mathbb{N}$, G a Banach space, $S : W_p^r(Q) \to G$ any mapping.

deterministic n-th minimal error:

$$e_n^{\mathrm{det}}(S) = \inf_{x_1,\dots,x_n \in Q,\, \varphi:\mathbb{R}^n \to G} \; \sup_{f \in B_{W_p^r(Q)}} \|S(f) - \varphi(f(x_1),\dots,f(x_n))\|_G$$

randomized n-th minimal error:

$$e_n^{\mathrm{ran}}(S) = \inf \sup_{f \in B_{W_p^r(Q)}} \mathbb{E}\,\|S(f) - \varphi_\omega(f(x_{1,\omega}),\dots,f(x_{n,\omega}))\|_G$$

the infimum taken over all tuples

$$\Big((\Omega,\Sigma,\mathbb{P}),(x_{1,\omega}\dots,x_{n,\omega})_{\omega\in\Omega},(\varphi_\omega)_{\omega\in\Omega}\Big)$$

such that $x_{1,\omega}\dots,x_{n,\omega} \in Q$, $\varphi_\omega : \mathbb{R}^n \to G$, and for all $f \in W_p^r(Q)$

$$\omega \to \varphi_\omega(f(x_{1,\omega}),\dots,f(x_{n,\omega}))$$

is Σ-to-Borel measurable and essentially separably valued

Theorem 1. *(Wasilkowski, 1984)*

$$e_n^{\mathrm{det}}(S_q, B_{W_p^r(Q)}) \asymp n^{-r/d+\max(1/p-1/q,0)}.$$

G. W. Wasilkowski, Some nonlinear problems are as easy as the approximation problem, Comp. & Maths. with Appls. 10 (1984), 351-363

$$\begin{aligned} e_n^{\mathrm{det}}(S_q, B_{W_p^r(Q)}) &\asymp e_n^{\mathrm{det}}(J, B_{W_p^r(Q)}, L_q(Q)) \\ &\asymp n^{-r/d+\max(1/p-1/q,0)} \end{aligned}$$

Theorem 2. *(H., 2018)*

$$e_n^{\text{ran}}(S_q, B_{W_p^r(Q)}) \asymp n^{-r/d+\max(1/p-1/q,-1/2)}.$$

Comparisons:

$$e_n^{\text{ran}}(J, B_{W_p^r(Q)}, L_q(Q)) \quad \asymp \quad n^{-r/d+\max(1/p-1/q,0)}$$

the same rate as $e_n^{\text{ran}}(S_q, B_{W_p^r(Q)})$ iff $p \le q$

$I : W_p^r(Q) \to \mathbb{R}$ integration, $If = \int_Q f(x)dx$

$$e_n^{\text{ran}}(I, B_{W_p^r(Q)}) \asymp n^{-r/d-\max(1/p-1,-1/2)}$$

the same rate as $e_n^{\text{ran}}(S_q, B_{W_p^r(Q)})$ iff $q = 1$ or $1/p - 1/q \le -1/2$.

Direct computation: Approximate $\int_Q |f(x)|^q dx$, then take the q-th root

Problems with the direct computation:

absolute value $|f|$ or fractional q may spoil the smoothness

But even for $q = 2k$, $k \in \mathbb{N}$ the power $|f|^q$ reduces integrability

and finally, there is still the (only Hölder continuous) q-th root to take

The algorithms:

Fix $0 < \delta < 1$, let $Pg = \sum_{j=1}^{\kappa} g(z_j)\psi_j \quad (g \in C(Q))$ be tensor product Lagrange interpolation operator of degree $\max(r-1, 0)$ w.r. to the uniform grid $(z_j)_{j=1}^{\kappa}$ on $[0, 1-\delta]^d$, $(\psi_j)_{j=1}^{\kappa}$ the respective Lagrange polynomials

Let $\theta = \theta(\omega)$ be a uniformly distributed on $[0, \delta]^d$ random variable on some probability space $(\Omega, \Sigma, \mathbb{P})$. For $\omega \in \Omega$ define

$$(P_\omega g)(x) = \sum_{j=1}^{\kappa} g(z_j + \theta)\psi_j(x - \theta) \quad (x \in Q)$$

Let $l \in \mathbb{N}_0$, $(Q_i)_{i=1}^{2^{dl}}$ partition of Q into $2^{dl} \asymp n$ cubes of sidelength 2^{-l} and of disjoint interior, x_i the point in Q_i with minimal coordinates. Define

$$(P_{l,\omega} g)(x) = \sum_{j=1}^{\kappa} g(x_i + 2^{-l}(z_j + \theta))\psi_j(2^l(x - x_i) - \theta) \quad (x \in Q_i)$$

– composite tensor product Lagrange interpolation on a randomly perturbed grid

We approximate $S_q(f) = \|f\|_{L_q(Q)}$ by three algorithms:

$$A_{1,\omega}(f) = \|P_{l,\omega} f\|_{L_q(Q)}$$

Deterministic setting: $A_{1,\omega}(f)$ for any fixed $\omega \in [0, \delta]^d$, e.g. $\theta(\omega) = 0$.

Randomized setting, case $1 \le p \le q \le \infty$: $A_{1,\omega}(f)$

Randomized setting, case $1 \le q < p \le \infty$:

Let $\xi_i = \xi_i(\omega)$ $(i = 1, \dots, n)$ be independent (also of θ), uniformly distributed on Q random variables on $(\Omega, \Sigma, \mathbb{P})$. Now put

$$A'_{2,\omega}(f) = \int_Q |(P_{l,\omega} f)(x)|^q dx + \frac{1}{n} \sum_{i=1}^{n} \Big(|f(\xi_i(\omega))|^q - |(P_{l,\omega} f)(\xi_i(\omega))|^q \Big)$$

Monte Carlo with separation of the main part / control variate

$$A_{2,\omega}(f) = \left| A'_{2,\omega}(f) \right|^{1/q}$$

For $r = 0$ we can drop the main part.

$$A_{3,\omega}(f) \;=\; \left(\frac{1}{n} \sum_{i=1}^{n} |f(\xi_i(\omega))|^q \right)^{1/q} .$$

Upper bound:

What is different from the usual estimate?

$p = \infty$ (say even $F = B_{C^r(Q)}$), $q = 2$, claimed rate $n^{-r/d-1/2}$

$$\begin{aligned}
&\mathbb{E}\,|S_q(f) - A_\omega(f)|^2 \le \ldots\ldots \\
&\le \; cn^{-1-2r/d}\|f\|_{L_2(Q)}^{-2}\left(\|f\|_{L_2(Q)}^2 + \underbrace{\mathbb{E}_\theta\|P_{l,\omega}f\|_{L_2(Q)}^2}_{\le c\|f\|_{L_2(Q)}^2}\right) \\
&\le \; cn^{-1-2r/d}
\end{aligned}$$

A deterministic interpolation operator P_l would not work:

$$\begin{aligned}
&\le \; cn^{-1-2r/d}\|f\|_{L_2(Q)}^{-2}\left(\|f\|_{L_2(Q)}^2 + \|P_l f\|_{L_2(Q)}^2\right) \\
&\le \; cn^{-1-2r/d}\left(1 + \frac{\|P_l f\|_{L_2(Q)}^2}{\|f\|_{L_2(Q)}^2}\right)
\end{aligned}$$

Lower bound:

Through Bakhvalov's average case technique:

w_i smooth bump functions, shrunk to Q_i

usually:

$$\nu \quad \text{distribution of} \quad n^{-r/d} \sum_{i=1}^{n} \varepsilon_i(\omega) w_i$$

(does not work, since $\left\| \sum_{i=1}^{n} \varepsilon_i(\omega) w_i \right\|_{L_q(Q)}$ is the same for all values of ε_i)

here:

$$\nu \quad \text{distribution of} \quad n^{-r/d} \left(1 + \sum_{i=1}^{n} \varepsilon_i(\omega) w_i \right)^{1/q}$$

Arithmetic cost

So far n-th minimal errors, i.e., the number of information functionals (function values) is bounded by n.

Question: implementation using a small number of arithmetic operations, say, $\mathcal{O}(n)$ or $\mathcal{O}(n(\log n)^\alpha)$ for some $\alpha > 0$

Real number model of computation: addition, subtraction, multiplication, division, comparison of real numbers, elementary functions $\ln x$ $(x > 0)$ and $\exp(x)$ $(x \in \mathbb{R})$ (all carried out exactly).

random number generator: produces at the k-th call a realization of the k-th element of a sequence of (ideal) independent, uniformly distributed on $[0, 1]$ random variables

each operation - cost 1.

What do we have to implement?

$$\int_Q |(P_{l,\theta(\omega)} f)(x)|^q dx = \sum_{i=1}^{2^{dl}} \int_{Q_i} |\zeta_i(x)|^q dx \quad (q < \infty)$$

$$\operatorname{ess\,sup}_{x \in Q} |(P_{l,\theta(\omega)} f)(x)| = \max_{1 \le i \le 2^{dl}} \max_{x \in Q_i} |\zeta_i(x)| \quad (q = \infty),$$

where $Q_i = \left[\frac{i-1}{2^l}, \frac{i}{2^l}\right]$ and

$$\zeta_i(x) = \sum_{j=1}^{\kappa} f(x_i + 2^{-l}(z_j + \theta(\omega)))\psi_j(2^l(x - x_i) - \theta(\omega)),$$

If $r = 0$ use algorithm $A_{3,\omega}$: $\mathcal{O}(n)$ operations

If $r = 1$, the ζ_i are constant – algorithms $A_{1,\omega}$ and $A_{2,\omega}$ need $\mathcal{O}(n \log n)$ operations

$r > 1$ one has to compute $\int_Q |(P_{l,\omega}f)(x)|^q dx$, i.e.

$\int_{Q_i} |\zeta_i(x)|^q dx$ for $2^{dl} \asymp n$ (multivariate) polynomials ζ_i of degree depending only on r

If q is an even integer, $|\zeta_i(x)|^q$ is a polynomial –
each of the $2^{dl} \asymp n$ integrals can be compute exactly in $\mathcal{O}(1)$ operations

Now we settle the case $d = 1$, $r \in \mathbb{N}$, $1 \leq q \leq \infty$. For a polynomial of one variable $\zeta(x)$ of degree m compute $\int_a^b |\zeta(x)|^q dx$

Step 1: use bisection to include the zeros into small open intervals – recursively, for $\zeta^{(m-1)}$, $\zeta^{(m-2)}$,$\ldots \zeta'$, ζ.

Remove these intervals – the remaining set is a union of at most 2^m disjoint closed intervals.

$q \in 2\mathbb{N}+1$: integrate $|\zeta(x)|$ over these intervals
($|\zeta(x)|^q$ is a polynomial on each of them)

$q = \infty$: take the maximum of $|\zeta(x)|$ over the endpoints of these intervals
($\zeta(x)$ is monotone on each interval)

$1 < q < \infty$, $q \neq \mathbb{N}$,

Step 2: Use the intervals from Step 1, i.e. $\zeta(x)$ is monotone and does not change sign on, say $[c, d]$

Expand z^q on on the interval $[2^{k-1}, 2^k]$ $(k \in \mathbb{Z})$ at the midpoint $z_k = 3 \cdot 2^{k-2}$ into its Taylor series, cut at term $N \in \mathbb{N}$:

$$\pi_{k,N}(z) = \sum_{n=0}^{N} \frac{q(q-1)\dots(q-n+1)}{n!} z_k^{q-n}(z - z_k)^n.$$

Include all x_k with $c < x_{k_0} < \cdots < x_{k_1} < d$ and

$$\zeta(x_k) = 2^k \quad (k_0 \leq k \leq k_1)$$

into small intervals, using bisection again, and remove these intervals. Then integrate (the polynomial) $\pi_{k_0-1,N}(\zeta(x))$ over each interval in the remaining set

Repeat steps 1 and 2 for ζ_i on $\left[\frac{i-1}{2^l}, \frac{i}{2^l}\right]$, $1 \leq i \leq 2^l \quad (\asymp n$ intervals),
$n(\log n)^3$ operations, optimal order of the error is preserved

Extensions, open problems:

1. Analogous results hold for bounded Lipschitz domains Q.
deterministic case: Novak, Triebel 2006,
stochastic case: different approximation operators $P_{l,\omega}$ (H. 2008)

2. For $r = 0$, $1 \leq q < p \leq \infty$ the upper bounds hold for arbitrary probability spaces $(Q, \mathcal{Q}, \varrho)$
– a matching lower bound holds if the probability space has some structure

3. Let $s \in \mathbb{N}_0$, $s \le r$, and assume that $W_p^r(Q)$ is embedded in $W_q^s(Q)$. Moreover, the algorithm can also use values of partial derivatives $D^\beta f(x)$ for $|\beta| \le s$.
Compute $S_q^{(s)}(f) = \|f\|_{W_q^s(Q)}$

$$e_n^{\text{det}}(S_q^{(s)}, B_{W_p^r(Q)}) \asymp n^{-(r-s)/d+\max(1/p-1/q,0)}$$

$$e_n^{\text{ran}}(S_q^{(s)}, B_{W_p^r(Q)}) \prec n^{-(r-s)/d+\max(1/p-1/q,-1/2)}$$

Lower bound: open

Open: Compute the norms of other function spaces
(Besov spaces, Triebel-Lizorkin spaces, ...)

Open: the randomized setting for linear information

Open: the arithmetic complexity of norm computation for $d > 1$